Norm-Referenced Versus Curriculum-Based Assessment: A Balanced Perspective

By Stefan C. Dombrowski

In anticipation of the 2003 IDEA reauthorization, there has been increased discussion about how to effectively enhance the process of identifying learning disabilities. Recent roundtable discussions (President's Commission on Excellence in Special Education, 2002; U.S. Office of Special Education [NJCLD] Roundtable Report, 2002) and articles (e.g., Aaron, 1997; Pasternack as cited in Harrison, 2002; Sternberg & Grigorenko, 2002; Bradley, Danielson & Hallahan, 2002) have emphasized the importance of basing LD policy decisions on effective diagnostic practices.

One of the more significant outcomes of these discussions is the recommendation to abandon mandated IQ testing and the IQ-achievement discrepancy approach for LD diagnosis. Although it is unclear whether this change will be included in the next reauthorization of IDEA (legislation is pending as of this writing), the field seems justified in taking this position when considering the preponderance of accumulated empirical, theoretical, and practical evidence against the current discrepancy model for learning disabilities classification (Aaron, 1997; Harrison, 2002; Kavale & Forness, 1984; Siegel, 1999; Sternberg & Grigorenko, 2002; Vellutino, Scanlon, & Lyons, 2000). A corollary of discarding the discrepancy approach is to recommend the elimination of intelligence tests for LD classification. There may be some resistance to this recommendation, given the long-standing tradition of using of intelligence tests in the LD identification process (Kirk, 1981). However, research indicates that the use of intellectual (i.e., cognitive) assessment in its traditional form via a discrepancy model has limited utility for LD diagnostic and intervention purposes (Siegel, 1989; Stanovich, 1991) and should be abandoned (Naglieri & Reardon, 1993).

Instead of the discrepancy model, leaders in the field recently proposed a comprehensive model based on the collection of data from a variety of sources and assessment approaches (Commission on Excellence in Special Education, 2002; OSEP, 2002). Considering this perspective, it will be important to review both the strengths and limitations of the two most common assessment methods used in the evaluation of learning problems: norm-referenced cognitive/achievement assessment and curriculum-based assessment models. There has been both widespread support for and an abundance of research documenting the utility of curriculum-based measures for LD intervention purposes (Deno, Fuchs, Marston, & Shinn, 2001; Gresham, 2001; Shinn, 1998). Of late, much of this research has focused on Curriculum-Based Measurement (CBM) methodology. On the other hand, there has been a relative lack of current research that discusses the utility of norm-referenced cognitive and achievement assessment for LD eligibility and intervention decisions. In light of this paucity, and given recent proposals to change how the field diagnoses learning disabilities, there is need for a balanced review of the merits and limitations of each approach.

Both norm-referenced assessment and functionally related academic assessment measures such as curriculum-based assessment (CBA) provide important information. Accordingly, the two assessment approaches should be thought of as interacting synergistically, rather than antagonistically. While it is increasingly evident that curriculum-based assessment should play a more formal role in LD identification and intervention decisions (Deno, 1985; Shapiro, 1990; Shinn, 1998), there is also evidence for the utility of norm-referenced cognitive and achievement assessment in this process (Mather & Roberts, 1994).
Strengths of Norm-Referenced Assessment

When reference is made to “norm-referenced assessment,” this refers to both norm-referenced cognitive and achievement assessment, unless otherwise specified. Further, it is important to be aware that there is considerable overlap between cognitive and academic abilities (Carroll, 1993; Flanagan, Ortiz, Alfonso & Mascaro, 2002), leading to a blurring of traditional cognitive/achievement assessment boundaries. For example, some abilities that were traditionally considered “achievement” are now included within cognitive assessment batteries (e.g., phonemic awareness; quantitative ability), while others that were traditionally considered “intelligence” may be found on achievement test batteries (e.g., expressive vocabulary; auditory discrimination) (Flanagan, 2000).

One of the more significant features of norm-referenced cognitive/achievement assessment is that it may be useful for the early detection of children who are at risk for learning difficulties. Recent research points to the existence of cognitive co-morbidities and achievement patterns that often accompany children’s learning difficulties (Fletcher et al., 2002; Gottardo, Stanovich, & Siegel, 1996; Geva & Siegel, 2001; Hammill, Mather, Allen & Roberts, 2002; Siegel, 1992, 1999). The existence of these conditions might serve as markers for the presence of learning disabilities and become important in early identification, should accumulating research evidence continue to substantiate their relevance (Stanovich & Siegel, 1994). For instance, Siegel (1999) describes two potential LD profiles that have emerged from cognitive co-morbidity research. The first, associated with reading, writing and spelling difficulties, includes cognitive processing deficits in phonological/phonemic awareness and short-term memory. The second profile is associated with difficulties with writing and arithmetic in what may be described as an output failure (Siegel, 1999). Within the second LD profile, individuals have solid oral language skills, some difficulty with spelling, and cognitive processing deficits in fine motor coordination, visual-spatial processing, and short and long term memory (Siegel, 1999). Lower standardized performance in these cognitive ability and achievement areas might serve as markers for the presence of learning disabilities that might otherwise go undetected (or undiagnosed under present diagnostic nosology) until later grades when a child begins to demonstrate failure on actual tests of reading, writing, mathematics and spelling achievement.

The early recognition of these cognitive co-morbidities will place educators and clinicians in a better position to remediate learning difficulties at an early age through proactive intervention (Stanovich & Siegel, 1994). Thus, LD profile research using standardized assessment should be encouraged, as it may continue to uncover cognitive/achievement patterns that will be useful for early identification and remediation purposes.

Measures of verbal IQ are also potentially useful for the early detection of future learning difficulties. Verbal IQ measures are predictive of academic success with correlations in the 0.50 to 0.75 range (Anastasi & Urbina, 1997; Kaufman & Lichtenberger, 2002). Thus, the assessment of verbal IQ might be utilized, for instance, as a screener for the detection of future academic problems among children in the early stages of literacy acquisition.

Norm-referenced assessment also provides for expedient (i.e., a few hours) understanding of a child’s profile of cognitive/academic strengths and weakness (Flanagan, Ortiz, Alfonso, & Mascaro, 2002). Such assessment may be useful, for example, when a child transfers from another school without academic records and is suspected of having learning difficulties. Data from this assessment may then be used to establish an initial intervention tailored to the child’s individual learning needs.

There are other important derivatives of understanding a child’s cognitive profile. First, understanding a child’s cognitive profile may serve to establish teacher and parent expectations for a child’s learning trajectory and perhaps increase empathy among teachers and parents for a child’s particular academic plight. Second, this understanding may be useful in transition planning for graduating students who may be unaware of their strengths/weaknesses and uncertain of the skills necessary for a particular vocational path. Using this information, the adolescent can be directed toward a particular vocational path that is aligned with his or her profile of strengths/weaknesses.
An additional strength of norm-referenced cognitive/achievement assessment is that many of the instruments employed in the assessment process have been rigorously constructed and then scrutinized in accordance with professional test standards (AERA/ APA/NCME test standards, 1999). Measurement experts have analyzed the reliability and validity of these instruments, ensuring the quality of their psychometric properties and their overall technical adequacy. This allows for interpretation of the results to be made with a greater degree of certainty than interpretations based upon assessment instruments or approaches that may not have been as rigorously scrutinized (Salvia & Ysseldyke, 1985). In the past, psychologists, researchers and even the entire field of psychology have come under scrutiny for making interpretations and basing conclusions upon poorly constructed assessment instruments (Eysenck & Kamin, 1981). Therefore, it is essential and ethically mandated that practitioners ensure the psychometric quality (e.g., reliability and validity) of any assessment instrument or approach prior to adopting it for use.

The use of norm-referenced assessment also allows for enhanced communication, both among and between practitioners and researchers over what construct is being evaluated. For instance, when a school psychologist in New Jersey indicates that a child performed at the 4th percentile on a norm-referenced reading comprehension measure, a school psychologist in Georgia or California will understand that this child struggles with reading comprehension.

A well-constructed standardized assessment instrument will also place a child’s performance in a developmental context and allow interpretation of progress relative to a national peer group (Cicchetti, 1994; Reynolds, 1984a). Relative to CBA models, this may be an important niche of norm-referenced assessment instruments. Whereas CBA models ask the question, “where does this child stand relative to expected performance in the local curriculum?”, norm-referenced tests have always been concerned with the question: “How does this child differ from other children his or her age at a broader (e.g., national) level?” (Kamphaus, 2001). Moreover, given the variations in child and adolescent development, norm-referenced assessment through established protocol removes much of the subjectivity and observational error from estimating how a child is progressing at a particular age (Bracken, 2000). The use of norm-referenced assessment is also considered by some to be more objective and reliable than relying on classroom performance or the judgment of a teacher, regardless of experience (Reynolds & Kaiser, 1990). Others, however, have countered this contention by reporting on the accuracy of teacher’s performance evaluations of students (Gerber & Semmel, 1984; Gresham, MacMillan & Bocian, 1998). Further, curriculum variations across schools and educators may contribute to decreased reliability of children’s performance evaluations when conducted outside the realm of standardized protocol. Thus, when clinicians, parents and educators are interested in data regarding where a child stands relative to other children of the same age in the general population, norm-referenced assessment can provide for such comparison (Kamphaus, 2001).

**Limitations of Norm-Referenced Assessment**

One of the biggest criticisms of norm-referenced assessment is that it lacks treatment validity (Deno, 1985; Shapiro & Derr, 1990; Shinn, 1998). In other words, some consider it only obtusely related to potential remediation as it is not directly linked to the academic curriculum in which a child is currently being taught (Deno, 1985). This lack of treatment validity spurred the development of CBA models (Fuchs & Shinn, 1989; Marston, 1989; Reschly, 1997; Shinn, 1998). From this perspective, some have criticized standardized assessment as inefficient and wasteful of a school’s resources as well as the psychologist’s and child’s valuable time (Shapiro, 1990; Shinn, 1998). Some contend that rather than spending time administering a series of assessment batteries, the psychologist should be facilitating intervention for children (Shinn, 1998).

Another noted problem of standardized assessment is its lack of sensitivity to the actual performance of a child in the classroom (Marston, 1989). A child may be receiving accommodations and yet standardized assessment may be unable to detect any short-term progress (Fuchs, Deno, & Mirkin, 1984; Fuchs & Shinn, 1989). Compounding this problem is the lack of availability of more than one or two versions of the test, which also contributes to the instrument limitations in measuring a student’s achievement progress repeatedly across time. Moreover, considering that many standardized assessment instruments are reformed approximately every ten years, the possibility exists that the
content of the achievement test may no longer represent the curriculum being taught in the classroom (Bracken, 2000). It is also possible that children may be directly taught the content of the assessment instrument, should an unqualified individual administer the test. This would obviously spoil the results derived from that instrument and violate ethical standards (AERA, APA, & NCME Test Standards, 1999).

**Strengths of Curriculum-Based Assessment Models**

The following section provides a general overview of the advantages of Curriculum-Based Assessment (CBA) models, with particular attention focused on Curriculum-Based Measurement (CBM). CBA is defined broadly as any testing strategy that uses the student’s curriculum as a basis for informing decision-making regarding the student’s learning needs (Howell, Kuns, & Antil, 2002). Deno (1987) describes CBA as “direct observation and recording of a student’s performance in the local curriculum as the basis for gathering information to make instructional decisions” (p. 41). Although CBA is generally viewed as an overarching term for methodologies that link assessment with intervention, many different models fall under this rubric, and the specific purposes and procedures of these models can sometimes vary dramatically.

With this understanding in mind, all Curriculum-Based Assessment models are connected by three common features (Fuchs & Deno, 1991; 1994). First, assessment is linked to the student’s curriculum. Second, the student’s progress in the curriculum is evaluated to determine instructional/intervention success. Third, information from the assessment is used to tailor instructional/intervention decisions to more appropriately suit the learning needs of the student (Shapiro & Derr, 1990).

One type of CBA model, Curriculum-Based Measurement (CBM), has recently received considerable attention in the literature as a proposed alternative approach for LD assessment/intervention. Shinn (2002) describes CBM as a “set of standardized and validated short duration tests that are used...for the purpose of evaluating the effects of instructional programs in the basic skills of reading, mathematics computation, spelling and written expression” (p. 671). In some respects, CBM is a hybrid approach that incorporates the tenets of CBA in addition to the measurement characteristics of more traditional norm-referenced achievement instruments. In fact, one of the most important features of CBM is its emphasis on standardization, technical adequacy and appropriate local norming procedures (Shinn, 1998, 2002). CBM, like norm-referenced instruments, contains a standardized protocol, including directions, scoring procedures and interpretation heuristics. Any deviation from standardized protocol limits the conclusions that may be drawn from the test. Similarly, CBM procedures require adequate sampling for the establishment and use of the local norms. An additional distinctive feature of CBM is that the tests used to measure progress, called “probes,” are short, thereby allowing for repeated measurement of a student’s progress (Shinn, 1998).

With an understanding of the distinction between CBA and CBM in mind, it is now appropriate to turn to a discussion of the strengths of CBA models. One of the often-
cited strengths of curriculum-based assessment is that it is directly linked with the curriculum in which a child is being taught (Shapiro, 1990; Shapiro & Derr, 1990). This means that CBA directly measures student progress on actual materials taught in that student's classroom, rather than on material that is most typically taught in a particular grade. One of the overriding goals of assessment should be to inform intervention/instructional practices, and the various models of CBA were designed with this specific purpose in mind (Deno, 1985; Shapiro, 1990). Most models of CBA offer standardized procedures for monitoring a student's progress that are directly congruent with IDEA's mandate to tailor instruction to an individual student's specific learning needs (Fuchs & Shinn, 1989; Howell, Kurns, & Antil, 2002; Shapiro, 1990; Shinn, 1998).

CBM in particular can be useful for this purpose. It establishes student growth standards, allows for repeated evaluation of the effectiveness of those goals/objectives and allows for the alteration of the intervention if progress has not been achieved (Allinder, Fuchs, & Fuchs, 1998; Fuchs & Shinn, 1989). When progress is not evident, the intervention may be tailored to more appropriately meet the needs of a particular child.

As CBA is directly linked to the pedagogical activities within the classroom, it is viewed as an efficient use of a school system's professional resources (Deno, 1985; Shapiro, 1990; Howell, Kurns, & Antil, 2002). Instead of focusing time and energy on assessment activities that may only be peripherally related to the academic curriculum, professionals are able to use assessment approaches that are functionally related to intervention (Shapiro & Derr, 1990; Deno, 1985). This allows school professionals to understand not only when to change instructional approaches for a student, but also how to change those approaches (Shapiro, 1990).

Given their high degree of treatment validity, some have suggested that CBA and CBM are highly ethical approaches to assessment and intervention (Deno, Fuchs, Martson, & Shinn, 2001; Fuchs, Deno, & Mirkin, 1984; Gresham, 2001; Shinn, 1998). They do not require a child to be evaluated on an extensive assessment battery that may only be indirectly related to intervention. In light of the advantages of CBA broadly and CBM specifically, some have even advocated for these models to supplant traditional norm-referenced assessment as the approach of choice for learning disabilities classification (Gresham, 2001). Whether or not one agrees with this perspective, ample evidence is available that documents CBA's link with intervention planning and monitoring.

**Limitations of CBA and CBM**

Despite the many advantages of CBA, the models have their limitations. First, Curriculum-Based Measurement (CBM) incorporates norm-referenced procedures, allowing for a comparison among children, and accomplished through the development of local school-based norms (Shinn, 1998). Hence, some of the same criticism that has been directed toward traditional norm-referenced assessment can be applied to CBM.
Assessment: A Balanced Perspective... continued from page 17

For instance, there is the danger of technical inadequacy (i.e., poor reliability/validity) through poor norming procedures and inadequate sampling. Without appropriate local norms, the value of CBM is certainly limited. School personnel will be unable to discern how much of a deficit in an academic area is beyond the normal range, or how much change over time should be expected as an indicator of student progress.

Furthermore, a sufficient population must be sampled and the teacher must adhere to the curriculum for a comparison to remain valid (Shinn, 1998; Stewart & Kaminski, 2002). Were a teacher or school to adopt a curriculum different from the one used during the local norming process, then the assessment results may be suspect (Shinn, 1998) and the assessment approach may not be ethically grounded (AERA, APA, & NCME Test Standards, 1999). However, there has been recent research that provides evidence that the effectiveness of CBM probing techniques does not depend on such strict adherence to the student's actual curriculum (Fuchs & Deno, 1994; Powell-Smith & Bradley-Klug, 2001).

Some CBA models do not derive local norms, while typical CBM models only provide data comparing children to same grade peers in the local area (Stewart & Kaminski, 2002). Thus, it may be difficult to distinguish through CBA or CBM where a child stands relative to other children on a national level. Many practitioners and parents may wish to understand not only how a child is progressing locally, but also how that child compares to other children his or her age at a national level. This type of comparison may be useful, for instance, as a benchmark for future performance on college entrance exams that are referenced at the national level.

Finally, CBM has been criticized as being less predictive of performance in the broader curriculum (Shinn, 1998). While CBM probes in reading and mathematics have demonstrated adequate predictive validity with overall reading and math achievement (Good & Jefferson, 1998; VanDerHeyden, Witt, Naquin, & Noell, 2001), they do not generalize to some of the broader skills that might underlie poor achievement (Shinn, 1998). These models, however, were not necessarily intended for such purposes (Shapiro, 1990; Shinn, 1998; 2002). Therefore, practitioners and researchers should be wary of viewing CBA/CBM as an assessment methodology that is useful for all assessment purposes and able to transcend its initial theoretical underpinnings without additional empirical validation. To do so would be to akin to the past inappropriate practice of placing primary emphasis on IQ tests as a means of diagnosing severe psychopathology, such as schizophrenia, or using the Bender as a means of diagnosing emotional disturbance (Rampheus, 2001).
Conclusions

Considering recent proposals (e.g., NJCLD Roundtable Report) to implement a new approach to LD classification, it is important to review the strengths and limitations of the two most common approaches to the assessment of learning difficulties: norm-referenced standardized cognitive/achievement and curriculum-based assessment, particularly CBM. There has been considerable discussion of the limitations of the discrepancy approach and IQ testing for LD eligibility decisions, with some calling for their elimination. Norm-referenced cognitive/achievement assessment, however, should still play a role in eligibility/intervention decisions. As mentioned earlier in this article, norm-referenced measures of verbal IQ have high predictive validity and might be utilized to screen potential learning difficulties for children in the emergent stages of literacy acquisition. Norm-referenced achievement assessment via large-scale screening might also be potentially useful in identifying those students who are at risk for academic problems. In addition, norm-referenced instruments might be useful for discerning patterns and profiles underlying learning difficulties. Finally, norm-referenced assessment instruments place performance within a national context, allowing for the transfer and communication of results beyond the local area.

Of late, the literature has focused on assessment approaches such as CBM, possibly overlooking the merits of traditional, norm-referenced assessment. Unbalanced coverage of this nature might serve to unduly influence policy and professional practice by overlooking the contributions of norm-referenced assessment and, for that matter, alternative assessment models that might be relevant, but infrequently discussed in the literature, such as DIBELS (Good & Kaminski, 1996) or Dynamic Assessment (Lidz, 1991). Although it is important to attain a balanced perspective toward the two most common assessment approaches through an understanding of the strengths/limitations of each model, it might also be prudent to uniformly investigate the relevancy of additional assessment approaches for informing LD decision making.

Such an investigation should not occur in an arbitrary and capricious fashion, but rather through the lens of professional test standards (APA/AERA/NCME, 1999). In this manner, the assessment approaches under consideration can be subjected to rigorous scrutiny that ensures their technical adequacy. Equally important, a comprehensive review of additional assessment approaches will encourage the predication of policy and practice upon all available information, rather than on narrowly focused, outdated and perhaps even myopic perspectives.

Further, it is incumbent upon trainers of future school psychologists to tailor their programs to present the assessment of learning difficulties from a much more balanced perspective. Programs that are heavily behaviorally oriented should seek to integrate methodology from more traditional assessment approaches, while programs that are more psychometrically oriented would serve their students well to incorporate more functionally related assessment methodology. School psychology programs that have not revised their programs in such a fashion may be considered anachronistic. While there is certainly the need for classification and diagnosis, there is also the need for intervention that is directly related to classification and diagnosis. Norm-referenced assessment procedures might serve the former purpose of classification, while CBA methodology seems better suited for the latter. In the end, CBA and norm-referenced assessment models are complementary and should be viewed as synergistic. The time has arrived for a balanced perspective toward both.

*continued on page 20*
● Assessment: A Balanced Perspective... continued from page 18

References


© 2003, National Association of School Psychologists. Stefan C. Dombrowski, Ph.D., is an Assistant Professor and trainer of school psychologists at Rider University, Lawrenceville, NJ.