Review of Assessment Procedures for Students with Moderate and Severe Disabilities

Ellin Siegel
University of Nebraska-Lincoln

Rose M. Allinder
Rafiki Children’s Center, Nairobi, Kenya

Abstract: Assessment procedures for students with moderate or severe disabilities cover a range of procedures and practices. This article reviews recommendations for assessment practices for these students. These recommendations are compared with the assessment actually conducted by a sample of 22 teachers of students with moderate and severe disabilities. Three assessment purposes were analyzed; diagnosis and placement, curriculum and program development, and evaluation. Results are discussed in terms of best practices and recommendations for assessment of students with moderate or severe disabilities.

Assessment procedures for students with moderate and severe disabilities represent a conundrum. Students must be assessed to be classified and placed into special education, yet these assessments rarely yield information useful to practitioners for designing or evaluating effects of programming (Snell & Brown, 2006). Sigafoos, Cole, and McQuarter (1987) noted this dilemma in their investigation of then-current assessment practices of students with severe disabilities. Their analysis of assessment practices indicated that over 50 percent of students with severe disabilities had been administered standardized, norm-referenced tests (NRTs), which evidenced varying degrees of reliability and validity. In contrast, use of criterion-referenced tests and adaptive behavior scales was much less common. As Sigafoos et al. point out, use of NRTs, while required for classification purposes, requires supplementation by other sources of data.

This is necessary to ensure that data generated by the assessment process would be “relevant for planning and tracking on going progress” (Wolf-Schein, 1998, p. 52-53).

The necessity of additional types of assessment becomes more apparent when the purposes of assessment are more than classification and identification. Two critical areas that require assessment procedures that are carefully chosen or created, administered and interpreted are program development and instructional evaluation (Grisham-Brown, 2000; Macfarlane, 1998; Siegel-Causey & Allinder, 1998; Snell & Brown, 2006). For each of these, a range of assessment practices is promoted to meet best-practice standards.

Program development, in which an individual education plan (IEP) for the student is developed, may be grounded in the assessment done for classification and placement purposes (Alper, Ryndak, & Schloss, 2001). Assessment for these purposes includes, at minimum, evaluation of intelligence and adaptive behavior. If data drawn from assessment for eligibility purposes are used to develop IEPs, several caveats must be issued. One relates to use of intelligence tests for these students. The appropriateness of these tests has been questioned for a number of reasons. First, for students who have accompanying disabilities, which make responding in standard format impossible, modification of response type makes validity and reliability unknown. Second, validity and reliability is...
questionable unless tests’ normative populations include students with the same range and level of cognitive functioning (Salvia & Ysseldyke, 2001). A final concern with using intellectual functioning for program development purposes, for students who are able to complete them, is that these tests were designed to compare them with their age peers, not to document what students do and do not know or can and cannot do.

In addition, caveats for using adaptive behavior assessment results must be noted. While this type of assessment also yields a norm-referenced score similar to that obtained in intelligence tests, it is possible to use adaptive behavior assessment for program development purposes by examining ratings of specific items within the domains of the adaptive behavior scale. Using adaptive behavior scales for this purpose, however, must consider if they are conducted only tri-annually. The information may be too out-dated to be of use for annual IEP evaluation and planning. Practitioners are urged to re-assess adaptive behavior more often when updating and revising IEPs.

While adaptive behavior scales may be used for program development purposes, other types of assessments should be used to develop each student’s IEP. These other assessments may include criterion-referenced tests (CRTs) and ecological inventories (Snell & Brown, 2006). CRTs measure the degree to which a student has obtained specific skills. As Beirne-Smith, Ittenbach, and Patton (2002) point out, CRTs vary with regard to what skills are assessed and the specificity with which skills are assessed. Practitioners are urged to use care in selecting CRTs to use with students with more severe disabilities and may find that creating their own CRTs is more beneficial for the student. Ecological inventories are a form of assessment in which activities and skills necessary for different environments, such as community recreation, vocational, domestic or general education settings are analyzed (Browder, 2001; Downing, 2002). A student’s performance on the ecological inventory contributes to the IEP both for the present level of performance and goal-setting sections.

Some students with moderate disabilities may require assessment of functional academic skills. CRTs and NRTs may provide insight into the degree of academic skill development as well as pinpointing specific functional academic skills to be targeted for instruction.

Using either CRTs or ecological inventories or some other method of assessing students for the purpose of educational programming can result in an overwhelming number of concerns. In this mix of data, family concerns, interests, and desires may be slighted. This should be avoided because creating education programs around family concerns, interests and desires may increase the likelihood that the student will obtain and maintain educational goals. The importance of building students’ programs around family concerns is underscored by IDEA 97 regulations that require that “concerns of the parents for enhancing the education of their child” (34 CFR 300.346) be incorporated into the IEP process. Several processes have been developed which help uncover families’ visions and hopes and dreams for their children and which help delineate important and meaningful goals for individual students. These processes are intentionally strengths-based, focusing not merely on the needs and deficits of the student with moderate or severe disabilities, but their strengths. These methods include personal futures (Mount & Zwerkik, 1988), Making Action Planning System (MAPS; Forest & Snow, 1987), and Choosing Outcomes and Accommodations for Children (COACH; Giangreco, Cloninger, & Iverson, 1998).

In personal futures or lifestyle planning, concerned family members, friends, advocates and professionals meet to collaboratively evaluate an individual’s quality of life and describe a desirable future for that person (Browder, 2001). MAPS (Forest & Snow, 1987) is a collaborative team approach to identifying critical information about the student with severe disabilities. Family members, professionals, advocates, and peers relate their dreams, nightmares, and descriptions of the student as well as their descriptions of the student’s strengths, needs, and ideal day. Responses to these questions can aid in developing and prioritizing goals for the student’s IEP (Forest & Lusthaus, 1990). Similarly, COACH (Giangreco et al., 1998) is used by school personnel and family members to identify IEP goals and objectives in accordance with family-centered...
priorities, using a structured family-interview approach.

After the educational program has been developed and instruction begins, it is critical to assess student progress. As Wolf-Schein (1998) points out, “educational assessment is useful insofar as it contributes to positive development of the child.” (p. 38) The best care in selecting, administering, and interpreting eligibility assessment information or program development information is pointless if the child does not make progress. Typical ways of assessing progress in other students may be of limited use for students with moderate or severe disabilities. It is important that assessments, which are sensitive to small changes, are used. This may be various ongoing assessments such as informal anecdotal recording keeping, task analytic assessment, discrete trial or opportunity assessment, curriculum-based measurement, or time-based assessment (Farlow & Snell, 1994; Macfarlane, 1998).

While expert advice is to craft assessment practices in alignment with best-practices standards (e.g., Macfarlane, 1998; Siegel-Causey & Allinder, 1998) it is not known to what degree practitioners implement these recommendations. Sigafoos et al. (1987) studied assessment practices for a large number of students with severe disabilities by examining students’ cumulative files. Results of that study indicated that tests used with these students were lacking in technical qualities such as reliability and validity and that measures other than NRTs such as CRTs and adaptive behavior measures were used less frequently than recommended. This study did not investigate informal assessment practices to collect student performance data in order to make educational decisions such as programming development and instructional evaluation. The current study was designed to do this. Specifically, we examined assessment practices of 22 teachers of students with moderate or severe disabilities across the different purposes of assessment.

Method

Participants

Participants were 22 special education teachers from different schools in an urban school district in the Midwest. Teachers were chosen from volunteers who had students with moderate or severe disabilities and who had not been trained previously to use a specific form of alternative assessment with their students. Both teachers and parent(s) of participating students completed consent forms.

All but one of the teachers were employed full-time. Median age of teachers was between 40–49 and; the mean number of years teaching experience was 12.1 (range 15-23). Thirteen had a bachelor’s degree and nine had a master’s degree.

Participating students were in Kindergarten through fourth grade with a mean age of 9 years 2 months (range 5.0–11.2). Each child was classified using state eligibility criteria as mentally handicapped; moderate or severe/profound; multi-handicapped; or autistic. Table 1 details student demographics.

Procedure

We were interested in learning about all assessments used with students in this study. Therefore, we examined student records from cumulative files and from teachers’ records, limiting assessment information to the previous five years. Information obtained was photocopied at the student’s school and student names were blacked out immediately and replaced with a code number. To obtain this information, teachers were asked to make available to project staff all assessment information for the target student. To verify that project staff obtained copies of all assessments for each child, two forms of follow up were conducted. First, project staff used a previously-generated list of possible assessment information and interviewed teachers at their schools to see if some items on the list were available for their student but had not been given to the project staff. Second, teachers were given a listing of possible assessments, which indicated what had been obtained by project staff for their student. Teachers were asked to review the listing and to contact staff if there were errors (e.g., assessments existed but had not been given to the staff).

Once all assessment information was obtained, we categorized each item into one of three categories: (a) diagnosis and placement, (b) program development, and (c) instruc-
tional evaluation (Grisham-Brown, 2000; Macfarlane, 1998; Siegel & Allinder, 1998; Snell & Brown, 2006).

According to Snell and Brown (2006), the process of diagnosis and placement is used to determine the type of disorder displayed by the child and to make decisions regarding eligibility, classification, and placement. Types of assessments that fall under this category include intelligence tests, adaptive behavior measures, and motor and sensory functioning measures (Taylor, 1997). These assessments usually are administered by a psychologist or specialist (e.g., physical therapist, vision specialist). Program development, the second category, refers to assessments that are used to ascertain skills that should be taught and are used in the development of the child’s IEP. This type of assessment is frequently conducted by the teacher in the classroom environment. Finally, assessments are used to evaluate the educational program of the child, including student progress. Assessments used here may include taking frequent probe data, curriculum-based measurement, or frequent and informal anecdotal recording keeping (Farlow & Snell, 1994). We compared types of assessment recommended by best practice literature for each assessment purpose with what this sample of teachers actually did.

Results

Diagnosis and Placement

By state statute, students identified as having a mental handicap must be administered a cognitive and an adaptive behavior measure. In our sample of students, seven participating students had an intellectual assessment. For seven other students, the multi-disciplinary team (MDT) had determined that formal cognitive assessment was not possible due to the severity disability. Specific tests administered for cognitive functioning varied and included: the Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983) \((n = 3)\); Bayley Scales of Infant Development (Bayley, 1984) (attempted with four children, but successfully completed for only two); Stanford-Binet Intelli-

<table>
<thead>
<tr>
<th>Student</th>
<th>Gender</th>
<th>Age</th>
<th>Grade</th>
<th>Primary disability label</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. D.</td>
<td>Male</td>
<td>11.2</td>
<td>5</td>
<td>Mental handicap: severe profound</td>
</tr>
<tr>
<td>B. L.</td>
<td>Male</td>
<td>8.0</td>
<td>2</td>
<td>Multiple disabilities</td>
</tr>
<tr>
<td>C. T.</td>
<td>Male</td>
<td>9.10</td>
<td>3</td>
<td>Other health impairment</td>
</tr>
<tr>
<td>C. T.</td>
<td>Male</td>
<td>11.0</td>
<td>4</td>
<td>Other health impairment</td>
</tr>
<tr>
<td>C. N.</td>
<td>Female</td>
<td>6.5</td>
<td>1</td>
<td>Multiple disabilities</td>
</tr>
<tr>
<td>E. E.</td>
<td>Male</td>
<td>7.6</td>
<td>2</td>
<td>Autism</td>
</tr>
<tr>
<td>E. T.</td>
<td>Male</td>
<td>8.2</td>
<td>3</td>
<td>Autism</td>
</tr>
<tr>
<td>E. T.</td>
<td>Male</td>
<td>9.9</td>
<td>3</td>
<td>Autism</td>
</tr>
<tr>
<td>H. Y.</td>
<td>Male</td>
<td>10.11</td>
<td>4</td>
<td>Mental handicap: severe profound</td>
</tr>
<tr>
<td>H. E.</td>
<td>Female</td>
<td>12.0</td>
<td>5</td>
<td>Mental handicap: moderate</td>
</tr>
<tr>
<td>H. S.</td>
<td>Female</td>
<td>7.9</td>
<td>K-1</td>
<td>Mental handicap: moderate</td>
</tr>
<tr>
<td>H. N.</td>
<td>Female</td>
<td>8.10</td>
<td>2</td>
<td>Orthopedically impaired</td>
</tr>
<tr>
<td>H. N.</td>
<td>Female</td>
<td>10.11</td>
<td>4</td>
<td>Mental handicap: moderate</td>
</tr>
<tr>
<td>K. A.</td>
<td>Female</td>
<td>5.0</td>
<td>K</td>
<td>Autism</td>
</tr>
<tr>
<td>M. Y.</td>
<td>Male</td>
<td>10.3</td>
<td>4</td>
<td>Mental handicap: severe profound</td>
</tr>
<tr>
<td>P. G.</td>
<td>Male</td>
<td>10.8</td>
<td>4</td>
<td>Mental handicap: moderate</td>
</tr>
<tr>
<td>P. T.</td>
<td>Male</td>
<td>10.6</td>
<td>4</td>
<td>Mental handicap: moderate</td>
</tr>
<tr>
<td>P. E.</td>
<td>Male</td>
<td>8.4</td>
<td>3</td>
<td>Autism</td>
</tr>
<tr>
<td>R. H.</td>
<td>Female</td>
<td>8.4</td>
<td>3</td>
<td>Multiple disabilities</td>
</tr>
<tr>
<td>R. Y.</td>
<td>Male</td>
<td>10.2</td>
<td>5</td>
<td>Mental handicap: moderate</td>
</tr>
<tr>
<td>S. N.</td>
<td>Female</td>
<td>10.3</td>
<td>5</td>
<td>Autism</td>
</tr>
<tr>
<td>Z. N.</td>
<td>Male</td>
<td>6.7</td>
<td>1</td>
<td>Mental handicap: moderate</td>
</tr>
</tbody>
</table>
gence Scale (Thorndike, Hagen, & Sattler, 1985) (attempted with two children but successfully completely for only one), Woodcock
Johnson Test of Cognitive Ability (Woodcock & Mather, 1989) (unsuccessfully attempted with one child), and Cognitive Abilities Test (Thorndike & Hagen, 1994) (n = 1).

In the domain of adaptive behavior, all but one student received some type of assessment and several students received more than one. The Vineland Adaptive Behavior Scale (Sparrow, Balla, & Cicchetti, 1984) was the most common (n = 18). Other measures were: Scales of Independent Behavior (Bruininks, Woodcock, Weatherman, & Hill, 1984) (n = 3), Maxfield-Buchholz Social Maturity Scale for Blind Preschool Children (Maxfield & Buchholz, 1957) (n = 2) and Developmental Assessment of Life Experiences (Barber et al., 1995) (n = 3).

In addition to assessing cognitive and adaptive behavior skills, most students had other NRTs administered. In the domain of speech/language, all but one had some assessment. Informal assessment, including observation of the child and/or interview of the child’s parent(s) and/or teacher, was common (n = 20). Among formal assessments administered, the Peabody Picture Vocabulary Test- Revised (Dunn, 1984) was the most prevalent (n = 6), followed by the Arizona Articulation Proficiency Scale (Fudala, 1987) (n = 4), the Preschool Language Scale - 3 (Zimmerman, Steiner, & Pond, 1979) (n = 4), the Nonspeech Test (Huer, 1988) (n = 4), the Expressive One-Word Picture Vocabulary Test (Gardner, 1979) (n = 3), the Rossetti Infant-Toddler Language Scale (Rossetti, 1990) (n = 2), Communication and Symbolic Behavior Scale (Wetherby & Prizant, 1993) (n = 2), and four others.

Only five students were administered formal tests of academic ability, although one subject was unable to complete the test. Assessments used included Woodcock-Johnson Psycho-Educational Battery, Tests of Achievement (Woodcock & Johnson, 1989) (n = 2) and the Psycho-Educational Profile- Revised (Schopler, Reichler, Bashford, Lansing & Marcus, 1990) (n = 2).

As part of their diagnosis and placement assessment procedures, 15 students were assessed in the domain of behavior. The majority of these (n = 11) were by means of inform-
Developmental skill levels of reading and math, which were completed and maintained for general education students.

Augmentative communication objectives and recommendations were used in the curriculum development for some students. In addition, for students who changed schools, a Transition Planning form or checklist was used to facilitate transition to the new school environment.

As an alternative to formal testing, individualized assessments may provide information that is more useful for a given child with special needs (Snell & Brown, 2006). Nearly one-third of the teachers created some form of a task or functional analysis, specific to their student’s needs. For example, one teacher used a functional analysis to define specific steps involved for a student who makes purchases at a grocery store in the local community.

Specialists often assisted to develop or improve programs for students in this sample. These specialists were school district personnel with expertise in programming for students with more severe disabilities (referred to as inclusion cadre personnel) or in behavioral interventions (behavioral specialists), and who consulted with individual teachers. Inclusion cadre was found for nearly three fourths of the students (n = 16). Additionally, behavior specialists and other specialists were involved in the program development of some children. Approximately one-third of the students (n = 7) attended clinics outside the school setting. Reports from these clinics suggested specific recommendations for the given student.

**Evaluation**

Not surprisingly, all students had documentation in the form of IEP progress reports or student progress reports (regular or narrative), which are both evaluative documentation required by the local school district.

Some teachers (n = 8) maintained a chart that listed IEP goals and objectives and used it to document ongoing progress. Other methods used by teachers to track growth included keeping a notebook of anecdotal records (n = 3) and keeping samples of student work to indicate progress (n = 6). Some teachers (n = 4) incorporated a home note system in order facilitate communication with the child’s caregiver(s). Additionally, several teachers designed charts to show improvement for specific skills such as behavior (n = 4), bathroom activity (n = 3), and classroom activities (n = 2). Spelling lists (n = 2) and high frequency word lists (n = 2) were used for a few students to show ongoing development in these specific literacy skills.

Besides the child’s homeroom or resource teacher, other individuals were involved in the evaluative process for some children. For example, four students had narrative reports by other teachers (art, PE, etc.). In addition, 14 students had progress reports by an occupational or physical therapist, seven students had reports by a speech/language therapist, and one student had a report regarding neurodevelopment. Finally, two students had evaluative reports for summer school.

**Discussion**

In our sample of students with moderate and severe disabilities, we compared the actual assessment processes with those recommended by best practice standards. This comparison yielded a mixed review of how these practices were implemented.

With regard to diagnosis/placement purposes of assessment, there was evidence that few students had been administered a formal norm-referenced test of intelligence. While this measure is required by state statute and definition of mental retardation, it appeared the MDTs demonstrated good clinical judgment in not administering the test to students whose functioning level was severely impaired or to students whose multiple disabilities made it difficult or problematic to ensure that results were valid. Problems associated with administering intelligence tests to students such as were in our sample is not uncommon.

An adaptive behavior measure was administered to 21 students. This is in contrast to results of Sigafos et al. (1987) that the majority of their sample were not administered this type of assessment. However, we echo the caveat of Sigafos et al. in stating that we cannot determine how the data from these
measures were used. It is possible that results for adaptive behavior measures were used to fulfill eligibility guidelines and not to plan educational programs.

In addition to these required measures, most students in our sample had additional assessment as part of the referral/placement process. The primary type was for communication, via tests for speech, receptive language, and expressive language.

With regard to assessment for curriculum planning, teachers appeared to follow best practice standards to a lesser degree. Five types of criterion-referenced tests were used with the students in our sample. Additionally, five teachers had conducted task analyses of IEP objectives and one teacher had completed an ecological inventory. Limited evidence of a broader range of CRTs, task analyses, and ecological inventories was disheartening and could lead to the conclusion that students’ IEPs might not have been as individualized, comprehensive, or specific as they might have been.

Only one participant had completed a procedure to ascertain family preferences, dreams, desires with regard to their child’s future, by completing MAPS. Five other teachers had used the school district one-page form querying families about what goal areas would they like included on their child’s IEP. The lack of broad based personal-planning procedures among our participants was discouraging given the emphasis placed on this in the professional literature in recent years and given that the participating school district had previously conducted inservices on this topic.

Regarding assessment practices for program evaluation purposes, the pattern of mixed results continued. Best practices for this assessment purpose include anecdotal recording keeping, task analytic assessment, discreet trial or opportunity assessment, curriculum-based assessment, or time-based assessment. In our sample, we found evidence that all of there were used to some degree. Anecdotal recording was common, both by teachers and by related services personnel such as physical therapists or behavior specialists. Furthermore, there was evidence that a small number of teachers used each of the other types of recommended assessments.

Conclusion

We examined assessment practices and procedures of 22 teachers of students with moderate or severe disabilities. From this examination we found a pattern of mixed evidence that these practitioners were following best practice recommendations. The one area in which these assessment practices appeared to be in closest alignment with best practice recommendations was that of diagnosis/placement. That 21 of 22 students had an adaptive behavior measure completed was encouraging. This is tempered by the supposition that these measures may have been conducted only to fulfill state statute and that results did not appear to have been used when designing students’ programs. The area of diagnosis/placement seemed to typify best practice standards for a second reason. Only two students had scores from tests that were normed for a younger age population. This is certainly an improvement over results reported a decade ago (Sigafoos et al., 1987). Additionally, not all students had intelligence scores reported. Given the extent and nature of the disabilities present in our sample of students, this appeared to have been good judgment of case of decision makers.

In contrast to the more positive findings in the diagnosis/placement area, assessments conducted for program planning and program evaluation were less encouraging. Most disappointing was the fact that only one student had results of MAPs reported. Absence of information of this type was perplexing. It is not possible to determine if this was due to lack of training of teachers, recency of students placed on these teachers’ caseloads, or difficulties because of more logistical reasons such as lack of time to conduct assessments.

Assessment for program evaluation, however, appears to be even more mixed. Not surprising, anecdotal recording was the most common among our sample of teachers. Much less common were assessments that were tied directly to IEP objectives and which would yield data informing instructional decision making. This was especially problematic from our viewpoint, given research stating that students with more severe disabilities make progress less easily and thus need their growth monitored more closely.
**Recommendations**

Teachers of students with moderate or severe disabilities should continue to articulate the need for manageable and educationally informative assessment, especially for program planning and evaluation purposes. Moreover, administrators who supervise and support these teachers should be well versed in appropriate assessment procedures and should be able to teach, model, and refine assessment procedures for these students. While assessment is one of the most important topics in education today, assessment encompasses well more than high-stakes assessment. Assessment used for planning individual education programs and evaluating effectiveness of instruction must be emphasized as well.

**References**


Received: 3 February 2004
Initial Acceptance: 31 March 2004
Final Acceptance: 25 August 2004