ASSESSMENT OF MOTOR DEVELOPMENT AND FUNCTION IN PRESCCHOOL CHILDREN

Beth L. Tieman,1* Robert J. Palisano,2 and Ann C. Sutlive1
1Georgia State University, Department of Physical Therapy, Atlanta, Georgia
2Drexel University, Programs in Rehabilitation Sciences, Philadelphia, Pennsylvania

The process of identification of children with delays or disorders in motor development includes developmental screening, examination, and reexamination. Throughout this process, various types of measures are used, including discriminative and evaluative measures. Discriminative and evaluative measures of motor development and function that are commonly used for preschool-aged children include the Bayley Scales of Infant Development II, Peabody Developmental Motor Scales, 2nd edition, Toddler and Infant Motor Evaluation, Pediatric Evaluation of Disability Inventory, and Gross Motor Function Measure. Selecting an appropriate measure is a crucial part of the examination process and should be geared toward the purpose of testing and characteristics of the child. Evidence of reliability and validity are important considerations for selection of a measure.

Key Words: motor development; motor function; children; assessment; measurement

Neurodevelopmental assessment of preschool age children (18 months to 4 years) includes examining motor development and function. Standardized measures are administered as part of the examination process. A number of standardized measures are available to measure motor development and function in preschool-aged children. Selection of an appropriate measure will depend on the purpose of testing and characteristics of the child, in addition to reliability and validity of the measure. The purpose of this article is to summarize the following: (a) the process of identification of preschool-aged children with a delay or disorder in motor development, (b) contexts for assessment of motor development and function, (c) selection of reliable and valid measures, and (d) interpretation of scores. Five measures are critiqued and cases are presented to illustrate the decision-making process.

IDENTIFICATION OF CHILDREN WITH DELAYS OR DISORDERS IN MOTOR DEVELOPMENT

Developmental Screening

The purpose of screening is to identify children who appear to have delays in development and to make an appropriate referral. Primary care providers, such as pediatricians, are often the first professionals to identify children who have delays in motor development. Developmental screening may be performed as a routine part of a primary care visit or may be initiated by parent concern. The American Academy of Pediatrics Committee on Children with Disabilities [2001] advocates that a role of primary care providers is to screen the development of infants and preschool-aged children. Many children with a history of neonatal complications are at risk for delays in motor development and thereby receive regular developmental screening through a neonatal follow-up program. Developmental screening tools include several developmental domains (e.g., language, motor, and cognition) and should be relatively short and inexpensive. One screening tool, the Denver Developmental Screening Test II, was designed for use in primary care settings [Frankenburg and Dodds, 1990]. Another screening tool, the Ages and Stages Questionnaire is administered via parent questionnaire [Bricker and Squires, 1999]. Referral for a multidisciplinary developmental assessment is recommended for children whose screen suggests delays in at least two areas of development. A child who appears to have a delay in motor development may be referred to a physical therapist or occupational therapist for examination.

Examination

Examination is the process of obtaining a history, performing a systems review, and selecting and administering tests and measures to gather data about the child [APTA, 2001]. The examination should begin with an interview, whereby family concerns and needs related to the child’s development are discussed. A standardized measure of motor development is administered by a physical or occupational therapist. While administering a standardized measure, the therapist simultaneously observes how the child performs tasks in order to evaluate quality of movement and motor planning. On the basis of these observations, the therapist identifies sensory, motor, and musculoskeletal system concerns that require examination (e.g., sensory processing, strength, balance, skeletal alignment, and range of motion). Recommendations are made based on the results of the comprehensive examination.

*Correspondence to: Beth L. Tieman, P.T., Ph.D., Assistant Professor, Georgia State University, Department of Physical Therapy, Atlanta, GA, USA. E-mail: btieman@gsu.edu

Received 3 November 2004; Accepted 12 July 2005
Published online in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/mrdd.20074

© 2005 Wiley-Liss, Inc.
Reexamination

Reexamination refers to the process of performing selected tests and measures to evaluate progress and to modify or redirect interventions [APTA, 2001]. Upon initiation of therapy, the date for reexamination is determined. Readministration of a standardized measure of motor development provides an indication of overall progress. Standardized measures should be complemented by evaluation of change in individualized goals that reflect the focus of therapy intervention.

TYPES OF MEASURES

The different types of measures of motor development and function for preschool-aged children include discriminative, evaluative, curriculum-based, individualized outcome measures, and comparative performance measures (Table 1). This article will focus on discriminative and evaluative measures.

Discriminative Measures

These are norm-referenced measures designed to identify children with delays in development. A child’s score is compared with the average performance of the normative sample (children of the same age who are typically developing). The results are used to document delay and determine eligibility for therapy services in early intervention and the public school system. Discriminative measures are generally not appropriate to use as outcome measures for children with identified delays or disorders in development [Rosenbaum et al., 1990]. The Peabody Developmental Motor Scales (PDMS-2) [Folio and Fewell, 2000] and the Bayley Scales of Infant Development II (BSID-II) [Bayley, 1993] are discriminative measures critiqued in this article.

Evaluative Measures

These are designed to measure change over time or in response to an intervention. Responsiveness, or the ability to measure meaningful change over time [Guyatt et al., 1987], is a type of validity that is necessary for evaluative measures. Most evaluative measures are criterion-referenced, meaning that a child is assessed on individual progress on a specific skill (i.e., walking) or domain (i.e., gross motor function), rather than comparing with norms for children without delays. For example, an evaluative measure may measure the child’s progress toward walking, rather than comparing the child’s ability to walk with typically developing peers. The Gross Motor Function Measure (GMFM) [Russell et al., 2002] is an evaluative measure critiqued in this article.

Curriculum-Based Assessments

These are a type of criterion-referenced measure that include multiple domains of development (e.g., language, motor, and cognition). The purposes of curriculum-based assessments are to plan interventions and to measure progress over time, usually for children enrolled in an early intervention or preschool program. Even though most curriculum-based assessments are based on the normal developmental sequence, they are not norm-referenced. Scores on curriculum-based assessments, therefore, are not compared with a normative sample, but instead measure a child’s achievement specific to the context of a school curriculum. Curriculum-based assessments are designed for administration by an interdisciplinary team, including therapists, teachers, and other educational professionals. There are many curriculum-based assessments available for preschool-aged children, including the Assessment, Evaluation, and Programming System for Infants and Children [Bricker, 2002], Battelle Developmental Inventory [Newborg, 2004], Carolina Curriculum [Johnson-Martin et al., 1990], and the Hawaii Early Learning Profile [Vort Corporation, 1999].

Individualized Outcome Measures

These are designed to measure an individual’s change in function related to the focus of intervention. Individualized outcomes provide a method to measure changes in the context of home, school, and community setting and therefore evaluate performance (i.e., what the child does do) in daily life. An individualized outcome measure, the behavioral objec-

<table>
<thead>
<tr>
<th>Type of Measures</th>
<th>Purpose</th>
<th>Comparison of Child’s Performance</th>
<th>Application of Results</th>
<th>Example Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discriminative measures</td>
<td>To identify children with developmental delays</td>
<td>To normative sample (children of same age who are typically developing)</td>
<td>To document delay and determine eligibility</td>
<td>Peabody Development Motor Scales (PDMS-2) [Folio and Fewell, 2000]</td>
</tr>
<tr>
<td>Evaluative measures</td>
<td>To measure change over time</td>
<td>To child’s previous scores of a specific skill</td>
<td>To measure change over time</td>
<td>Gross Motor Function Measure (GMFM) [Russell et al., 2002]</td>
</tr>
<tr>
<td>Curriculum-based assessments</td>
<td>To measure progress over time in a curriculum</td>
<td>To child’s previous scores of a skill in a curriculum</td>
<td>To plan interventions and instructional activities</td>
<td>Carolina Curriculum [Johnson-Martin et al., 1990]</td>
</tr>
<tr>
<td>Individualized outcome</td>
<td>To measure an individual’s change in function related to interventional focus</td>
<td>To child’s previous performance (individualized to each child)</td>
<td>To evaluate progress on intervention goals</td>
<td>Goal Attainment Scaling [King et al., 1999]</td>
</tr>
<tr>
<td>measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparative performance</td>
<td>To aggregate scores for a group of children</td>
<td>To child’s scores at admission and discharge</td>
<td>To compare outcomes within and between health care facilities</td>
<td>Functional Independence Measure for Children (WeeFIM II)</td>
</tr>
<tr>
<td>measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
tive, is a statement that specifies an individualized therapy goal in measurable terms. This includes selection of behaviors that are observable and repeatable, specification of the conditions under which performance is measured, and a time frame for achievement of the outcome. Goal Attainment Scaling (GAS) is another type of individualized measure with unique features that are advantageous for measuring change [King et al., 1999]. GAS format uses five possible outcomes as opposed to one. Scores of $-2$ and $-1$ represent the two less favorable levels of attainment, a score of 0 represents the expected level of attainment, and scores of $+1$ and $+2$ represent the two levels of attainment that exceed expectations. A change score expressed as a T-score with a mean equal to 50 and a standard deviation of 10 can be computed for multiple goals. Palisano et al., [1992, 1993] provide evidence of the validity of GAS to measure change in motor function in young children receiving therapy services.

Comparative Performance Measures

These are designed to aggregate scores for groups of children with a similar health condition for purposes of program evaluation. Comparative performance measures are intended to be relatively brief and administered at admission and discharge. Scores are used to evaluate efficiency of care, standards of care, and outcomes. The process involves contracting with a vendor who analyzes data and provides a comparison with outcomes reported by other institutions or a benchmark. Comparative performance measures have primarily been used with children receiving in-patient rehabilitation. The Functional Independence Measure for Children (WeeFIM II) is the most widely used outcomes management system for children with acquired or congenital disabilities [Uniform Data System for Medical Rehabilitation, 2004]. The WeeFIM II is designed to measure the need for assistance and severity of disability in children (6 months to 7 years) and includes three domains: self-care, mobility, and cognitive.

**SELECTION OF A MEASURE**

The most important decision in developmental assessment is selection of an appropriate measure. The choice of which measure to use will depend on responses to the following questions:

1. *What is the purpose?* For identifying a developmental delay or for eligibility for services, a discriminative measure is most appropriate. For measuring the effect of therapy intervention over time, an evaluative measure, curriculum-based assessment, or individualized outcome measure is most appropriate.

2. *What are the characteristics of the child?* The age of the child is very important, as most measures are intended for use within a specific age range. The functional abilities of the child will impact on the selection process. For example, the measure selected will differ between a child who walks but has limitations in gross motor skills and a child who uses a wheelchair for mobility. The cognitive and language abilities of the child also affect selection, especially if a measure involves complex verbal instructions.

3. *What are the developmental or functional areas to be examined?* Some assessments measure gross motor or fine motor development, whereas others measure both domains. Other assessments measure self-care, play, and functional mobility.

4. *In which setting will the examination occur?* Therapists examine children in a variety of settings, including the child’s natural environment (e.g., home, daycare, and preschool), rehabilitation clinic, or in-patient hospital. Certain measures may be more appropriate for a given setting. In addition, therapists should consider the distinction between capability (what a child can do) in a controlled environment versus performance (what a child does do) in everyday settings (e.g., home, school, and community) [Tieman et al., 2004]. Most measures of development assess capability. There is a need for measures that identify environmental factors that facilitate and restrict performance of daily activities and routines.

5. *What are the external constraints of testing?* External constraints of testing may include the time available to administer a measure, examiner training, space and equipment required, and the cost of purchasing a measure.

**RELIABILITY AND VALIDITY**

Evidence of reliability and validity are important considerations for selection of a measure. The test manual, and subsequent published articles, should present evidence of reliability and validity. Reliability is the accuracy and consistency or repeatability of measurements. Reliability studies are conducted during development of a new measure. To obtain a reliable score, it is essential to follow the standardized procedures for administration and scoring of each item. The primary two types of reliability are interrater (the agreement of scores between two or more examiners) and test–retest (the consistency of repeated measurements separated in time). Test–retest reliability is critical for evaluative measures to ensure that scores on readministration represent changes in development and not measurement error.

Validity is the extent to which an assessment tool actually measures what it is intended to measure. Validity, therefore, is specific to each measure, depending on its intended purpose. For example, a valid measure of gross motor development should include postures and whole body movements of increasing difficulty. For discriminative measures, scores should systematically increase with age. The types of validity include the following: construct validity (the extent to which items reflect the theoretical construct of interest), content validity (the extent to which items are representative of the domain that is measured), concurrent validity (the extent to which scores are related to another measure of the same construct), and predictive validity (the extent to which scores will predict future outcomes).

**EXAMINER QUALIFICATIONS**

Qualifications for administration of standardized measures of motor development and function include a strong background in child development, experience testing in young children with disabilities, and an understanding of tests and measures to interpret scores correctly. Prior to administering a measure for the first time, the manual should be reviewed thoroughly, including scoring procedures. Administration and scoring should be practiced on several children, and should include training and feedback from a mentor experienced in using the measure. Clinicians must be aware of the factors within the child (e.g., interest and behavior), environment (e.g., noise and distractions), and clinician (e.g., rapport and observation skills) that can negatively affect the reliability of results, even when
there is evidence that the measure has good reliability.

**INTERPRETATION OF SCORES**

The raw score on a norm-referenced measure is converted into one of several scores using the tables provided in the test manual. A **percentile rank** compares the raw score obtained by the child to the raw score of the normative sample. An **age equivalent** reflects the average age of children in the normative sample who achieved the raw score. **Standard scores** are interpreted in terms of the normal distribution of scores, with a certain mean and standard deviation for each score. Most developmental measures have a mean standard score of 100 and a standard deviation of 15. Standard scores below 85 generally are interpreted as a delay. Standard scores are preferred when the purpose is to identify whether a child has a developmental delay.

For criterion-referenced measures, scores reflect achievement on the domain or skill of interest. For some measures, raw scores can be converted into a **percentile score**, which reflects the percentage of passed items out of a total possible score. Percentile scores can be compared across time periods to measure progress. In addition, criterion-referenced measures may have undergone a mathematical procedure, Rasch item response analysis, to place items along a continuum of difficulty [Wright & Masters, 1980]. Individual items are ranked from less difficult to more difficult, allowing the tracking of a child’s progress in a more meaningful way. Rasch item response analysis was used in development of two measures critiqued in this article, the Pediatric Evaluation of Disability Inventory (PEDI) [Haley et al., 1992] and the GMFM [Russell et al., 2002]. For both measures, items are arranged in order of difficulty and the score reflects both the number and difficulty of items the child is able to perform.

**CRITIQUE OF SELECTED MEASURES**

We have selected five measures of motor development and function for preschool-aged children, 18 months to 4 years, to review in this article (see Table 2). These measures are the following: BSID-II, PDMS-2, Toddler and Infant Motor Evaluation (T.I.M.E.), PEDI, and GMFM. These measures were selected, as they are commonly used to measure motor development and function and have evidence of reliability and validity. Long and Toscano, [2002] provide a comprehensive review of motor assessments for children from infancy through school-aged.

**Bayley Scales of Infant Development II**

The second edition of the BSID [Bayley, 1993] consists of a mental scale (language and perceptual), motor scale (gross and fine motor), and behavioral scale. The purpose is to identify children, 1 month to 3.5 years, with developmental delay. Norms for the BSID-II were developed from a sample of 1,700 children. Normative data also have been collected on high-risk children such as those with Down Syndrome, prematurity, HIV, and prenatal drug exposure. A limitation of the BSID-II motor scale is the inability to differentiate gross and fine motor development. Furthermore, the number of items (111 total) does not provide for a comprehensive measure of motor development. Items are scored as credit or no credit without accounting for emerging ability. For preschool-aged children, the age range is somewhat limited such that another measure would need to be administered after age 3½ years.

The BSID-II has concurrent validity with the PDMS-2 for age equivalent scores but not for standard scores [Provost et al., 2000; Provost et al., 2004]. The motor scale of the BSID-II is a valid discriminative measure that is recommended for identification of motor delay or determination of eligibility for early intervention services.

**Peabody Developmental Motor Scales, 2nd edition**

The PDMS-2 [Folio and Fewell, 2000] includes comprehensive norm-referenced gross motor and fine motor scales intended to determine whether children of age 6 years and younger have delayed motor development, based on a large normative sample (n = 2,003). The scales consist of 127 gross motor items and 122 fine motor items, divided into six subtests. The PDMS-2 is a valid measure for determining a child’s eligibility of services in early intervention and preschool programs. While Folio and Fewell, [2000] indicate that the PDMS-2 can be used for both discriminative and evaluative purposes, there is no evidence of responsiveness to change for children with disabilities. As an evaluative measure, therefore, the PDMS-2 may be appropriate for children with motor delays, but not for children with specific neuromotor impairments [Palisano et al., 1995].

The PDMS-2 is the second edition [Folio and Fewell, 2000] and is substantially different than the first edition [Folio and Fewell, 1983], with a new normative sample, different test structure, and more specified scoring criteria. Scores from the first and second editions, therefore, cannot be directly compared. The PDMS-2 has excellent reliability and validity as a discriminative measure. The score sheet provides abbreviated instructions for each item and can be used for repeated administrations. Separate scores are obtained for gross and fine motor development, allowing for determination of relative differences in gross motor and fine motor development.

**Toddler and Infant Motor Evaluation**

The T.I.M.E. [Miller and Roid, 1994] is a norm-referenced tool that uses an observational, play-based approach to assess quality of movement while encouraging parental involvement. The T.I.M.E. is detailed and comprehensive, especially regarding motor skills and transitional movements. The T.I.M.E. has unique clinical subtests that rate the qualitative aspects of atypical motor performance, including hypertonicity and hypotonicity. The test also includes a functional performance subtest, including self-care and community function. A social–emotional abilities subtest includes behavioral state, attention, and emotional reactions.

There are five primary subtests and three optional clinical subtests, each with different scoring systems. The entire test is lengthy, but can provide detailed information about the child’s movement patterns. The scoring system is complex and may initially be cumbersome and lengthy for some examiners, especially for novice practitioners [Long and Tie man, 1998]. Therapists may need to videotape test administration so that the detailed scoring system can be completed later [Rahlin et al., 2003].

While the primary purpose of the T.I.M.E. is to identify children with motor delays, the authors suggest that the tool may be used for evaluative purposes as well [Miller and Roid, 1994]. Standard scores that compare a child’s score to the normative sample can be obtained for the five primary subtests. In addition, for the motor organization subtest, a “growth score” can be used to measure change over time. Research on responsiveness to change is needed, especially because of the structural and scoring problems that have been detected in the T.I.M.E. [Rahlin et al., 2003].
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purposes</strong></td>
<td>● To identify developmental delay</td>
<td>● To estimate motor competence relative to peers</td>
<td>● To identify children with motor problems</td>
<td>● To measure functional capabilities and performance</td>
<td>● To measure change over time in gross motor function in children with cerebral palsy</td>
</tr>
<tr>
<td></td>
<td>● To monitor developmental progress</td>
<td>● To compare gross and fine motor skills</td>
<td>● To identify patterns of movement</td>
<td>● To monitor progress in rehabilitation</td>
<td>● To use for intervention planning</td>
</tr>
<tr>
<td><strong>Type of test</strong></td>
<td>Norm-referenced</td>
<td>Norm-referenced</td>
<td>Norm-referenced</td>
<td>Norm-referenced, criterion-referenced</td>
<td>Criterion-referenced</td>
</tr>
<tr>
<td><strong>Age range</strong></td>
<td>1 month to 3.5 years</td>
<td>Birth to 6 years</td>
<td>4 months to 3.5 years</td>
<td>6 months to 7.5 years</td>
<td>Not specified</td>
</tr>
<tr>
<td><strong>Test structure</strong></td>
<td>Motor scale (fine and gross motor); mental scale (language and perceptual); behavior rating scale</td>
<td>Fine motor and gross motor scales; 6 subtests (GM scale); ● Reflexes</td>
<td>Primary subtests: ● Social-emotional</td>
<td>Three scales (functional skills, caregiver assistance, modifications) in 3 content domains: self-care, mobility, and social function</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Therapist administers elicited items using standardized instructions and materials</td>
<td>● Stationary</td>
<td>● Mobility</td>
<td>Five dimensions: ● Lying/rolling</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Therapist observes and guides parent-child interaction</td>
<td>● Object manipulation</td>
<td>● Motor organization</td>
<td>● Sitting</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Parental interview</td>
<td>● Grasping</td>
<td>● Stability</td>
<td>● Crawling/kneeling</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Therapist places child in starting position and elicits item</td>
<td>● Visual-motor integration</td>
<td>● Functional performance</td>
<td>● Standing</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Motor organization</td>
<td>● Clinical subtests: ● Quality rating</td>
<td>● Walking/running/jumping</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Stability</td>
<td>● Component analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Functional performance</td>
<td>● Atypical positions</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Normative sample</strong></td>
<td>$n = 1,700$</td>
<td>$n = 2,003$</td>
<td>$n = 731$ (typical children); $n = 144$ (children with disabilities)</td>
<td>$n = 412$ (normative sample); $n = 102$ (clinical sample)</td>
<td>Not applicable</td>
</tr>
<tr>
<td><strong>Procedure</strong></td>
<td>Therapist completes via parental report, structured interview, or professional observation of the child’s functional behavior</td>
<td>Therapist observes and guides parent-child interaction</td>
<td>Parental interview</td>
<td>Therapist places child in starting position and elicits item</td>
<td></td>
</tr>
<tr>
<td><strong>Time required</strong></td>
<td>25–60 min</td>
<td>45–60 min</td>
<td>10–55 min</td>
<td>20–60 min</td>
<td>45–60 min</td>
</tr>
<tr>
<td><strong>Item scoring</strong></td>
<td>● Binary: credit/no credit</td>
<td>3-point scale: 0,1,2 (1 for partial credit)</td>
<td>● Social-emotional: 5-point scale</td>
<td>● Functional skills: 0/1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Refused, omit, and caregiver report recorded</td>
<td>● Mobility: record sequence of movement</td>
<td>● Caregiver assistance: 5-point scale</td>
<td>● Caregiver assistance: 5-point scale</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Motor organization: credit/no credit</td>
<td>● Modifications: total number (frequency)</td>
<td>● Modifications: total number (frequency)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Stability: highest item passed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Functional performance: 2-point scale</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Types of total scores</strong></td>
<td>Age equivalent, developmental index (mental and motor); percentile rank (behavioral)</td>
<td>Age equivalent, percentile rank, standard score (subtests), composite quotient (fine and gross motor)</td>
<td>Standard score, percentile rank, growth score (motor organization subtest)</td>
<td>Normative standard score, scaled score (criterion-referenced), fit score (compared with expected Rasch fit)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$r = 0.75$ (motor); $r = 0.96$ (mental)</td>
<td>$r = 0.96$ (total test)</td>
<td>$r = 0.88–0.99$</td>
<td>GMFM-88: percentage score; GMFM-66: interval-level total score</td>
<td></td>
</tr>
<tr>
<td><strong>Interrater reliability</strong></td>
<td></td>
<td></td>
<td>Normative sample: ICC = 0.96–1.00; clinical sample: ICC = 0.84–1.00</td>
<td>GMFM-88: ICC = 0.87–0.99</td>
<td></td>
</tr>
</tbody>
</table>
Pediatric Evaluation of Disability Inventory

The PEDI [Haley et al., 1992] is a comprehensive measure designed to identify functional limitations and monitor progress in children with disabilities. The PEDI is unique in that it was designed to measure functional abilities of children with physical disabilities or a combination of physical and cognitive disabilities. The PEDI has been used to evaluate outcomes of in-patient rehabilitation for children with traumatic brain injury, nontraumatic brain injury, orthopedic conditions, and neurological conditions [Haley et al., 2002; Dumas et al., 2002]. In contrast to norm-referenced measures of development, the mobility and self-care domains emphasize functions important for daily life, including transfers, locomotion, eating, grooming, dressing, and bowel and bladder control. The PEDI can be used either as a discriminative measure (when norm-referenced scores are used) or as an evaluative measure (when criterion or scaled scores are used). Items were standardized on both a normative sample of 412 children and a sample of 102 children with disabilities. The PEDI is administered by structured interview with the parent or observation of the child. The PEDI has rating scales for modifications (e.g., assistive devices) and caregiver assistance for each domain (e.g., mobility, self-care, and social function). The scales are independent and are scored separately, with the option to use only the scales that are relevant to the child and the purpose of testing.

The PEDI has strong psychometric properties, although in terms of validity, there may be a “floor effect” because young children (6 months to 2 years) would not be expected to achieve many items, including dressing, using utensils, and toileting. Scaled scores are used to measure change over time that represent newly achieved skills or new levels of independence [Iyer et al., 2003]. A scaled score change of 11 points or more represents an important clinical change [Iyer et al., 2003]. Alternatively, self-care classification levels can be used to report change scores [Dumas et al., 2001].

Gross Motor Function Measure

The GMFM [Russell et al., 2002] is an evaluative measure designed to measure change over time in children with cerebral palsy (CP). The GMFM-88 was originally developed with 88 items and validated not only for children with CP but also for children with Down Syndrome [Russell et al., 1989; Russell et al., 1994; Russell et al., 1998]. Items on the GMFM were recently scaled on a sample of 537 children with CP using a Rasch measurement model [Russell et al., 2000]. This resulted in the GMFM-66, which places items in order from easy to difficult and created an interval level score, the Gross Motor Ability Estimator. The GMFM-66 should be used only for

Table 2. Continued

<table>
<thead>
<tr>
<th>Test-retest reliability</th>
<th>Construct validity</th>
<th>Content validity</th>
<th>Concurrent validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>r = 0.78 (motor); r = 0.87 (mental); r = 0.55–0.70 (behavior)</td>
<td>From BSID (1st ed.): • Factor analysis • Constructs become more differentiated with age • Correlations of items within each scale reported</td>
<td>Items reviewed by experts</td>
<td>With PDMS-2 subtests r = 0.85–0.97 (motor, AE scores), r = 0.67–0.76 (motor, quotient scores) (Provost et al., 2004)</td>
</tr>
<tr>
<td>ICC = 0.67–1.00 (Nichols and Case-Smith, 1996)</td>
<td>• Factor analysis • Subtest scores correlated with age • Age trends for more items mastered and mastery of difficult</td>
<td>GMFM-88 and GMFM-66; ICC = 0.99</td>
<td>Not reported</td>
</tr>
<tr>
<td>r = 0.96–0.99</td>
<td>• Scores of functional behaviors increase with age • Caregiver assistance and functional skill scales reflect different dimensions of function</td>
<td>• Change scores correlated with parent and therapists ratings • Scores “responsive” to changes • Change scores related to the child’s age and severity of CP</td>
<td>With the Battelle r = 0.70–0.73 (total sample)</td>
</tr>
<tr>
<td>r = 0.89 (total test, age 2–11 months); r = 0.96 (total test, age 12–17 months)</td>
<td>• Rasch analysis • Internal consistency α = 0.72–0.97</td>
<td>GMFM-88: Pilot testing with therapists; GMFM-66: Rasch analysis and reliability of item difficulties</td>
<td>With the WeeFIM r = 0.80–0.97 (total sample)</td>
</tr>
<tr>
<td></td>
<td>• Internal consistency α = 0.95–0.99</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
children with CP. A scoring CD is included with the GMFM-66 test manual, which is used to obtain a score, in addition to providing visual item maps of the child’s function.

The GMFM has been used extensively in clinical practice and research studies to assess a variety of interventions [Damiano and Abel, 1998; McLaughlin et al., 1998; Bower et al., 2001]. The GMFM can be used for clinical decision-making and intervention planning. Because the items are administered in a controlled environment, scores on the GMFM reflects capability (what a child can do) instead of performance (what a child does do) in everyday settings [Tieman et al., 2004]. Quality of movement is assessed with a companion measure the Gross Motor Performance measure [Boyle et al., 1995].

CASE EXAMPLES

The following cases illustrate the selection and use of measures in clinical practice.

Case 1: Greta is a 3-year-old child who attends a preschool program. Greta’s family and teacher notice that Greta has difficulty keeping up with her classmates (e.g., falling frequently, running slowly, and difficulty jumping). Her parents request a comprehensive examination of motor development. In this case, a discriminative measure is recommended as part of the comprehensive examination to determine if motor development is delayed and eligibility for services. The PDMS-2 is norm-referenced measure and provides a detailed assessment of gross motor development. The BSID-II includes fewer items and does not differentiate gross motor and fine motor development. In addition, the BSID-II and T.I.M.E. have an upper age of 3.5 years and thus may not measure Greta’s full range of abilities. The PEDI and GMFM do not match with the purpose of the examination.

Case 2: John is a 4-year-old child who had a traumatic head injury. He has general delays in all developmental domains, including cognition, language, and motor. John’s family has questions about his functional performance, such as eating, mobility, and communicating, compared with his peers, in order to decide on an appropriate school placement for John. In this case, the PEDI would provide information on self-care, mobility, and social function, as compared with the normative sample. Furthermore, because the PEDI is also criterion-referenced, the scaled scores can be used upon reexamination to measure progress in these functional areas. The T.I.M.E. has a functional performance subtest, but it is unclear whether this subtest can be administered in isolation of the other subtests. The BSID-II and PDMS-2 predominantly measure motor development, rather than functional performance, and therefore do not match with the purpose of examination.

Case 3: Maria is a 2-year-old child with CP who is referred for physical therapy services. The GMFM-66 would be appropriate to administer, as this tool was designed to measure change of gross motor function in children with CP. The results are shared with Maria’s parents. Standing and walking are identified as areas for intervention. Six months later, the therapist conducts a reexamination in order to measure Maria’s progress in standing and walking. Because the GMFM-66 is criterion-referenced, Maria’s score on reexamination is compared with her previous performance. Maria is classified at level II on the Gross Motor Function Classification System [Palisano et al., 1997]. Her scores are plotted on the motor development curve for children in level II [Rosenbaum et al., 2002] to estimate her progress compared with other children with CP in level II. In this case, a norm-referenced test, such as the BSID-II or PDMS-2, is not necessary to administer, since the therapist does not need to measure Maria’s gross motor function compared with her age-matched peers, to measure her progress in therapy.

CONCLUSION

Selecting an appropriate measure is a crucial part of the examination process and should be geared toward the purpose of testing and characteristics of the child. The examination process includes screening, examination, and reexamination. To identify a developmental delay or for eligibility for services, a discriminative measure is most appropriate. To measure the effect of therapy intervention over time, an evaluative measure is most appropriate. Evidence of reliability and validity are important considerations for selection of a measure. Measures commonly used for preschool-aged children, 18 months to 4 years, include the PDMS-2, BSID-II, T.I.M.E., PEDI, and GMFM. ■

REFERENCES


