
The Peabody Developmental Motor Scales (PDMS), developed over a 10 year period, is a significant addition to the inventory of movement product tests. It is designed to measure both gross and fine motor development of children in the age range from birth through 83 months. It is based on the concept that motor development proceeds in an orderly sequence. The PDMS, like other movement product scales, has established motor milestones and includes the age range in which they should normally appear. The total test can be administered in about 45 to 60 minutes with individual scales requiring approximately half that time. Instructions are provided that help examiners to administer the test to groups of children in a station format whenever the number of examinees requires a more expedient process.

The gross motor scale contains 170 items with 10 items at each of 17 age levels. The items are classified into five skill categories: (a) reflexes, (b) balance, (c) non-locomotor, (d) locomotor, and (e) receipt and propulsion of objects. The fine motor scale consists of items in four skill categories: (a) grasping, (b) hand use, (c) eye-hand coordination, and (d) manual dexterity. The authors provide norms for each skill category at each age level as well as total composite scale scores. The test kit provides materials needed for the fine motor items while the gross motor items only require materials that are commonly found in motor programs. Fifteen copies of the response/scoring booklet are also included in the kit while others can be ordered separately.

The test designers list eight functions that the PDMS can serve. The Standards of Educational and Psychological Tests (APA, 1974) require that validity data be presented for each intended use. Data were presented in the manual to empirically support three uses: (a) developmental screening, (b) provide information on what motor tasks have been mastered, motor skills currently developing, and those not in the child’s behavior repertoire, and (c) determining unique motor strengths and weaknesses. I do not necessarily doubt that the test can be used for the other stated purposes but support the continuous process of test validation. Most of the other functions could be supported by the collection of a sample of teacher and therapist judgements.

The activity cards that accompany the PDMS are matched to the individual test items. This completes the comprehensive assessment-based programming system advocated by Folio and Fewell.

Administration, Scoring, and Interpretation

The PDMS was designed to be administered by a wide range of professionals including special education teachers, physical and occupational therapists, and physical educators. Folio and Fewell suggest that program aides and volunteers who have had experience following specific testing instructions can also learn to administer the test. An interesting graduate thesis might be to investigate how much training is actually needed by aides to become a reliable test administrator.

Examiner qualifications, specific administrative procedures, and appropriate
cautions are clearly stated in the test manual. Subsequent to practicing with several children, the test user should have little problem duplicating the administrative conditions under which the norms and the data on reliability and validity were secured. This conclusion is based on a review of the step by step administrative process described in the manual. Directions for administering the test to handicapped children for eligibility and placement purposes (although not specifically listed as a formal purpose) are described along with instructions for adapting the procedures to make instructional programming decisions only. The developers correctly recommend that the normative data should only be used if the standardized procedures were followed.

The scoring of the PDMS is one of the major advantages that this test has over other published movement product tests (e.g., The Denver Developmental Screening Test and the Bayley Scales). The norms are based on scoring each items as 0, 1, or 2. A 0 score indicates that a child could not or would not attempt the item, or the attempt does not show that the skill is emerging. A 1 indicates that the student's performance shows a clear resemblance to the item criterion but does not fully meet the criterion. A 2 rating refers to a child that performs the item according to the specified criterion. The test developers indicate that the major area of judgement is whether a response should be scored as 0 or 1 but they provide several excellent examples to help the test administrator make this judgement. The scoring system of the PDMS along with the provision of a clear description of the behavioral criterion for each item are the elements that distinguish this test from other published movement product tests. The three level scoring system allows for individual differences found in all children and especially the handicapped. The behavioral criterion for each item will definitely help to improve the reliability of measurement which is an area of concern for other early screening instruments.

The authors encourage test administrators to record comments concerning the child's behavior during assessment. These comments should aid teachers in prescribing instruction based on the results of testing. One example might be the child that appears distractable or to have a short attention span. The teacher will need to manipulate the instructional environment to reduce distractions if maximum learning is to occur.

**Test Development**

According to the test manual, PDMS was developed to improve upon the existing motor assessment instruments. It was designed to meet the need for accountability and for specified guidelines in programming for handicapped children in general and in physical education in particular. The authors mention that other instruments that are available do not provide a means for planning objectives-based programs. It has been my experience that the OSU Sigma (Loovis & Ersing, 1980) and I CAN (Wessel, 1976) provide extremely specific motor assessment data and include curricular materials for prescriptive purposes. Both of these systems are widely used with exceptional learners.

The first edition of the scales incorporated motor items resulting from a review of a number of validated motor development scales. New items were created based on studies of children's growth and motor development. The original scales used a 5-point scoring system which was modified to the current 3-point system based on results obtained during field testing. Other changes were made during the extensive field testing resulting in several revised experimental editions. The final result was the 1979 Field Test Edition which was used in the standardization procedures conducted from December 1981 through April 1982.
In describing the longitudinal changes in fundamental gross motor patterns as reflected in the test items, the authors have obviously ignored the major contributions of motor developmentalists in the past decade (e.g., Roberton, 1978; Roberton & Halverson, 1977). One example of this can be found in items 78, 83, 94, 95, 122, 126, and 150 which purported to evaluate throwing a ball. All but one item (Item 122) evaluates the accuracy of the throw from various distances, while most motor development texts would suggest that the movement pattern or process being used during the early childhood period is a more critical concern. Distances and accuracy should only be a concern after a mature pattern has been mastered.

Standardization

The sampling design and the characteristics of the norming sample are described in the test manual. The basic sampling procedure was a stratified quota design. The sample was stratified by geographical region, race, and sex and totaled 617 children between 0 and 83 months of age. Twenty states were selected in the four major census regions consisting of 37% from the Southern region, 20.4% from the Northeast, 24.3% from the North Central, and 18% from the West.

When the sample distribution is examined in terms of gender and race, it closely approximates the 1980 census data. Test sites were selected on the basis of two major considerations: The schools had to reflect the urban-rural characteristics of their respective region, and also had to represent a range of socioeconomic levels. The test administrators selected were physical therapists, occupational therapists, psychologists, teachers, and paraprofessionals.

Prospective motor skill test developers should feel confident in using the PDMS standardization plan as a model for future test development purposes with one exception. Since the norms for certain age intervals are based on a relatively small number of observations the possibility of sampling error is present. Norms are presented for the 17 age groups while 11 have less than 35 subjects. Most measurement texts suggest a minimum of 100 per age group (Salvia & Ysseldyke, 1978).

Reliability

Folio and Fewell report the essential aspects of reliability as suggested in the Standards for Educational and Psychological Tests (APA, 1974). The standard error of measurement which is extremely important for interpreting test scores is reported for each age group. According to the data presented for gross and fine motor scores, the small standard errors of measurement suggest that the PDMS has excellent precision of measurement across all age intervals.

Test-retest stability was evaluated by assessing 38 children in the original norming sample on a second occasion. Both evaluations were conducted by the same individual. The reliability coefficients presented for those items that were actually administered to each subject for both the gross and fine motor scores were .95 and .80 respectively. Interrater reliability was determined by having one test administrator give the test to 36 subjects and score the performances, and a second individual observe and also score the performances. The reliability coefficients for item scores were .97 for the gross motor scale and .94 for the fine motor scale.

Although the data presented indicates very high stability I have several concerns about the reliability and the specified purposes of the scales. First, the test manual or a related journal article on reliability should describe how much training the test administrator had prior to the study. A high reliability coefficient obtained from highly
skilled or experienced test administrators does not mean the same as one obtained from
a test administrator who had little or no formal training. Users of most standardized
tests will have little experience or no formal training. A test should be evaluated under
conditions in which it will frequently be used. Secondly, the test designers make a strong
case for the usability of the scales by a wide range of professionals including
paraprofessionals. Reliability should be studied using a variety of test administrators.
Thirdly, the authors attest to its application to handicapped infants and children,
therefore it would be interesting to investigate the stability of the two scales with
handicapped subjects. Finally, the use of classical test theory in studying reliability of a
motor performance test is not as precise as the use of generalizability theory. The
correlation of two data sets which is the procedure in classical test theory, averages over
all possible sources of error (or unreliability). Generalizability theory using ANOVA to
determine the relative magnitude of each individual source of error (e.g., raters,
occasions) provides information that is necessary for both the test user and developer.
By analyzing the interactions between the various sources of error the developer can
determine if certain test items are more difficult to rate thus requiring more training.
The results of this type of reliability information will be better assessments and
educational decisions. All of these concerns should be addressed on a post hoc basis and
included in separate journal articles or the second edition of the scales.

Validity

Information on content, construct, and criterion-related validity is presented in the test
manual. The PDMS manual states that content validity was verified by research into
children's motor development, a review of other validated motor development tests,
and by an examination of Harrow's (1972) Taxonomy of the Psychomotor Domain.
According to Brown (1980), claims of content validity should be based on a systematic
study of the relation of the test items to the description of the skill domain. The test
manual does an excellent job of clearly describing the content domain being measured
but confuses face validity with content validity. Independent raters should be used to
judge whether the test items adequately sample the various skill domains. I believe they
do, but until a systematic investigation is reported this remains a potential area of
concern.

Construct validity was established by determining whether the gross and fine
motor test scores improved as a function of age. Several tables present data to support
that significant improvement occurs with age for total scores and individual skill scores.
Data are also presented to support the prediction that the performances of nonclinical
children are consistent across skill areas. The final source of construct validity was a
comparison of children with developmental delays with a group experiencing normal
development. The results suggest that the developmentally delayed group performed
significantly below the normal developing children on the PDMS.

Criterion-related validity was established by evaluating the concurrent validity
and involved determining whether a group of subjects performed similarly on the
PDMS as they did on the Bayley Scales (Bayley, 1969) and on a motor development test
designed by a New York State Rehabilitation Hospital. The pattern of correlations
between the performances provides good support for concurrent validity. Gender and
cultural validity was reported to be very high thus allowing the performances of both
genders as well as Blacks and Hispanics to be evaluated by using the norms presented in
the test manual.
The Activity Cards

The PDMS test manual includes a chapter on instructional programming that educators and therapists will find extremely helpful for using the results obtained from testing to design individualized education programs for gross and fine motor development. This chapter presents four case studies that include excellent examples of real test information and IEPs. Folio and Fewell also include teaching strategies that will help to facilitate programmatic interventions with the mentally retarded, sensory impaired, behaviorally disordered, and learning disabled student.

The activity cards that are included in the tab-indexed card file can be used as a resource for instructional programming. The file includes 170 activity cards for gross motor and 112 for fine motor skill development. Each card is matched to a test item and provides a behavioral objective and four instructional strategies for facilitating the objective. The professional with little knowledge in motor skill activities will find these cards to be beneficial, but most will need to supplement them with one of the many excellent preschool or elementary physical activity texts.

Concluding Comments

The Peabody Developmental Motor Scales and Activity Cards are designed to provide a wide range of professionals responsible for the gross and fine motor development of children with an assessment-based programming system. The standardization sample is generally representative of the U.S. population, and the test manual should be used as a model by future test developers. The strengths of the PDMS include its three level scoring system, the inclusion of both fine and gross motor components, the standardization plan, and the technical characteristics of the instrument. The major weakness is the movement product orientation of the test items. Most items measure the distance or accuracy of movement rather than the motor pattern being used by the young child. Most motor developmentalists suggest that motor patterns should be emphasized before the product of performance is evaluated. In conclusion, the PDMS is a significant improvement over other tests of this nature, but Folio and Fewell or their associates need to continue to investigate and publish studies on the technical characteristics of the scales under various conditions.

References


Reviewed by Dale A. Ulrich, Physical Education Department, Southern Illinois University at Carbondale.