A Handbook of Evaluation Instruments for Use With Children from Birth to Three Years

by

SARA ANN SNYDER

B. S., Kansas State University, 1980

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Family and Child Development

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1984

Approved by:

Major Professor

Table of Contents

Table of Contents (Continued)

Evaluation of children during the period of birth to three years is of value in: 1) screening for potential delays; 2) determining the child's current level of functioning in different developmental areas; and 3) planning a program to meet the needs of the child as revealed by the evaluation procedure. It also serves to measure change in the child's performance over the course of the program.

The teacher is the logical person to be involved in the evaluation process for a number of reasons. It is most often the child's teacher who will be involved in the evaluation procedure at some point in time (e.g., making the original referral, making use of evaluation results). The benefits of the teacher filling the role of evaluator include: 1) teacher-child rapport is already established; 2) the child is in a familiar environment; 3) the teacher is familiar with the child's personality and learning style; 4) evaluation is ongoing; and 5) the teacher is in the position to re-evaluate goals (Cross & Goin, 1977).

There are many evaluation instruments currently available, designed for use with young children from birth to three years. These instruments cover a broad range of behaviors and uses (e.g., screening, diagnosis, determining individual strengths and weaknesses for program planning).

The professional responsible for choosing an instrument for use in an early childhood program often does not have the time to evaluate these instruments on an individual basis and has limited access to a range of instruments. Many times, the only available information about an instrument is the summary description in the publisher's catalogue.

This handbook was designed primarily to provide a reference for early childhood educators for choosing appropriate evaluation instruments for their programs. The information should also be of value to other professionals involved in the evaluation of very young children. The evaluations of the instruments included in this handbook will enable the reader to compare the intended purposes of the available instruments to the specific needs of a particular early childhood program.

Types of Instruments

Screening, evaluation of current level of functioning, and measurement of change as a result of programming intervention are generally conducted utilizing one or more of the following types of instruments; 1) norm-referenced tests; 2) criterion-referenced instruments; and 3) curriculum-referenced measures.

Norm-referenced tests. Those measures which compare a given child's performance to the performance of his peers are referred to as norm-referenced. "The intent during development of norm-referenced tests is to maximize the variance among students so as to allow very specific differentiations among

students in terms of their knowledge or skills" (Goodwin & Driscoll, 1980, p. 58). Emphasis is placed on the child's relative standing in his peer group, rank ordered from those who have mastered many skills to those who have mastered few, rather than on the mastery of particular skills (Salvia & Ysseldyke, 1978). Norm-referenced tests involve a standardization process whereby norms are established, representing descriptions of the test performance of reference groups of children. The reference group for some tests consists of a national standardization sample, while for tests intended for local use only, a local sample may be used as the reference group (Goodwin & Driscoll).

Criterion-referenced instruments. These tests are a recently developed type of instrument measuring the development of a child's skills in terms of absolute levels of mastery (Salvia & Ysseldyke, 1978). Individual differences are de-emphasized and an individual's performance is compared against pre-determined criteria rather than to a reference group. Children are generally evaluated as either having mastered or not mastered the skills listed (Goodwin & Driscoll, 1980) and these data are then used to design individual program plans. Criterion-referenced instruments are sometimes mistaken as norm-referenced because they are usually developed by selecting items from various standardized tests. The assessment items frequently include age levels (taken from the norm-referenced tests) at which they should be passed. However, these are not norms, but rather serve as guidelines to the evaluator for which items are appropriate to administer. Criterion-referenced instruments are generally more comprehensive in developmental areas assessed and more detailed within developmental areas than norm-referenced tests.

Curriculum-referenced instruments. Curriculum-referenced instruments measure where the child performs in a sequence of objectives listed for the general curriculum of an early childhood program (White, Edgar, & Haring, 1978). The sequence and developmental ages are validated by curriculum practices and current pupil tests (Brigance, 1978).

Curriculum-and-criterion-referenced instruments may or may not be the same thing. Curriculum-referenced tests are designed to reflect skills taught in existing curricula, whereas criterion-referenced instruments are designed to assess developmental sequences, the results of which are used to compose individualized program plans (i. e., curricula). The fine distinction being made here is: Is the child assessed and made to fit into an existing curriculum or is the curriculum made to fit the child, based on his/her assessment?

Criteria Used to Evaluate Instruments

The "usability" of an instrument (Goodwin & Driscoll, 1980) should be determined for the targeted early childhood program. The test manual should provide specific information regarding: 1) costs; 2) equipment required; 3) the purpose of the measure; 4) the test format; 5) administration procedures; 6) scoring procedures; 7) interpretation of test scores; 8) validity and reliability of the measure; and 9) a bibliography.

Cost. Along with the cost of the evaluation instrument and score sheets, additional costs must be considered. Some of these costs may include the purchase of equipment for testing, training and paying test administrators, time lost from classroom instruction, and the cost of

inaccurate results (Stangler, Huber & Routh, 1980).

Equipment. The test manual should include a list of the equipment necessary for administering the test. This allows the evaluator to determine what additional equipment would be required to administer the test and determine the portability of the equipment, if testing will not be conducted in one place.

Purpose. The purpose of the evaluation suggests the choice of which type of instrument is to be used in an early childhood program. Norm-referenced tests are generally most suited for screening, grouping, funding, and program evaluation purposes. Criterion-referenced instruments are a means to evaluate the child's progress in mastering developmental skills, and are useful for writing objectives for individual children and program planning. Curriculum-referenced instruments are used to assess the child's initial position and progress in a given program curriculum.

Format. Consideration of the test format includes such items as the instructions to the child, the kinds of responses required of the child, the organization of the manual and score sheets, and the visual appeal of the stimuli.

Administration. Information concerning administration of the measure should include training procedures required to learn to use the test, testing group size, the amount of time required to administer the test, specific instructions, and testing procedures.

Scoring. Scoring instructions and types of scores should be included in the manual to inform teachers of the product obtained from the assessment, the complexity involved in obtaining that product, and the time that must be committed to scoring itself. Objectively scored tests are more practical than subjectively scored tests, requiring less training of scorers and resulting in greater reliability of test scores. Instruments having long, involved scoring procedures are less practical in early childhood programs because of the time commitment required.

Interpretation. Specific evidence of validity, reliability, and characteristics of standardization samples is necessary for accurate interpretation of scores. In the norm-references tests, norm tables, standard scores, age equivalents, and referral criteria should be provided in the test manual.

Standardization, validity, and reliability. A primary consideration in the evaluation of an instrument is the adequacy of the standardization, validity, and reliability of the instrument, including a comparison of the group used to develop the instrument with the group to be evaluated and a comparison of traits included on the instrument to the goals of the targeted early childhood program.

Information regarding the standardization group is essential in determining the appropriateness of the use of an instrument for a particular group. This judgment is based on a comparison of the demographic (e.g., sex, age, socio-economic status, area of residence) data of the reference group with the characteristics of the target group. This comparison

can not be made unless a comprehensive description of the reference group is included in the manual.

"Validity refers to the extent to which a test measures what its authors or users claim it measures" (Salvia & Ysseldyke, 1978, p. 95). It also involves a judgment of the appropriateness of the use of the instrument for the targeted group of children. This judgment is based on the measure's content validity, criterion-related validity, and construct validity.

The first of these, content validity, is usually associated with achievement tests and is a judgment of how well the items represent the content and processes of a particular subject domain. This is usually determined by having subject-matter experts judge whether the test items are representative of the described subject area. One approach is to have two or more judges rate each task in relation to the objective it is supposed to measure. Another approach commonly used is to have judges place test items in a pre-constructed matrix to determine whether the items measure what they were intended to measure. In both methods, inter-judge agreement is computed (Goodwin & Driscoll, 1980; Salvia & Ysseldyke, 1978).

The second type of validity, criterion-related, is an indication of the relationship between two tests measuring the same traits (concurrent validity) or the relationship between test scores and future performance (predictive validity). This type of validity is determined for shortened versions of existing instruments (e.g., McCarthy Screening Test and McCarthy Scale of Children's Abilities), new instruments measuring traits for which well-established valid instruments already exist, diagnostic instruments, and developmental surveys (Goodwin & Driscoll, 1980).

The third type, construct validity, is concerned with how well a test measures the theoretical construct (e.g., intelligence, self-esteem) it intends to measure. This can be determined "when a test correlates substantially with variables with which it theoretically should correlate (convergent validity), and also when it does not correlate with variables from which it should differ (discriminant validity), or that the latter correlation is less than the former" (Goodwin & Driscoll, 1980, p. 87). For example, one might expect a high correlation between WISC-R scores and academic achievement and a low correlation between WISC-R scores and popularity with peers. Examples of construct validity are difficult to find in instruments designed for young children (Goodwin & Driscoll, 1980).

An otherwise valid measure may be inappropriate in interpreting scores if the "test" sample is not representative of the targeted group of children, has an inadequate number of cases, or has a different testing purpose (Salvia & Ysseldyke, 1978).

The reliability of an instrument is a measure of the accuracy and consistency of the scores obtained from the use of the instrument (Goodwin & Driscoll, 1980). Reliability is generally expressed as a correlation coefficient, a numerical indication of the relationship between two variables. A coefficient of 1.0 indicates a perfect relationship between variables, while a coefficient of 0.0 indicates a complete lack of a relationship between variables. Goodwin and Driscoll (1980) suggest that in evaluating

the reliability of tests, "r's in the .30 range and below may be considered low; those in the .40's and .50's, moderate; those in the .60's and .70's substantial; and those .80 or above, high" (p. 52). It should be emphasized that this is an arbitrary interpretation of the significance of reliability coefficients. Categorizing correlational values as low, medium, or high is dependent on the type of relationship under consideration (e.g., test-retest reliability, coefficient of equivalence) and the characteristics of the sample. In evaluating instruments, the types of reliability which should be considered include test-retest reliability, inter-rater reliability, and internal consistency.

The first type, test-retest reliability, is a measure of stability or the consistency of scores over time for the same group of children. This stability is expressed as a correlation coefficient where the closer the coefficient is to 1.0, the greater the stability. As a rule, there is an inverse relationship between the length or time between testing and the test-retest coefficient (Goodwin & Driscoll, 1980).

A second type of reliability, inter-rater, is determined by having two or more judges independently score a set of tests. The correspondence between the scores is expressed as a reliability coefficient. Inter-observer reliability is similar to inter-rater reliability in that it is a measure of the correspondence between independent observations of the same behaviors by two or more observers (Goodwin & Driscoll, 1980; Stangler, Huber & Routh, 1980).

The final type of reliability, internal consistency, is a measure of consistency among tasks in evaluating the content domain of an instrument from the single administration of a test. Split-half reliability is a form of internal consistency and is determined by comparing scores obtained from two equal parts of the same test. The half-test reliability can be used to compute internal consistency reliability for the full test (Goodwin & Driscoll, 1980; Salvia & Ysseldyke, 1978).

The consideration of the adequacy of validity and reliability becomes more complex when considering the criterion-referenced instrument. Historically, validity and reliability have been determined using the Pearson product-moment correlation coefficient which requires interval data. Criterion-referenced instruments are designed to assess whether the child's performance meets criterion levels on given tasks and do not yield interval data; therefore, this particular statistical analysis is inappropriate. However, there are a number of non-parametric analyses which may be utilized. Goodwin and Driscoll (1980) suggest that content validity can be established by a logical analysis of the content of the test, similar to the method used to establish content validity for norm-referenced tests. These authors also suggest that a comparison with other tests measuring similar skills can provide evidence of criterion-related and construct validity. Reliability for criterion-referenced tests can be determined by analyzing the consistency of classifications of children made from criterion-referenced test results rather than from consistency of classifications of children in test scores themselves (Goodwin & Driscoll). In the case of criterion-related validity, construct validity, and reliability, nonparametric methods would be appropriate.

Bibliography. A test manual should contain a list of all references used in the development of the instrument. This information is important in determining the theoretical perspectives on which the test is based. It also provides the reader with the origin of some of the items which is a relevant piece of information since so many instruments tend to build on one another.

Instruments Evaluated

The instruments included in this handbook were chosen due to their availability, coverage of a range of developmental areas, moderate to low cost (with a few exceptions), and their design for use with very young children. The sources for these instruments included: Buros' (Ed.) Tests in Print, 1974; Cross and Goin's (Eds.) Identifying Handicapped Children: A Guide to Casefinding, Screening, Diagnosis, Assessment, and Evaluation, 1977; Johnson's Tests and Measurements in Child Development, 1976; and Southworth, Burr and Cox's Screening and Evaluating the Young Child, 1980.

The instruments in this handbook were evaluated based on the preceding criteria and following an evaluative format designed by Goodwin and Driscoll (1980).

References

Brigance, A. H. Brigance Inventory of Early Development. North Billerica, MA: Curriculum Associates, 1978.

Buros, O. K. (Ed.), Tests in Print: A Comprehensive Bibliography of Tests for Use in Education, Psychology, and Industry. Highland Park, NJ: The Gryphon Press, 1974.

Cross, L., & Goin, K. W. (Eds.), Identifying Handicapped Children: A Guide to Casefinding, Screening, Diagnosis, Assessment, and Evaluation. New York: Walker and Company, 1977.

Goodwin, W. L., & Driscoll, L. A. Handbook for Measurement and Evaluation in Early Childhood Education. San Francisco: Jossey-Bass, 1980.

Johnson, O. G. Tests and Measurements in Child Development: Handbook II (Volume 2). San Francisco: Jossey-Bass, 1976.

Salvia, J., & Ysseldyke, J. E. Assessment in Special and Remedial Education. Boston: Houghton Mifflin Company, 1978.

Stangler, S. R., Huber, C. J., & Routh, D. K. Screening Growth and Development of Preschool Children: A Guide for Test Selection. New York: McGraw-Hill, 1980.

Southworth, L. E., Burr, R. L., & Cox, A. E. Screening and Evaluating the Young Child: A Handbook of Instruments to Use from Infancy to Six Years. Springfield: Charles C. Thomas, 1980.

White, O. R., Edgar, E., & Haring, N. G. Uniform Performance Assessment System. Columbus: Charles E. Merrill Publishing Company, 1978.

Overview of Instruments Evaluated

In recent years, the demand for evaluation measures suitable for the developmentally young child has been increasing. This need for assessment devices has arisen from the growth in preschool/day care programs; recognition of the advantages of early detection and identification of developmental delays; and the implementation of Public Law 94-142 which mandates individualized educational planning for all handicapped children.

Current Status of Instruments

The response to the demand for suitable evaluation measures has been in the form of a large number of recently developed instruments with wide diversity in content, quality, and usability.

Variation in content. Research in child development has produced extensive, detailed information about the development of young children's cognitive, gross and fine motor, language, social, and self-help skills. Unfortunately, many evaluation instruments cover only the "milestones" which are suitable for the detection of gross delays but is not of great value to the teacher in the classroom working with the child on a daily basis. Instead, s/he is concerned with how and where the child is performing within the sequence of behaviors which make up a particular skill. For example, when a child is learning colors, the teacher is interested in whether the child is only matching colors, pointing to the named color, or is able to verbalize the color of an object.

Some instruments evaluate only one or two developmental areas but in great detail; other instruments cover a broad range of areas but include only a few items per age interval; and yet other instruments include all developmental areas and contain a large number of tasks for each age interval.

Technical quality versus usability. The rudimentary state of early childhood evaluation is most dramatically reflected by the issue of technical adequacy.

From a psychometric standpoint, standardization of task administration, normative, validity, and reliability data are not only desirable but essential characteristics of an evaluation instrument (Goodwin & Driscoll, 1980; Salvia & Ysseldyke, 1978). In addition, the manual designed for use with the instrument should include this technical information. Of the instruments evaluated in this handbook, eleven were judged to fulfill the above criteria; however, for the majority of instruments available the manuals provide little or no information regarding their psychometric characteristics. While in some cases it is simply a failure to include the information (e.g., Revised Developmental Screening Inventory, Vineland Social Maturity Scale) it is more often a reflection of the author's failure to conduct any studies to determine whether or not the instrument actually measures what it claims to measure, and yields accurate scores, which may be interpreted in a culturally-fair way (Goodwin & Driscoll, 1980; Salvia & Ysseldyke, 1978). The remaining seventeen instruments

evaluated in this handbook reflect this lack of standardization, reliability, and validation.

Other technical deficiencies for some of the instruments reviewed in this handbook include the computation of derived scores but a lack of a normative sample as a basis for these scores (e.g., Delco-Elfman Developmental Achievement Test, Preschool Language Scale), and in the case of screening instruments, the lack of criteria for referral for more extensive assessment (e.g., Birth to Three Developmental Scale, The Boyd Developmental Progress Scale, Communicative Evaluation Chart, Infant Evaluation Scale, KNI Developmental Scale).

The opposing viewpoint does not necessarily view a lack of technical support as a fatal flaw since the primary focus is simply being able to evaluate the child's current level of functioning in terms of strengths and weaknesses. The argument is that while standardization procedures should be followed in order to accurately compare a child to the original reference group, deviation is sometimes necessary in order to obtain any type of evaluation and especially one that reflects what the child is capable of doing (Bagnato & Neisworth, 1981). This deviation from standardized procedures is particularly necessary in the case of handicapped preschoolers and infants (Bagnato, 1981).

A supporting argument against the use of standardized instruments is that the characteristics of the young child do not appear to be suited to the standardized assessment measure. Infants and young preschoolers have a limited behavioral repertoire, are easily distracted, are not motivated to perform on demand, develop at different rates, and in general, reflect instability in their levels of development (Bagnato & Neisworth, 1981; Dunst & Rheingrover, 1981). These traits are compounded in the handicapped child with the addition of his/her functional disabilities.

For the teacher in the classroom, concerned with developing plans for the children in her care, the technical support for an instrument is of very limited value, if s/he is unable to obtain a "fair" evaluation of a child's skills through the standardized use of that measure.

This conflict between the technical quality and the usability of instruments may account for the current status of evaluation instruments for infants and preschoolers. Some of the seventeen assessment tools reviewed in the second section of instruments in this handbook illustrate this point in that they rate high in usability but are lacking in technical qualities. These measures often claim inherent validity based on developmental literature.

New Trends in Instruments

Since 1975, many of the instruments for use with developmentally young children have been designed to meet the requirements of Public Law 94-142 (e.g., Brigance Diagnostic Inventory of Early Development, Uniform Performance Assessment System) and have taken the form of criterion-referenced or curriculum-referenced measures. These instruments usually contain explicit administration and scoring criteria but have no established validity and reliability beyond informal field testing, for scores or

classifications made using these measures. Criterion-referenced instruments generally cover a broader range of tasks than norm-referenced instruments and provide more information to the teacher than the derived score obtained from the norm-referenced device. However, this development of criterion-referenced instruments appears to be a regeneration of old instruments in new forms. Frequently, the tasks are appropriated from instruments lacking in or with only limited validity and reliability. Two instruments often used as a source of tasks are the Cattell Infant Intelligence Scale (1940) and the Vineland Social Maturity Scale (1965). These two instruments have been available for a relatively long period of time and since publication, subsequent validity and reliability studies have been conducted which have established strong validity and reliability for these measures. However, the original normative samples were limited in size and were unrepresentative of the general population. In the case of the Cattell Infant Intelligence Scale, the standardization sample was composed of 274 middle class children with 35 children in each of eight age groups; "norms" were estimated for the other 12 age groups (Thomas, 1970). The Vineland Social Maturity Scale was standardized in 1935, utilizing a sample of 620 individuals, ranging in age from 0 to 30 years of age (20 per year), and all living in southern New Jersey (Doll, 1965). It would therefore appear that these "new" instruments are being based on a small and, perhaps, weak foundation, thus perpetuating possible inadequacies.

Summary

At the present time there appears to be a large number of instruments available for use in the evaluation of young children in early childhood programs. The presence of technical quality and, simultaneously, practicality in their use by teachers in the classroom is often lacking in many of these measures. Evaluations must be accurate and consistent to a certain degree in order to be of value but should also provide useful information to the teacher for planning purposes.

In view of the abundance of instruments (however limited) available to early childhood professionals, but in recognition of the value of accurate and consistent scores or classifications, it is this author's opinion that rather than discarding these instruments and developing still more evaluative measures, teachers can make a decision to use those instruments which appear to be most practical for their purposes. Subsequently, teachers should develop their own "norms", validity, and reliability for the instruments within their classrooms and make program and placement decisions based on these local "norms."

References

Bagnato, S. J. Developmental Programming for Infants and Young Children. In J. T. Neisworth (Ed.), Topics in Early Childhood Special Education: Assessing the Handicapped Preschooler, 1981, 1, 77-79.

Bagnato, S. J., & Neisworth, J. T. Linking Developmental Assessment and Curricula: Prescriptions for Early Intervention. Rockville, MD: Aspen Systems, 1981.

Doll, E. A. Vineland Social Maturity Scale. Circle Pines, MN: American Guidance Service, 1965.

Dunst, C. J. & Rheingrover, R. M. Discontinuity and instability in early development: Implications for assessment. In J. T. Neisworth (Ed.),

Topics in Early Childhood Special Education: Assessing the Handicapped Preschooler, 1981, 1, 49-60.

Goodwin, W. L. & Driscoll, L. A. Handbook for Measurement and Evaluation in Early Childhood Education. San Francisco: Jossey-Bass, 1980.

Salvia, J., & Ysseldyke, J. E. Assessment in Special and Remedial Education. Boston: Houghton Mifflin Company, 1978.

Thomas, H. Psychological assessment instruments for use with human infants. Merrill-Palmer Quarterly, 1970, 16, 179-223.

Instruments Reporting Validity and Reliability Data

Albert Einstein Scales of Sensori-Motor Development

Authors	Harvey H. Corman and Sibylle K. Escalona
Publisher	Sibylle K. Escalona Rose Fitzgerald Kennedy Center for Research in Mental Retardation and Human Development Albert Einstein College of Medicine Yeshiva University Eastchester Road and Morris Park Avenue Bronx, New York 10461
Date of Publication	Unavailable
Materials Available	Three scales of sensorimotor development (object permanence, prehension, space) Score sheets Films demonstrating the use of the scale are available for sale or rent. Testing equipment is not available from the publisher but is often present in infant programs or should be locally available.

Purposes/Traits Assessed

The Piagetian based measure was designed to assess infants to determine their level of sensorimotor functioning. The Object Permanence Scale assesses the infant's stage of development in the acquisition of the concept of the object. The Prehension Scale assesses the development of hand use (prehension) for Stages II and III of sensorimotor development. The Space Scale assesses the infant's constructs of space a s/he manipulates objects. All three scales assess problem solving skills rather than actual mastery of tasks.

Intended Test Population

The scales are appropriate for use with children functioning in the sensori-motor stage of development, which is generally from one month to two years.

Individual or Group Administration and Time Required

One or more of the scales may be administered to the infant individually. No guidelines are given for determining the length of time for administering the scales. However, like most infant testing, the child should be allowed to rest as needed so that testing occurs only when the infant is in an alert and comfortable state.

Instrument Format

Items on the scales are developmentally sequenced by stages of sensori-

motor development. The Object Permanence Scale includes 25 items and covers stages II through VI; the Prehension Scale includes 24 items and covers stages II and III; the Space Scale includes 34 items and covers stages III through VI.

The Object Permanence Scale is constructed so that "the degree of object permanence is reflected by the infant's capacity to find objects hidden from view" (Corman & Escalona). Behaviors evaluated by the use of this scale include the infant's removal of objects blocking his/her vision, visually following the path of a moving object, finding an object which the child has observed being hidden (visible displacement), and finding a hidden object which was first concealed from the child's view (invisible displacement). The Prehension Scale includes grasping, coordination of gaze and reaching, and swiping behaviors. Responses to items on the Space Scale include the child's searching for hidden objects, following the path of a moving object, releasing objects into a container, and retrieving an object by using another object (e.g., pulling a string, using a stick).

A list of suggested materials is described for each scale. Materials are easily transportable to a testing site and were chosen to be attractive to infants.

Administration Procedures

Training in the use of the scales is not specified in the manual; however, films are available which demonstrate the administration of the Object Permanence and Space Scales. In addition, knowledge of Piaget's sensorimotor stages of development and infant behaviors is essential.

General procedures for administering the items on the scales and the infant responses expected for each item are included in the manuals. Examiner directions are described for each item and the appropriate infant position is described when required. The authors suggest that two examiners be used, one to administer the items and one to score the infant's responses. There is a suggested sequence for administering each scale.

Scoring Procedures

Training for scoring the scales is not specified by the authors. Criteria for scoring are included for each item and stage of development and a definition of mistrials is included.

The highest achieved level of an item must be observed twice before the infant can be credited with the task unless specified otherwise in the directions.

Interpretation Procedures

The manual does not contain any information concerning training in the interpretation of the results of the use of these scales.

A total score is not derived for any of the scales; scores are

recorded in terms of Piaget's stages of sensorimotor development to indicate the infant's level of functioning.

Item Selection

Age norms are reportably available, from the Child Development Project staff, for many items on the Prehension Scale. Standardized procedures and scoring criteria for the Prehension Scale are taken from the Gesell Developmental Examination and the Cattell Infant Intelligence Test.

Items included on the scales were probably adapted from Piaget's Origins of Intelligence in Children, 1952.

Validity

Validity data for the scales are not reported in the manuals.

In a series of cross-sectional validation studies, Guttman scalogram analyses employing Green's index of consistency yielded Consistency Indices of 1.00 for object permanence (N=113), .66 for prehension (N=51), and .98 for spatial relationships (N=83). All of the scales met the criteria of being a true scale. Of 45 infants participating in longitudinal studies, all progressed in the expected sequence of the three scales (object permanence, N=15; prehension, N=14; space, N=16) (Corman and Escalona, 1969).

Reliability

Although reliability data are not reported in the manuals, Corman and Escalona (1969) report that interobserver reliability (three observes for each test of 119 infants) was .94 for the Object Permanence Scale and .95 for the Space Scale. For the Prehension Scale, Escalona (1976) reports interobserver reliabilities of .96 during pilot studies and .97 during spot reliability checks in validity studies.

Cautions and Comments by Authors

The infant must be in an alert state; the examiner should adapt the testing situation to the infant.

For the Prehension Scale, the examiner should avoid being seen by the infant to prevent distracting the infant's attention from the test materials during the testing situation.

Reviewer's Overall Judgment of the Measure

Assets. The description of test equipment is adequate for duplicating the materials. The descriptions of the examiner's actions and the criteria for scoring the infant's responses are explicit, resulting in objective scoring.

The scales are reasonably priced for most early childhood programs and little additional expense would be required for testing materials, as most materials are already present in infant care programs.

Limitations. In the past, there has been an accessability problem with these scales due to a lack of information about where to obtain the scales. At this time they are available from the address provided at the beginning of this review.

Examiners, unfamiliar with Piaget's stages of development, would benefit from a more detailed description of the characteristics of the sensorimotor period as related to object permanence, prehension, and sapce. More information would be helpful in completing the score sheet (e.g., definition of abbreviations, types of scores).

Bibliography

A bibliography is not included in the manuals.

References

Corman, H. H., & Escalona, S. K. Albert Einstein Scales of Sensorimotor Development.

Corman, H. H., & Escalona, S. K. Stages of sensorimotor development: A replication study. Merrill-Palmer Quarterly of Behavior and Development, 1969, 15, 351-361.

Escalona, S. K. In O. G. Johnson's Tests and Measurements in Child Development: Handbook II (Volume 1). San Francisco: Jossey-Bass, 1976.

Bayley Scales of Infant Development

Author	Nancy Bayley Institute of Human Development University of California Berkeley, California
Publisher	The Psychological Corporation 304 East 45th Street New York, New York 10017
Date of Publication	1969
Materials Available	Manual Record forms Test materials including a carrying case

Purposes/Traits Assessed

The norm-referenced assessment instrument is composed of three parts designed to provide a complete evaluation of the infant's development and to provide a means for comparison with other infants of the same age.

Traits assessed by the Mental Scale include: 1) sensory-perceptual acuties; 2) object constancy; 3) memory; 4) problem-solving abilities; 5) vocalization; and 6) ability to generalize and classify.

The Motor Scale evaluates: 1) body control; 2) large muscle coordination; and 3) manipulation skills.

The Infant Behavior Record is a subjective series of rating scales assessing the infant's interaction with his environment, including: 1) attitude; 2) interest; 3) emotion; 4) energy; 5) activity; and 6) approach/withdrawal from stimulation.

Intended Test Population

The BSID was designed for normal and exceptional children functioning in the birth to 30 month developmental age range.

Individual or Group Administration and Time Required

The measure is individually administered with an average testing time of approximately forty-five minutes. Additional time is required for computing the scores and completing the Infant Behavior Record.

Instrument Format

Test materials are standardized and are included in the kit with the exception of a stopwatch, 8½" x 11" paper, facial tissues, a standard set of stairs, and a standard walking board. Test materials are attractively designed and most are durably constructed. Replacement items are available from the publisher. With the exception of the stairs and the walking board,

all testing equipment can be transported to the test site in the carrying case included in the kit.

The directions for administering the mental and motor items and the tasks on the record forms are developmentally sequenced. The mental and motor subscales are color coded in the manual, and on the record forms. Items which require the same materials or involve similar observations are arranged by situation in two tables in the manual.

The Mental Scale consists of 163 tasks; behaviors elicited by items on the Mental Scale include responding to sound, visually regarding stimuli, manipulating small objects, completing formboards, demonstrating pre-writing skills, labelling stimuli, and following directions.

There are eighty-one fine and gross motor tasks on the Motor Scale including grasping, sitting, standing, walking, throwing, jumping, and balancing skills.

Administration Procedures

The Bayley is generally administered by a certified psychologist. Training of examiners should be conducted by someone who is familiar with the BSID and is familiar with testing young children. Trainees should have a background in "the theory of measurement and the interpretation of test results" (Bayley, 1969, p. 5). Trainees, in groups of three to six, should begin by observing demonstration tests. Each trainee then tests several infants, of varying ages, under supervised conditions. After this, Bayley suggests that further practice involve the testing of one or more infants at each month from 0 to 12 months and at each 3-month age level from 15 to 30 months. A potential examiner must be able to motivate infants and control the testing situation so that the child does not lose interest or become frustrated.

Administration directions provided in the manual include: the name of the item/skill, the situation code and order of difficulty, standardized directions for presenting the item, cautionary notes, criteria for passing the task, and age placement (50% passing) and age range (5% to 95% passing).

Evaluation on each scale begins with items in age placements about one month below the chronological age of the child unless other evidence is present to indicate a different starting point. The basal level is the item preceding the earliest failure (earliest in terms of age placement) and the ceiling is the item representing the most difficult success. Ten successive items passed or failed should be sufficient to establish basal levels or ceilings for the Mental Scale and six successive items for the Motor Scale.

The Mental Scale is generally presented before the Motor Scale; however, the sequence for administering the items should be adapted to the infant's responsiveness. Tasks on the Motor Scale which use the same materials as Mental Scale items (indicated by a dagger on the score form) may be presented during the mental evaluation.

Scoring Procedures

Scoring is generally done by psychologists who are very familiar with tests and measurements. The manual describes the scoring procedure in detail.

Separate record forms are used for the Mental and Motor Scales. The format of each form consists of biographical data, space to record the raw scores and developmental indices, item number, age placement and range, situation code, item title, score, and notes about the child's performance on each item.

Basal and ceiling items are circled and the words "basal" and "ceiling" are written in the margin next to the corresponding item. Each task tested is scored as passed, failed, omitted, refused by the child, or reported by the mother but not observed during the testing session. Only passes are credited when computing the raw score.

Behavior exhibited at any time during the testing situation may be credited on the record form. Because the items are developmentally sequenced, a child may fail an item because s/he has attained a higher developmental level and no longer exhibits the earlier behavior. When a child "surpasses" an item s/he should be credited with passing this earlier skill.

The Infant Behavior Record (IBR) is to be completed after the child has left the testing situation. A summary of characteristic behaviors for each age group in included in the manual.

Raw scores (the total number of passes) for the Motor and Mental Scales are recorded on the face of each record form. The child's chronological age (months and days) is used to convert the raw scores to the Mental Development Index (MDI) and the Psychomotor Development Index (PDI). Tables are included in the manual for converting the scores.

Interpretation Procedures

The manual provides guidance in interpreting the resulting test scores. The value of the developmental indices is in establishing the child's current level of functioning and in determining the significance of any deviation from age norms. The indices are not to be used for predicting later abilities.

The mean value, for each age level, is 100 with a standard deviation of 16. The standard score range of 50 to 150 covers more than three standard deviations on either side of the average MDI or PDI for each age. Performance by an individual child can be compared with the standardized scores of his age peers. In addition, a difference of 20 points between a child's MDI and PDI is considered important.

A table is included in the manual which illustrates the ratings for the IBR standardization sample. The table may be used as a guide for interpreting behaviors observed during the evaluation.

Item Selection

Items on the Mental and Motor Scales are based predominantly on: The California First-Year Mental Scale (Bayley, 1933), the California Preschool Mental Scale (Jaffa, 1934), the California Infant Scale of Motor Development (Bayley, 1936), and items selected from the 1958-1960 version of the Mental and Motor Scales which were tested on 1,400 children from 1 to 15 months of age and 160 children from 18 to 30 months of age.

The Infant Behavior Record is based on rating scales used to derive the old California Mental and Motor scores and a 1958 version of the behavior rating scales used on a sample of 1,350 children.

Norms for the current Mental and Motor Scales (1969) were developed from a sample of 1,262 children ranging in age from 2 to 30 months. Testing was done at 14 age levels. Children could not be more than 1 months premature, had to be considered "normal," and had to be living at home. Subjects were chosen to meet the proportions of children as described in the 1960 United States Census of Population for such factors as sex, race, urban-rural residence, and education of head of household. Demographic data are reported in the manual. (Rural children tended to be underrepresented in the actual sample).

Validity

Concurrent validity was determined by comparing scores from the Mental Scale with IQ scores from the Stanford-Binet for a sample of 120 California children (24, 27, and 30 months of age). The coefficients of correlation ranged from .47 to .64 with an overall r=.57.

The relationship between Mental and Motor Scale scores is illustrated in the manual for each of the 14 age groups in the standardization sample. Raw score coefficients ranged from .24 to .78 and standard score coefficients ranged from .18 to .75. The median coefficient across age groups for both raw scores and standard scores was .46.

Reliability

The manual includes data on the split-half reliability coefficients for both the Mental and Motor Scales for each of the 14 age groups in the standardization sample. The reliability coefficients for the Mental Scale had a median value of .88 (range .81 to .93). The reliability coefficients for the Motor Scale had a median value of .84 (range .68 to .92).

Werner and Bayley (1966) reported a tester-observer reliability study using the 1958-1960 version of the scales with 90 8-month-old infants. The mean of the percentages of agreement was 89.4 for the 59 items on the Mental Scale which fell in the 6 to 12 months age range and 93.4 for the 20 tasks on the Motor Scale covering the same developmental range. The same study reported test-retest data for 28 infants (8 months of age, 1 week test interval). The mean percentage of agreement for the 59 mental tasks was 76.4 and 75.3 for the 20 motor items.

Cautions and Comments by the Author

Accurate interpretation of test results is dependent upon obtaining the child's exact age. In addition, any change in the standardized administration of the test items will invalidate scoring results.

The scores derived on the BSID do not support a computation of an intelligence quotient.

Reviewer's Overall Judgment of the Measure

Assets. The manual includes adequate data to establish reasonable reliability and validity for the instrument.

The standardization sample is representative of the U. S. population, which makes the instrument useful for a wide range of children.

Standardization of test materials contributes to the validity and reliability of the test. Replacement materials are available from the publisher at a reasonable cost, materials are attractive, most materials are durably constructed, and materials (except for the stairs and walking board) are easily transported to the testing site in the carrying case provided with the kit.

The color coding of the directions for administering the tasks, the record forms, and the norm tables in the manual should reduce the possibility of recording error. The listing of testing materials by age level should decrease the amount of time required in preparation for testing.

Testing time requirements are short enough that, in most cases, the child should not become too tired or frustrated. The age placement and age range for each scale item provides a guide for the examiner in establishing a basal level and a ceiling. The presence of the mother during testing should contribute to the child's feeling of being at ease during the evaluation.

Limitations. Usability of this instruments is restricted in the sense that some states will not accept results from the BSID unless it is administered and scored by a certified psychologist.

The carrying case has been designed to store testing materials in loose form. A sectioned design, with a specific space for each item, would make it easier to locate materials during the testing situation and make it easier to notice missing materials.

Bibliography

A list of 20 references is included in the manual; included among these are the earlier instruments developed by Bayley and by Jaffa at Berkeley in the 1930's. Additional references include the use of this instrument with children in the 1960's.

References

Bayley, N. Bayley Scales of Infant Development. New York: The Psychological Corporation, 1969.

Werner, E. E., & Bayley, N. The reliability of Bayley's Revised Scale of Mental and Motor Development during the first year of life. Child Development, 1966, 37, 39-50.

Denver Developmental Screening Test (DDST)

Authors	William K. Frankenburg, M.D. Josiah B. Dodds, Ph.D. Alma W. Fandal University of Colorado Medical Center Boulder, Colorado	
Publisher	LADOCA Project and Publishing Foundation, Inc. East 51st Avenue and Lincoln Street Denver, Colorado 80216 (303) 629-6379	
Date of Publication	1975	
Materials Available	Reference Manual (revised 1975 edition) Manual/Workbook for Nursing and Paramedical Personnel Instructional Module: manual/workbook, film, practice assignments, and proficiency test. Test forms (package of 100) Preassembled kit of materials	 $11.00 $ 4.00 $12.00

Purposes/Traits Assessed

The DDST is a screening tool intended for the detection of developmental delays in young children. The 105 tasks are arranged in four developmental categories: personal-social, fine motor/adaptive, language, and gross motor.

Intended Test Population

The test was designed for use with children from birth to 6 years of age, but it is limited in its utility in the 5 to 6 year age range.

Individual or Group Administration and Time Required

An individually administered test, it may be quickly administered in 15 to 20 minutes (Frankenburg, Goldstein, and Camp, 1971).

Instrument Format

Administration directions for individual tasks are developmentally sequenced in each of the four categories. Individual items in the four areas are represented on the score form by bars arranged along an age continuum. The percentage of normal children passing each item (25%, 50%, 75%, 90%) is represented by points along the bar.

The standardized test materials will be familiar to most children and are easily transported to the test site in the small carrying case provided.

The back of the score form includes additional stimulus items (pictures for identification) and notes for administering some of the tasks. Space is provided at the bottom of the form for notes about the child's behavior during the screening session.

In the personal-social category (23 items), the child behaviors evaluated include response to a stimulus, grooming and feeding skills, and peer and adult interaction skills. Fine motor/adaptive developmental responses (30 items) include drawing and fine motor manipulation. The 21 tasks in the language category are evaluated through response to sound, imitation of sounds, pointing, following directions, and expressive language skills. Gross motor responses (31 items) require the child to exhibit sitting, standing, walking, and balancing skills. Tasks which the child refuses to do or require extensive observation may be scored by parental report.

Administration Procedures

The "Instructional Module" was designed for self instruction in the use of the DDST. The manual/workbook guides the trainee by the use of examples, references to the film, practice questions, and practice tests, in addition to the administration and scoring directions provided for evaluating the child. After familiarization with the manual, the trainee is advised to practice testing twelve children.

The first step in administering the DDST is to calculate the child's chronological age (examples provided in the manual) and plot the age line on the score form. Testing begins below the child's chronological age and continues until the child fails three consecutive tasks in that area or until all tasks, through which the age line passes, have been evaluated. In addition, each sector should contain at least three passed and three failed tasks.

The child is allowed three trials to perform each task. To help the child feel at ease with the tester and to avoid having the child become too active before the end of the session, it is suggested that personal-social tasks be presented first, followed by fine motor, language, and gross motor tasks, respectively.

Scoring Procedures

Self-instructional training, including examples and practice tests, is provided in the manual/workbook.

Criteria for passing each task are provided in the manual. Each task administered is scored as a pass, failure, refusal, or no opportunity to perform the skill. On the score form, an "R" beside the task indicates it may be scored by parental report. A number beside the task refers the examiner to notes on the back of the score form. Observations of the child's behavior are recorded on the back of the score form. Delays (task failures to the left of the age line on the score form) are "emphasized by coloring in the right end of the bar of the delayed item" (Frankenburg, Dodds, & Fandal, 1975, p. 23).

Re-evaluation may be scored on the same form by using a different colored pen.

Interpretation Procedures

Example test forms, practice score forms, and interpretation criteria are included in the manual/workbook for self-instructional training in interpreting screening results.

Results, based on the number of delays, are interpreted as normal, questionable, abnormal, or untestable. Rescreening, in two or three weeks, is indicated for those children evaluated as abnormal, questionable, or untestable. Referral for diagnosis is recommended if a "normal" interpretation is not made at the time of the rescreening.

Item Selection

A total of 240 items were selected from over 12 developmental tests and preschool intelligence tests (Frankenburg & Dodds, 1967) based on the simplicity of testing materials required, ease of administration, and clarity of required task response. The 240 items were administered to a sample of 200 children (data unreported) and the number of items was reduced to 105, based upon their ease of administration, clarity in scoring, and relatively narrow distribution (i.e., age range for children passing the item was relatively narrow).

With high risk children eliminated, a sample of 543 males and 493 females, aged 2 weeks to 6.4 years, from the Denver population was tested to standardize the DDST. Children were selected to represent proportionately the demographic characteristics of the 1960 Denver census population, although families with the father holding a professional, management, or sales position were over-represented slightly.

The percentage of children (25, 50, 75, 90) passing each task is reported in the manual for the total sample, by sex, and by occupational group of the father.

Validity

In a preliminary study, Frankenburg and Dodds (1967) compared the scores of 18 children (4 to 68 months) on the DDST and the Revised Yale Developmental Schedule and found a correlation of .97. They reported that no child was rated "normal" on the DDST and "retarded" or "borderline" on the Revised Yale Developmental Scale.

A concurrent validity study using the DDST, the revised Bayley Infant Scale (for children under 2 years) and the Stanford-Binet Form LM (for children over 2 years) is reported in the manual and by Frankenburg, Goldstein, and Camp (1971). Numerical inconsistencies exist between the information in the Frankenburg et. al. (1971) article and the manual and between the table and test regarding validity in the manual itself. Of 2,000 children (1 month to 6 years) screened by the DDST, 236 children (237 reported by Frankenburg, Goldstein & Camp) were evaluated from 1 to 3 weeks later, using either the Bayley or the Stanford-Binet. Using the revised screening method for the DDST, copositivity (abnormal rating on the screening test and reference test) was .73 and conegativity (normal rating on the screening test and reference test) was .92. The rate of overreferral was 7.2% and the rate of underreferral was 2.95%. (Reported as .68, .92, 11%, and 3% respectively, by Frankenburg, Goldstein & Camp).

A concurrent validity study of 246 children from a sample of 1,292 was also conducted by Frankenburg, Goldstein, and Camp (1971). After screening on the DDST, 116 children were administered the Revised Bayley Infant Scales and 130 were given the Stanford-Binet. Copositivity was reported as .92, conegativity was .97, the rate of oyerreferral was 3.2% and the rate of underreferral was .4%. Goodwin and Driscoll (1980) calculated an overreferral rate of 40%; Stangler, Huber, and Routh (1980) calculated a 3% rate of overreferral in the total sample and a 40% rate of overreferral for the total number of referrals.

Lichtenstein (1981) conducted a concurrent validity study of 40 preschool children given the DDST, the Developmental Indicators for the Assessment of Learning (DIAL), the Stanford-Binet Intelligence Test, the Peabody-Picture Vocabulary Test, and the Woodcock-Johnson Psycho-Educational Battery. The 24 boys and 16 girls were from lower and middle class homes, ranged in age from 43 to 59 months, and were 65% white, 22% black, and 13% Asian, Hispanic, and Native American. Testing took place within a 3-week interval. A correlation of .82 was obtained between DDST and DIAL total scores (.52 between gross motor sections, .65 between fine motor sections, and .31 between DDST language and DIAL communication). The total number of passed items on the DDST were then summed to obtain "raw score totals." The DDST raw scores and the Stanford-Binet mental age scores correlated at .65; for DDST raw scores and the PPVT raw scores, r=.58; for DDST raw scores and the Woodcock-Johnson preschool scale of cognitive ability scores, r=.54; for DDST raw scores and the Woodcock-Johnson knowledge cluster scores, r=.54; and the DDST raw scores and the Woodcock-Johnson skills cluster scores, r=.39.

A predictive validity study by Camp, van Doornick, Frankenburg, and Lampe (1977) was conducted to determine whether the DDST could predict later "school problems." A group of 65 children (4 to 6 years of age at the time of testing) from lower class homes were followed up 3 years after testing. School problems were identified for 88% classified as "abnormal" on the DDST, 66% classified as "questionable" and 32% classified as "normal."

Reliability

In a study of test-retest reliability (reported in the manual), the same examiner administered the DDST to 20 children (2 months to 5½ years), twice at a 1-week interval. The percentage of agreement of scores for individual tasks ranged from 90% to 100% with a mean of 95.8%.

A sample of 186 children (1.5 months to 76 months), was tested twice, with a 1-week interval between testings, to determine test-retest reliability (Frankenburg, Goldstein, & Camp, 1971). The percentage of agreement of the ratings of normal, questionable, or abnormal between the first and second tests was 97%.

A study of interrater reliability is reported in the manual where 48 children (from the original standardization sample of 1,036) were screened twice by two of four examiners. The percentage of agreement on scoring items ranged from 80% to 95% with a mean of 90%.

Frankenburg, Camp, VanNatta, and Demersseman (1971) also conducted a

study of tester-observer reliability. The sample was composed of 76 children (38 boys and 38 girls), aged .4 months to 77.4 months. Each child was given the DDST by two examiners (testing time divided approximately in half so that each served as examiner and observer). Tester-observer agreement ranged from 81% to 100% with a mean of 97.7% for those tasks passable by report and ranged from 86% to 100% with a mean of 96.1% for items not passable by report.

Cautions and Comments by the Authors

The DDST is a screening test and should not be used as an intelligence test. Items are standardized and any deviation in administration or scoring will invalidate the results.

Although the test was standardized using a population from a restricted area, the authors suggest it is suitable for use in a wide geographic area.

Actual observation is preferred to parental report, even when parental report is permissable. A refusal should be scored only if the examiner feels the child is capable of performing the task; otherwise it is scored as a failure.

Accuracy in placement of the chronological age line is critical due to the dependency of interpretation results upon the correct location of the line.

Reviewer's Overall Judgment of the Measure

Assets. The guidelines provided in the manual for administration, scoring, and interpretation allow any professional working with young children to use the screening test without extensive training.

A revised score form has been recently developed to be used when a shorter scoring session is desirable. It is quite similar to the old version but age line intervals are vertically located rather than horizontally placed on the form. Labels on horizontal item bars are slightly smaller and therefore somewhat harder to read. The ascending pattern of the bars has no bearing on the test results but may influence the parent's perception of the scores recorded by the examiner during the screening session.

Limitations. Due to the restricted geographical area in which norms, validity, and reliability were established, the screening test should be carefully evaluated for use in other areas.

Bibliography

A bibliography is not included in the manual.

References

Camp, B. W., van Doornick, W. J., Frankenburg, W. K., & Lampe, J. M. Preschool development testing in prediction of school problems. Clinical Pediatrics, 1977, 16, 257-263.

Frankenburg, W. K., Camp, B. W., VanNatta, P. A., & Demersseman, J. A. Reliability and stability of the Denver Developmental Screening Test. Child Development, 1971, 42, 1315-1325.

Frankenburg, W. K., & Dodds, J. B. The Denver Developmental Screening Test. The Journal of Pediatrics, 1967, 71, 181-191.

Frankenburg, W. K., Dodds, J. B., & Fandal, A. W. Manual/Workbook for Nursing & Paramedical Personnel. Denver: LADOCA, 1975.

Frankenburg, W. K., Goldstein, A. D., & Camp, B. W. The revised Denver Developmental Screening Test: Its accuracy as a screening instrument. The Journal of Pediatrics, 1971, 79, 988-995.

Goodwin, W. L., & Driscoll, L. A. Handbook for Measurement and Evaluation in Early Childhood Education. San Francisco: Jossey-Bass, 1980.

Lichtenstein, R. Comparative validity of two preschool screening tests: Correlational and classification approaches. Journal of Learning Disabilities, 1981, 14, 68-72.

Stangler, S. R., Huber, C. J., & Routh, D. K. Screening Growth and Development of Preschool Children: A Guide for Test Selection. New York: McGraw-Hill, 1980.

Developmental Profile

Authors	Gerald D. Alpern and Thomas J. Boll	
Publisher	Psychological Development Publications P. O. Box 3198 Aspen, Colorado 81611	
Date of Publication	1972	
Materials Available	Manual Scoring and Report forms (package of 10)	$14.25
	No testing materials are required.	

Purposes/Traits Assessed

The norm-referenced profile was designed for use as a screening instrument to evaluate five areas of development: physical, self-help, academic, social, and communication skills.

Intended Test Population

The instrument can be used to screen normal children from birth to nine years of age and developmentally handicapped children to the time when they are performing at the eleven-to-twelve year-old skill level.

Individual or Group Administration and Time Required

The profile is completed by conducting a parent interview (child observation is used when the parent is unable to answer). A time frame of 20 to 40 minutes is suggested to administer and score all five subscales when interviewing the parent of an "average" 3 to 9 year old. If a younger or older child is being evaluated, time requirements will be less. If only a part of the profile is administered, the examiner can plan on one-fifth of the total time for each subscale used.

Instrument Format

A total of 217 tasks were chosen to evaluate development on the five subscales and they are arranged by age intervals. A 6-month age interval is used from birth to 3½ years and a 1-year interval for children 3½ years and older. There are generally three items per age range for each subscale (sometimes only two items) for a total of 15 evaluative tasks at most age intervals.

The spiral binding of the manual makes it easy to use during the interview. Arrangement of tasks by age, both in the manual and on the score form, decrease the amount of time required for the interview. Color coding of the administration directions and the subscale score forms facilitates locating specific categories and helps prevent accidental use of the wrong score form. Extra care is necessary when using the subscale forms due to the small amount of space available for scoring individual tasks.

All of the tasks on the profile may be scored by parental report. Items on the physical subscale include assessing the child's ability to move about in his/her home, use scissors, and throw and catch. Skills that are evaluated on the self-help scale are grasping, dressing, feeding, grooming, and independent living skills. The social scale includes attention getting behaviors, following the rules of a game, peer interaction skills, and awareness of the emotions of others. Tasks on the academic scale include an understanding of object permanence, pointing to or naming objects, counting skills, and reading skills. Tasks which evaluate the child's development of communication skills include imitation of sounds and words, naming objects, singing songs, and rhyming words.

Administration Procedure

While basic interviewing skills are desirable for administering the profile, specialized training is not necessary; the authors of the manual suggest that three to five practice interviews be conducted.

Administration begins with completing the biographical section on the score form and determining the child's chronological age (chart provided in the manual). The child's parent may be interviewed or a teacher can self administer the profile. If the interview technique is used, questions may be rephrased as long as the meaning of the questions does not change.

Two methods of administration may be used in completing the profile. In the regular method all tasks, beginning with the youngest level, are administered. The "short cut" method may be used where a double basal level is established (the child passes all items at two consecutive age levels) and screening continues until the child fails all items at two consecutive age levels (double ceiling).

Scoring Procedures

Specialized training in scoring the profile is not required. Criteria for scoring each task are included.

Basal credit is recorded for each subscale on the appropriate record form; all additional credits (sum of months credit beyond basal level) are listed in the additional credit space. Summing the basal and additional credit scores results in the subscale score ages, which are recorded on the front page of the "Scoring and Report Form."

Interpretation Procedures

Training procedures are not included in the manual for interpreting results, although guidelines (tables for each scale) are provided for referral of a child for more complete diagnostic evaluation.

Delays or advances may be determined by comparing the child's chronological age with each of the child's subscale developmental ages. Tables are included for use as a guideline in determining whether a difference can be considered normal, in the borderline range of potential significance, or in the danger range (requiring referral).

Intelligence Quotient Equivalency (IQE) scores (determined by dividing

the child's Academic Age, expressed in months, by the child's chronological age, expressed in months, and multiplied by 100) are not valid for those children not represented in the standardization population.

Item Selection

The original 318 tasks were selected from a review of actual scales in print and child development normative data, or developed as original items designed to reflect profile concepts.

A standardization study was conducted to identify those tasks which 1) were passed by 75% of the children at one age level, but were generally failed by most children in the preceding age interval and passed by most children in the following age interval; 2) were nondiscriminatory by sex, race, or social class; and 3) had a high percentage of agreement between parental report and actual performance by the child. This study resulted in the final 217 tasks included in the profile. The sample was composed of 3,008 "normal" children (1,527 boys and 1,481 girls) with a minimum of 200 subjects for each year from birth to 12 years. Subjects resided in the states of Indiana and Washington. Demographic data and item analysis of the percentage of children (by sex, race, and social class) passing each item for each subscale are included in the manual. The racial composition of the sample was 84% white, 14% black, and 2% other (comparable to 1970 national census statistics); 80% were from middle class homes, 9% from lower class, and 11% from upper class (lower class was more heavily represented in the birth to 2 year age intervals); 91% were from the state of Indiana, 9% from Washington; 89% were residing in areas with a population greater than 25,000, 9% with a population of 2,500 to 25,000, and 2% with a population under 2,500.

Validity

The authors conducted a study comparing parental report using the profile with actual performance by the child. The study involved 100 children (55 boys and 45 girls), aged 3 months to 12 years. Of the 217 tasks on the profile, 116 were judged suitable for interview and observation techniques. The percentage of agreement between the mother's report and actual skills demonstrated by the child on these 116 items averaged 86% (physical area 87%, self help area 88%, social area 87%, academic area 86%, communication area 84%).

Scores obtained from the Stanford-Binet Intelligence Test were compared with the IQE scores from the academic scale of the Profile for 16 "normal" children (10 boys and 6 girls, 4 to 10 years of age). The average IQ was 119 (range 87 to 148), while the average IQE was 108 (range 84 to 131) with a correlation between the scores of .49; 50% of the subjects had IQ and IQE scores within ten points of each other.

For a group of 54 "retarded" children (30 boys and 24 girls, 2½ to 11 years), an average score of 51 (range 20 to 103) was obtained on an individually administered intelligence test (not identified). Although scores obtained (from teacher interview) on the Developmental Profile academic scale were not reported in the manual, a correlation coefficient of .85 was reported between the IQ and IQE scores for all subjects.

Reliability

A study of inter-observer reliability was conducted with 35 Headstart preschool teachers (untrained in the use of the profile). The teachers observed an interview with the mother of a 4½ year old boy, covering 30 of the 71 items on the academic scale. Seventy-one percent of the teachers scored the results identically, 89% scored within one point of the interviewer's score, and 100% were within two points.

A test-retest reliability study involved 11 mothers (children ranged in age from 11 months to 9 years, 11 months), interviewed twice at two to three day intervals. Identical scores were obtained for 22%, scores within one point of each other were obtained for 50%, 68% of the scores were within two points of each other, and 92% were within three points. The average difference among all scale scores for all subjects was 1.74 points.

Cautions and Comments by the Authors

The development of black and white urban children is primarily reflected by the standardization sample; use with rural children and other racial groups is not recommended at this time.

The IQE appears to err in the direction of underestimating the IQ of above-average children and overestimating the IQ of below-average children.

When screening a child, the lack of a double ceiling means the test can not measure the child's developmental age but the child is "at least" at the level indicated for that scale.

Rather than using the "short cut" method, the regular administration procedure should be followed when a handicap is suspected.

Reviewer's Overall Judgment of the Measure

Assets. Use of the profile may be quickly learned by a wide range of professionals for use with a broad age range of children. While the standardization sample is biased toward black and white urban children living near the states of Indiana and Washington, adequate demographic information is included in the manual for judging the appropriateness of the use of the profile with other groups of children.

The parent interview format requires only limited time to administer the profile and may increase parental involvement/committment to the program.

Limitations. Validity and reliability studies are limited by the small number of subjects involved. The physical scale tasks appear to favor the upper end of the age intervals (e.g., 2½ to 3 years: hop on one foot for five feet). Scores on the academic scale are heavily dependent on local school curricula in the upper age intervals.

Use of the profile is further weakened by the tendency of IQE scores to inaccurately reflect the abilities of other than "average" children and by the limited number of tasks included on the academic scale used for

deriving this IQE score.

Bibliography

A list of references is not included in the manual for the profile.

References

Alpern, G. D., & Boll, T. J. Developmental Profile. Aspen, Colorado: Psychological Development Publications, 1972.

Developmental Programming for Infants and Young Children

Editors	D. Sue Schafer and Martha S. Moersch	
Publisher	The University of Michigan Press University of Michigan Ann Arbor, Michigan	
Date of Publication	1977	
Materials Available	Assessment and Application (Volume 1) Early Intervention Developmental Profile (Volume 2) Stimulation Activities (Volume 3)	$14.50
	Test materials, although extensive in number, should be locally available and are generally found in infant/preschool programs.	

Purposes/Traits Assessed

The criterion-referenced instrument was designed for infant assessment and program planning. The 274 tasks evaluate development in the areas of perceptual-fine motor, cognition, language, social-emotional, self-care, and gross motor skills.

Intended Test Population

Assessment items are designed for use with children functioning in the birth to 36 months developmental age range.

Individual or Group Administration and Time Required

The profile may be administered to the individual child in less than one hour.

Instrument Format

Volume One includes the history and goals of the measure, contains directions for administration, the criteria for scoring the profile, reports on validity and reliability studies, guidelines for program planning, and a list of the reference material used in developing the instrument.

Volume Two contains general instructions in the use of the profile, a list of test materials, and the score form.

Volume Three is a collection of activities designed for use in the instructional program by parents or in the classroom.

Tasks, in the directions for administration and on the profile, are divided into developmental areas and are arranged by age intervals.

Responses to tasks are varied; for the perceptual/fine motor area, the child manipulates objects, demonstrates pre-writing skills, builds

with cubes, and cuts with scissors. Problem-solving skills, matching, and pointing to demonstrate understanding are included in the tasks on the cognitive scale. Response to sound, imitation of sounds, demonstration of understanding through pointing or following directions, labelling, and the use of different parts of speech are the response modes in the area of language. Social-emotional tasks are assessed by the child's attendance to stimuli; interactions with adults and peers, and expression of emotions. The child demonstrates feeding, toileting, and dressing skills in the self-care category. Gross motor tasks require the assessment of reflexive behaviors and skills in sitting, standing, walking, and climbing.

Administration Procedures

"The profile was designed to be administered by a multidisciplinary team which includes a psychologist or special educator, physical or occupational therapist, and a speech or language therapist" (Brown and Donovan, 1977, p. 6). Training includes familiarization with the profile, a team evaluation of a normal infant, and the evaluation of several handicapped infants or preschoolers.

After a brief observation period, items are administered from the fine motor and cognitive categories, followed by language, social-emotional, and self-care; gross motor tasks are administered last.

Scoring Procedures

Additional training is not required for scoring; scoring criteria for each item are included in Volume One.

Task performance is scored as either a pass, a pass-fail if the skill is emerging, or a fail if the skill is not demonstrated or if the task is omitted.

The basal level is established at the age range preceding the child's earliest failure and the ceiling level at the point where the child fails either six consecutive items or all items in two consecutive age ranges.

The child's performance in each category may be plotted on the "Profile Graph" to indicate the child's strengths or to indicate general developmental levels (example provided in the manual).

Interpretation Procedures

Results of the evaluation are to be used in developing activities in the child's instructional plan. No training is provided in interpreting results, but it is suggested that specialists in child development and therapists provide ongoing consultation.

Item Selection

Tasks must have appeared in at least two other recognized scales (specific references not cited for individual items), or be an original item based on current developmental theory. Items had to be developmental milestones for the age intervals on the profile and be representative of all developmental aspects for that category. Item assignment to specific

age intervals was based on research from other instruments and standardizations of other instruments. "The developmental-profile materials were constructed and field tested with both nonhandicapped and developmentally disabled infants." (Bagnato, 1981).

Validity

Concurrent validity is reported in the manual. Category scores for 7 to 14 children (characteristics unknown) were correlated with scores on the Bayley Mental Scale, the Bayley Motor Scale, the Vineland Social Maturity Scale, REEL, and a clinical motor evaluation. Correlation coefficients ranged from .33 to .96; highest coefficients were between scores representing the same developmental categories (e.g., .95 for gross motor and Bayley Motor, .96 for cognition and Bayley Mental).

Additional validity studies were not found in developmental literature.

Reliability

Reported correlation coefficients among the six profile scale ranged from .59 to .95.

Inter-observer reliability ranged from .80 to .97 with a mean of .89 for nine trained observers evaluating three videotaped profile assessments.

Correlation coefficients for test-retest reliability (scores for 15 at three, 3-month intervals) ranged from .93 to .98 at 3 months and .90 to .97 at 6 months.

Studies concerning the reliability of this instrument were not found in current literature.

Cautions and Comments by the Authors

The profile should not be used for diagnosis of handicapping conditions nor for predicting future capabilities.

The number of passes and failures may be condensed to a single age level for each category for comparison across categories. However, this reduces the amount of information available for program planning.

Reviewer's Overall Judgment of the Measure

Assets. This is a comprehensive plan for assessment and program planning for very young children.

Limitations. Reported results on validity and reliability studies are limited in value due to the small number of subjects involved and the lack of descriptive information about those subjects.

A physician or physical therapist should be part of the evaluation team due to the large number of reflexive behaviors included for evaluation on the profile.

Bibliography

A list of forty-five references is included in Volume One.

References

Bagnato, S. J. Developmental Programming for Infants and Young Children. In J. T. Neisworth (Ed.), Topics in Early Childhood Special Education: Assessing the Handicapped Preschooler, 1981, 1, 77-79.

Brown, S. L., & Donovan, C. M. Stimulation Activities, Volume 3. Ann Arbor: University of Michigan Press, 1977.

Rogers, S. J., D'Eugenio, D. B., Brown, S. L., Donovan, C. M., & Lynch, E. W. Early Intervention Developmental Profile, Volume 2. Ann Arbor: University of Michigan Press, 1977.

Infant Evaluation Scale (Six Weeks to Six Months of Age)

Authors	William Gingold, Phyllis Gingold, Kay McDonald, George Flamer, and William Hoehle
Publisher	Red River Human Services Foundation Publications Division 15 Broadway, Suite 510 Fargo, North Dakota 58126
Date of Publication	1978
Materials Available	Manual Score Sheets Preassembled test materials are not available from the publisher but should be locally available.

Purposes/Traits Assessed

This criterion-referenced screening instrument is meant to be used by parents to provide a systematic report of their child's performance on psycho-motor and cognitive tasks to professionals.

Intended Test Population

The scale was designed for use with infants functioning between the ages of six weeks and six months of age.

Individual or Group Administration and Time Required

The child is evaluated individually when s/he is in a cooperative mood; therefore, testing time will vary according to the child's needs and interest.

Instrument Format

The infant is evaluated on fifty-two tasks which include reflexive behaviors (e.g., palmar grasp, sucking), fine motor skills (e.g., picking up small objects), gross motor skills (e.g., crawling, walking, standing), problem solving skills (e.g., uncovers toy, removes cube from cup), and vocalization (e.g., babbling).

The score sheet consists of one page directions to the parent completing the form, followed by the items, which are listed in random order to avoid possibly suggesting to parents items which their infant should be passing.

Materials used in observing the infant are often found in the home and will therefore be familiar to the parent and child.

Administration Procedures

The scale has been designed for completion by the parent by following the one page of directions on the score sheet and the instructions provided for each task.

The procedure for administering/observing the infant is described beside the task, to be scored, on the score form.

The parent is asked to score all items on the score form with the reassurance that the infant is not expected to pass all the items.

Scoring Procedures

The parent scores the tasks using the criteria accompanying the directions for administering the items.

Parents are asked to score the infant on each task as having mastered it, not mastered it, or performing poorly due to environmental interference; a total score is not computed.

Interpretation Procedures

Parent-scored forms are to be reviewed by pediatricians, clinicians, and other child-oriented professionals in forming an evaluation of the infant's current developmental status. Guidelines, for these professionals, in interpreting the scores are not included in the manual.

Item Selection

The manual does not report the criteria used for including these fifty-two tasks on the scale; however, many of the items appear to be taken from the Bayley Scales of Infant Development.

Validity

Developmental age placement, for the professional's interpretation, is reported for 44 of the 52 tasks. Age placement was based on the scores obtained from 157 infants ranging in age from 4 to 48 weeks of age and representing 36 states (additional biographical data are included in the manual).

In addition, these 44 items correlated with the age placement of similar tasks found on the Bayley Scales of Infant Development. Developmental ages for passing each of the 44 tasks for the scale and the Bayley may be found in the manual; the percentage of agreement (where placement on the IES fell within the acceptable age range on the Bayley) between the two scales was 81%.

No information regarding the validity of this instrument was found in current developmental literature.

Reliability

In a test of inter-observer reliability, a trained nurse observed 20 mothers administer the scale to their infants (5 to 30 weeks of age). The mean of the percentages of agreement for all items was 97.7%.

Information concerning the reliability of this instrument was not found in developmental literature.

Cautions and Comments by the Authors

The child's exact chronological age is critical in interpreting the results of the evaluation, as is following administration directions for each of the tasks.

No clinical interpretation should be made based solely on the results of an evaluation using this scale.

Reviewer's Overall Judgment of the Measure

Assets. The scale has been designed to be administered by parents, which has resulted in a simple scoring format, explicit criteria for administering and scoring tasks, and easily understood directions for using the scale.

Parents will have access to most materials used in evaluating the infant.

Limitations. Reliability and validity samples are quite small. It would appear that this instrument is ready for more vigorous validation and reliability studies.

Bibliography

There is no bibliography included in the manual for this scale.

References

Gingold, W., Gingold, P., McDonald, K., Flamer, G., & Hoehle, W. Infant Evaluation Scale. Fargo, ND: Red River Human Services Foundation, 1978.

Ordinal Scales of Psychological Development

Authors	Ina C. Uzgiris and J. McV. Hunt
Publisher	University of Illinois Press Urbana, Illinois
Date of Publication	1975
Materials Available	Assessment in Infancy: Ordinal Scales of Infant Psychological Development Six demonstration films: "Ordinal Scales of Infant Psychological Development" Testing equipment is not available from the publisher. The authors make suggestions about equipment to use.

Purposes/Traits Assessed

The instrument was designed to indicate the infant's current level of cognitive organization, increase knowledge of psychological development, and determine the factors which foster development in infants.

The six scales which assess sensorimotor development are: 1) development of visual pursuit and the permanence of objects; 2) development of means for obtaining desired environmental events; 3) development of imitation (vocal and gestural); 4) development of operational causality; 5) construction of object relations in space; and 6) development of schemes for relating to objects.

Intended Test Population

The scales may be used with children functioning in the sensorimotor stage of development.

Individual or Group Administration and Time Required

The scales are individually administered to the infant in an environment familiar to him/her. The manual does not contain any information regarding the amount of time required to administer the instrument. It is not necessary to administer all of the scales or all the items within a scale to the child.

Instrument Format

The text contains information regarding the theoretical background for the scale, reliability data, general directions for administering the items, descriptions of suggested testing materials, the record forms for the six scales, and a bibliography.

The fifteen items in Scale I require visual tracking behaviors and finding hidden objects (including visible and invisible displacement). For Scale II (12 items), the infant performs in order to "cause events or obtain objects" (e.g., hand watching, use of a string or a stick to obtain an object). Scale III is divided into vocal and gestural imitation.

Vocal imitation (six tasks) ranges from imitation of familiar sounds to imitation of new words. The range of gestural imitation (four tasks) is from simple familiar movement to novel "invisible" action (e.g., blinking the eyes). Scale IV includes seven tasks in which the infant attempts to cause an interesting event to reoccur (e.g., hand watching, giving a mechanical toy to an adult to be activated). Scale V (eleven items) includes visual regard, coordination of gaze and grasping, following the path of a moving object, recognition of the reverse side of an object, and coordination of actions to obtain an object (e.g., pulling a string to obtain a toy, moving around on obstacle). The three tasks in Scale VI include the infant's interaction with objects (e.g., playing with more than one toy, using a cup to pretend to drink, naming the toy).

Administration Procedures

Six demonstration films (one for each scale) are available for rent or purchase as aids in learning to use the scales. The authors do not include explicit training directions in the text for using the scales.

Uzgiris and Hunt imply that two persons should conduct the evaluation: one as the examiner and one as observer.

A chapter is included in the book which gives general directions for administering the scales and a description of suggested testing equipment. Separate chapters provide directions for administering each scale. Each chapter includes information on: the category of the tasks, a description of each task, the position of the infant when presenting the item, suggested test equipment, the directions for administering the task, the number of trials for presenting the task, and possible infant responses.

It is not necessary to administer all of the scales to an infant and there is not a suggested sequence for administering the items. It is suggested that a group of similar tasks be presented at one time.

Scoring Procedures

The authors do not include directions for training in scoring the scales.

Two types of record forms are provided: examination and summary. The former are completed during the evaluation by checking the infant's response in each trial of the task. The latter are completed after the evaluation and indicate the highest level of response by the infant for each step in the scales.

Scale VI is an exception in that the infant's response to a specific object is noted (previous scales assume that responses are generalized for testing materials; Scale VI assumes specific objects elicit particular schemes).

Interpretation Procedures

Uzgiris and Hunt do not include any guidelines for interpreting the record forms of children evaluated through the use of these scales. The authors state that scores may be used for comparison of infants and groups

of infants. Age norms are not included for the scales to support these comparisons.

The summary record forms reveal the steps the infant has passed and the highest step attained for each of the scales administered to the infant.

Item Selection

Items were selected to reflect Piaget's description of psychological development and were organized into a hierarchy of simple to complex with the infant building on early skills to accomplish more complex tasks.

Individual items were chosen to represent landmarks of development with the responses of the infant to the tasks indicating the infant's level of cognitive organization.

Validity

The authors state that the scales have intrinsic validity based on the ordinal nature of the tasks and their reflection of landmarks in the developmental process.

King and Seegmiller (1973) compared performances on the Uzgiris-Hunt scales and the Bayley Scales of Infant Development for black, firstborn males residing in the Harlem area. The infants were tested at 14 (N=51), 18 (N=37), and 22 (N=33) months of age. Predictive validity for each scale was determined for each age interval. Correlations ranged from .01 (means) to .55 (causality) between 14 and 18 months, from .09 (causality) to .42 (vocal imitation) between 18 and 22 months, and from .00 (causality) to .40 (means) between 14 and 22 months. A correlation of .56 is reported for scores obtained at 14 and 22 months for the total of the scales.

Wachs (1975) reported the performance of 23 infants on the Uzgiris-Hunt scales at 3-months intervals between 12 and 24 months and the Stanford-Binet at 31 months of age. Correlations between scores obtained at 12 months (N=21) and 31 months ranged from -.12 (vocal imitation) to .52 (object permanence). Total scale scores correlated .64 with Stanford-Binet scores. At 15 months (N=20), correlations ranged from .17 (gestural imitation) to .43 (schemata); the overall multiple correlation between all 15-month subscales and the subsequent Stanford-Binet scores was .76. At 18 months (N=22), correlations ranged from .36 (object permanence) to .66 (schemata) with a multiple scale correlation of .86 with the Stanford-Binet. At 21 months (N=20), correlations ranged from .12 (schemata) to .48 (causality) and a correlation of .70 was obtained for the total of the scales with the Stanford-Binet. At 24 months (N=17), correlations ranged from .44 (causality) to .64 (objects in space) and a correlation of .92 was obtained for the total of the scales with the Stanford-Binet.

Reliability

The authors report reliability for this instrument. A total of 84 infants (four infants representing each month of age up to one year and at least four infants per two months of age from one year to two years) were examined in their own homes with their mothers present. Infants were evaluated two times within a 48-hour period by one of three pairs

of observers. The averaged percentage (Note 1) of inter-observer agreement was from 72 to 100% for the 157 infant actions for which evidence was available. The averaged percentage was above 90% for 87% of the actions. The means for Scales I through VI were 96.7%, 96.2%, 95.7%, 93.7%, 96.9% and 93%, respectively. In determining test-retest reliability, the authors found the mean of the averaged percentages between scores on the two evaluations to be 79.9% (range 42 to 100%). The averaged percentages were above 70% for 80% for the 157 actions produced by the infants. The means for Scales I through VI were 83.8%, 75.5%, 70%, 71.2%, 84.6%, and 79%, respectively.

Kahn (1976) reports reliability for use of the instrument with severely and profoundly retarded children. Test-retest reliability (N=30, tested at a one-week interval) ranged from .88 to .96 for the six scales. Interexaminer reliability (N=15, tested at a one-week interval by two examiners) ranged from .82 to .95. A scalogram analysis for the scales (excluding the Schemes Scale) resulted in an index of consistency ranging from .812 to 1.0.

Cautions and Comments by the Authors

The examiner should make inferences about the infant's abilities based on observation of the infant's behavior rather than simply recording the presence or absence of behaviors.

Uzgiris and Hunt state that "competence is based on a hierarchial organization of a number of abilities and motive systems with several relatively independent branches" (Uzgiris & Hunt, 1975, p. 15).

Reviewer's Overall Judgment of the Measure

Assets. This appears to be a well developed instrument; it is supported by validity and reliability studies, and can be used for observing infant actions within an early childhood program.

Limitations. The examiners must be very familiar with the administration procedures for each of the tasks and be well-versed in Piaget's theory to fully utilize the scales.

While background information for these scales was valuable for this author's purpose, a manual containing a brief description of the development of the scales and validity and reliability data would be more practical for using this instrument in an early childhood program.

The use of two examiners will increase the costs of administering the scales.

Bibliography

An extensive bibliography is included in the book.

References

Kahn, J. V. Utility of the Uzgiris and Hunt scales of sensorimotor development

with severely and profoundly retarded children. American Journal of Mental Deficiency, 1976, 80, 663-665.

King, W. L., & Seegmiller, B. Performance of 14-to-22-month-old black, firstborn, male infants on two tests of cognitive development: The Bayley Scales and the Infant Psychological Development Scale. Developmental Psychology, 1973, 8, 317-326.

Uzgiris, I. C., & Hunt, J. McV. Assessment in Infancy: Ordinal Scales of Infant Psychological Development. Urbana, IL: University of Illinois Press, 1975.

Wachs, T. D. Relation of infant's performance on Piaget scales between twelve and twenty-four months and their Stanford-Binet performance at thirty-one months. Child Development, 1975, 46, 929-935.

Content Note

1. Since the inter-observer agreement for the first session closely matched that for the second session, and since neither session yielded consistently higher agreement than the other, the two percentages of agreement for each infant action were averaged.

Preschool Attainment Record (PAR)

Author	Edgar A. Doll, Ph.D.
Publisher	American Guidance Service, Inc. Publishers' Building Circle Pines, Minnesota 55014
Date of Publication	1966 (Research Edition)
Materials Available	Manual Data Record

No testing materials are required.

Purposes/Traits Assessed

The PAR was designed for use in assessing children who are difficult to test because of sensory impairments, speech or language problems, emotional disturbance, neuromuscular problems, test resistance, or cultural problems. It was also designed to facilitate educational planning, treatment, and management. An interview procedure is used to evaluate the child's physical (ambulation and manipulation), social (rapport, communication, responsibility), and intellectual (information, ideation, creativity) functioning in his day-to-day environment.

Intended Test Population

Handicapped or disadvantaged children (0 to 84 months) are the targeted population, although the measure can also be used with nonhandicapped children.

Individual or Group Administration and Time Required

The assessment process is individualized; time requirements are not reported in the manual, but it has been reported elsewhere as taking 20 to 30 minutes (Southworth, Burr, & Cox, 1980).

Instrument Format

The measure includes 16 items per 1-year age interval (two per category, eight categories) with items listed in developmental sequence.

Information for completing the score form is obtained from a reliable informant supported, when possible, by observations of the child. Gross motor skills (walking, climbing, jumping, etc.) are evaluated in the ambulation section. Tasks on the manipulation scale include grasping, putting parts together, and pre-writing skills. The rapport category includes attending, discrimination, and play behaviors. Tasks which are included on the communication scale are babbling, labelling, the ability to carry on a conversation, and reading simple phrases. The responsibility category includes feeding, dressing, and toileting skills and the ability to obey rules. Included on the information section are: recognition of quantity, knowledge of time, and the ability to provide personal information.

The ideation scale evaluates the child's ability to discriminate between familiar and unfamiliar stimuli, count objects, and discriminate by attribute. The creativity category tasks include investigation of the environment, construction, drawing, and performing musical skills.

Administration Procedure

The measure does not require special training for use. The examiner may interview a person familiar with the child, observe the child, or elicit a response in order to complete the Data Record.

In an interview situation, general questions are asked and then, if necessary, information about specific behaviors is requested. The entire record is administered in all situations.

Biographical information about the child is recorded on the front cover of the Data Record. Life age means are included for each task. The last page of the record is a Summary and Profile Sheet.

Scoring Procedures

Training in scoring the PAR can be accomplished by attending a workshop on the administration of the PAR or studying the test manual. Criteria guidelines for scoring each item are included in the manual. Scores are not adapted when handicapping conditions impair a child's performance.

The basis for an individual item score (interview, observation, tested) is recorded for each task. Items are scored as mastered, intermittent, unsatisfactory, or the child was not given an opportunity to perform the task.

A raw score is computed from the total of successfully accomplished items (one-half point for intermittent scores, one point for mastered scores). The raw score is divided by 16 (the number of items per age interval) or multiplied by .75 (the month value of each item) to compute the Attainment Age (AA). The AA (expressed in months and years) is divided by the child's Life Age (LA) and multiplied by 100 to determine the Attainment Quotient (AQ). The Summary and Profile Sheet is completed (example provided in the manual) by crossing out the tasks passed by the child. The total passes are recorded on the profile by category and by age periods.

Interpretation Procedures

It is suggested that clinicians, school psychologists, or other testing specialists should interpret results (Southworth, Burr, & Cox, 1980).

The Summary and Profile Sheet provides a visual illustration of scores by category (demonstrates category performance variance) and by age intervals (demonstrates consistency of age progression).

Doll (1966) suggests that results may be used for "estimates of the child's native intelligence" (p. 21) and that comparisons may be made between these results and valid Mental Age and IQ scores obtained from other tests.

Item Selection

The PAR is the result of revision and extension of the Vineland Social Maturity Scale (Doll, 1965) and a scale by Doll and Edward L. McKnight designed to evaluate handicapped preschoolers.

No attempt was made to norm reference the measure; instead items were placed according to current developmental research.

Validity

No validity data are reported in the manual.

Hug, Barclay, Collins, and Lamp (1978) evaluated 148 Head Start students (134 black, mean age 72.7 months; 14 white, mean age 71.9 months) using the PAR and the Slosson Intelligence Test (SIT). A correlation of .46 (p .001) was obtained between the PAR (mean AQ=90.65) and SIT (mean IQ=91.21).

Another study of concurrent validity was conducted by Krasner and Silverstein (1976) with 27 cerebral palsied children (1 year, 4 months to 4 years, 11 months) classified as "multiply handicapped and retarded to various degrees." The children were individually administered the Vineland Social Maturity Scale (one parent interviewed) and the Gesell Developmental Schedules by a psychologist at two different sessions. The PAR was individually administered by one of six different teachers. Testing was completed within a 6 month interval. Social Quotient scores on the VSMS ranged from 7 to 94 with a mean of 35.51; Attainment Quotients on the PAR ranged from 9.9 to 150 with a mean of 55.82. A correlation of .88 was obtained for the PAR and the VSMS. The physical section of the Gesell (motor and adaptive combined) correlated with the PAR at .87, the Gesell language section correlated with the PAR at .84, and the Gesell personal-social section and the PAR had a correlation of .89.

In a study by Lederman and Blair (1972), PAR scores were obtained for 28 5-year-olds and 20 4-year-olds in a Title III program from interviews with their mothers and teachers. Higher ratings were reported for the 5-year-olds than for the 4-year-olds by mothers and teachers. The point-biserial correlation between age and ratings for the Attainment Quotient was .78 for teachers and .91 for mothers. Two subtests of the Metropolitan Reading Test (word knowledge and numbers) were given one year later to 30 children remaining in the program. Correlations between MRT scores and PAR ratings by teachers and mothers were .69 and .46, respectively. Controlling for age reduced the coefficient for the teachers to .62 and for the mothers to .24

Reliability

No reliability data are reported in the manual.

Owens and Bowling (1970) conducted a study to determine the internal consistency among the eight subtests of the PAR. The sample consisted of 68 boys (28 months to 140 months) and 32 girls (41 months to 150 months), all classified as retarded. The intercorrelations among the eight subtests ranged from .244 (ambulation and communication) to .767 (information and ideation).

Cautions and Comments by the Author

When differences occur between the child's usual behavior and optimal performance, those differences should be noted on the Data Record.

In evaluating or planning for the handicapped or disadvantaged child, gross range data from the Vineland Social Maturity Scale may be correlated with the more refined results from the PAR.

Reviewer's Overall Judgment of the Measure

Assets. This is a "quick to administer" evaluation which can be used with children who can not be evaluated, or only with great difficulty, on other standardized instruments.

Limitations. The lack of standardized norms and administration directions reduces the value of the Attainment Age and the Attainment Quotient. The Data Record is not designed to record results of a re-evaluation.

The limited number of items per catagory for each interval can provide only a gross estimate of the child's skills and abilities. There are no guidelines for using this measure for planning purposes as suggested by the author. Instructions for administering the measure are quite limited.

Doll justifies, in part, the development of the PAR because of the large number of culturally and economically disadvantaged preschool children; however, he presents no evidence of having developed an unbiased measure.

Bibliography

The manual does not contain a bibliography.

References

Doll, E. A. Preschool Attainment Record. Circle Pines, MN: American Guidance Service, 1966.

Hug, N., Barclay, A., Collins, H., & Lamp, R. Validity and factor structure of the Preschool Attainment Record in Head Start Children. The Journal of Psychology, 1978, 99, 71-74.

Krasner, P. R., & Silverstein, L. The Preschool Attainment Record: A concurrent validity study with cerebral palsied children. Educational and Psychological Measurement, 1976, 36, 1049-1054.

Lederman, E., & Blair, J. R. Comparison of the level and predictive validity of Preschool Attainment Record ratings obtained from teachers and mothers. Psychology in the Schools, 1972, 9, 392-395.

Owens, E. P., & Bowling, D. H. Internal consistency and factor structure of the Preschool Attainment Record. American Journal of Mental Deficiency, 1970, 75, 170-171.

Southworth, L. E., Burr, R. L., & Cox, A. E. Screening and Evaluating the Young Child: A Handbook of Instruments to Use from Infancy to Six Years. Springfield: Charles C. Thomas, 1980.

Preschool Language Scale (PLS)

Authors	Irla Lee Zimmerman, Violette G. Steiner, and Roberta Evatt Pond	
Publisher	Charles E. Merrill Publishing Company 1300 Alum Creek Drive Columbus, Ohio 43216	
Date of Publication	1979 (Revised Edition)	
Materials Available	Manual (includes Spanish version) Picture Book Test Booklet (25 score forms)	$32.95
	Additional PLS score forms (package of 12)	$ 9.95
	Other materials required include twelve one-inch colored blocks, sandpaper, coins, and a watch with a second hand.	

Purposes/Traits Assessed

The PLS was developed as a screening instrument to detect language strengths and weaknesses in the areas of auditory comprehension and verbal ability and as an aid in the development and evaluation of language programs.

The early stages of language development are evaluated by two categories: auditory comprehension and verbal ability. The child demonstrates comprehension through nonverbal responses to stimuli presented. The verbal ability section measures vocabulary, verbalized memory span, stages of concrete and abstract thought, concept acquisition, articulation, and ability to use grammatical features of the language.

Intended Test Population

The PLS was designed for use with English or Spanish speaking children functioning at a preschool or primary language level (one to seven years).

The authors suggest the auditory comprehension section is appropriate for assessing a deaf child's ability to read lips and measuring the comprehension level of pathologically or psychologically speech impaired children.

Individual or Group Administration and Time Required

The child is evaluated individually within a suggested time interval of fifteen minutes (Simonek, 1976).

Instrument Format

Test items are developmentally sequenced by sections in the administration directions and in the testing booklet. Each age interval contains four tasks per category (i.e., comprehension, production).

Auditory comprehension behaviors include pointing to pictures and following directions to demonstrate understanding of body parts, prepositions, size, functions of objects, actions, categories, color, weight, and quantity. Tasks in the verbal ability section include imitation, verbalizing needs, labelling, using pronouns, repeating sentences, using plurals, counting, articulating consonants, and composing sentences.

The Preschool Language Scale Picture Book is durably constructed of lightweight cardboard. Individual pages include the plate number, the corresponding scale item number, section number, and age interval. Illustrations are realistic and colorful. Pictures up to the 2½ year age level are large with only a few illustrations per page; after 2½ years, illustrations are smaller and more numerous for each page.

The first page of the test booklet is the Preschool Language Scale Profile. It provides a visual illustration of the child's language development in relation to his/her chronological age. The last page of the test booklet is the Preschool Language Scale Checklist which is a compilation of all scores from each of the scales.

An appendix is included which contains a Spanish version of the scale for use with children bilingual in Spanish and English.

Administration Procedures

The measure was designed for use by child development specialists, psychologists, speech therapists, teachers, and administrators. The suggested criteria for effectively using the instrument are: 1) knowledge of the instrument and the needs of the children; 2) ability to follow administration procedures; and 3) ability to correctly score responses.

The directions for administering the scale include the section number, age interval, the number of items in the section, the points scored per item, and the age credit per item. Information for each task in a section includes the name of the item, materials required, procedure for administering, scoring criteria, rationale, and the reference.

As a general rule, tasks are to be administered in the order arranged on the scale. Testing begins at a level six months below the child's assumed language age. If the child fails any test item in this section, the examiner tests at a lower level until the child passes all items in an age level. The highest interval at which all tasks are passed is the base age. (Even if the child does not pass all tasks at the one year level, his base age is one year). Testing continues until the child fails all tasks at an age interval, the ceiling age. (A ceiling age may not be established for older children).

Scoring Procedures

The manual does not contain any specialized training procedures necessary for scoring the PLS. Criteria are included for passing each task and examples of scores are provided in the manual.

The test booklet is composed of the PLS Profile, the PLS scales, and the PLS checklist. General directions for administering the scales are included in the booklet. Information is also provided regarding the section number,

age interval, point score column, item name, and criteria for passing the task. Space is provided (for appropriate items) to indicate the question to which the child responded correctly.

An alternative point credit system is used for scoring test items which are numerically sequenced. The child receives one point (items 1-32) or two points (items 33-40) for each item passed. The total point scores for the auditory comprehension and the verbal ability sections are obtained by adding the total points for all passed items. Total point scores are converted into an Auditory Comprehension (AC) Age Score and a Verbal Ability (VA) Age Score using a table in the manual. A Language Point Score is obtained by adding the AC Age and the VA Age which is then converted into a Language Age (LA) Score using a table in the manual.

The Auditory Comprehension Quotient (ACQ) is derived by dividing the AC Age by the child's chronological age and multiplying by 100. The Verbal Attainment Quotient (VAQ) is obtained by dividing the VA Age by the chronological age and multiplying by 100. The Language Age Quotient (LAQ) is derived by dividing the Language Age by the child's chronological age and multiplying by 100.

The Preschool Language Scale Profile is completed by drawing straight lines across the profile at the child's chronological age, base age, and ceiling age. The examiner then circles the items on the scale which were failed by the child.

The PLS Checklist is completed by scoring items from the scale as passed or failed. Scores are recorded for the AC Point score, AC age, ACQ, AC base age, AC ceiling, VA Point score, VA Age, VAQ, VA base age, VA ceiling, Language Point score, Language Age, and LQ.

Interpretation Procedures

Special training in interpreting results is not required. The authors suggest the results will indicate the child's strengths and weaknesses, and emerging skills and help teachers in planning. Guidelines for using this information for planning purposes are not included in the manual.

The child is considered "at risk" if s/he scores below his age level on the PLS. Low scores on the verbal ability section only suggest the child would benefit from speech therapy. When low scores are obtained for both sections, individual passes and failures should be examined to determine whether the delay may be due to a hearing loss, mental retardation, or lack of exposure to a language-rich environment (left to the examiner's own interpretation). Failure to produce consonants at the appropriate age intervals suggests the need for remediation.

When the child's chronological age line is "far" above the child's base line on the PLS Profile, developmental delay is indicated. The Profile is also indicative of performance in the areas of sensory discrimination, logical thinking, grammar, memory span, temporal/spatial relations, and self image.

For children bilingual in Spanish and English, administration of the

scale in Spanish and English reveals which language is primary and whether performance is the result of actual impairment or the lack of mastery of the English language.

Item Selection

This edition is adapted from the 1969 version of the PLS. Tasks were selected based on the theory that the child's language skills develop sequentially according to individual capacity, maturation, and individual experience. The scale focuses on the assessment of consonants since research suggests that two-thirds of speech therapy is the result of articulation disorders.

Age placements were determined by developmental research and use of the PLS with children enrolled in Head Start, nursery schools, and child development programs. However, this instrument was not formally standardized. The authors stated that the PLS "is ready for vigorous normative studies" and also suggested examiners "develop their own norms."

Validity

The authors claim "intrinsic content validity" for the measure based on the detailed description and rationale provided for each of the items.

A number of studies were conducted to establish concurrent validity for this measure (summarized in a table in the manual).

The PLS and the Illinois Test of Psycholinguistics (ITPA) were administered to 15 cerebral-palsied children (4 to 10 years). The mean ITPA IQ was 82.5 and the mean PLS LQ was 80.6 with a correlation of .97.

In a sample of 32 children (3 and 4 years) given the Utah Test of Language Development (mean LQ 104) and the PLS (mean LQ 123), the two scales correlated .70.

The Peabody Picture Vocabulary Test (PPVT) and the PLS were administered to 22 children (3 to 5 years) with a resulting PLS mean LQ of 123, a PPVT mean IQ of 177, and a correlation of .29.

A study involving 15 children (3 to 6 years) with "possible language delays" resulted in a PPVT mean IQ of 92.8 and a PLS mean LQ of 86.1. Correlation between scores obtained by the two measures was .59.

Initial scores at the beginning of the program on the PPVT (median IQ, 81) and the PLS (median LQ, 86) for three successive classes of Head Start students (N=171) resulted in a correlation of .59. Retesting at the end of the 12 month program (PPVT median IQ, 101; PLS median LQ, 105) resulted in a correlation of .32.

Lass and Golden (1975) compared scores on the Auditory Comprehension section of the PLS with the Peabody Picture Vocabulary Test for 24 children (15 boys, 9 girls; 2 years, 8 months to 8 years, 8 months). All of the children had speech and language disorders without hearing impairment and were enrolled in a preschool program. A correlation of .72 was obtained for the two instruments.

The 1960 version of the Stanford Binet LM and the PLS were administered to 60 3 to 5 year olds with a resulting Binet mean IQ of 93.2 and a mean PLS LQ of 101.6 and a correlation coefficient of .66. For 8 TMR children (6 to 11 years), a correlation of .70 was obtained (Binet mean IQ of 43, PLS mean LQ of 37).

Twenty-five 4 year olds were tested on the Columbia Mental Maturity Scale (CMMS) and the PLS. The mean CMMS IQ was 74.2 and the mean PLS LQ was 89.3 with a correlation coefficient of .67.

The authors of the manual report three studies conducted to establish predictive validity for the initial version of the PLS. PLS scores for Head Start children were compared with scores obtained one year later on the Lee Clark Reading Readiness Test. Sixty-five percent were identified as "average or above" or "below average" on both measures; 10% scoring "average or above" on the PLS later scored "below average" on the Lee Clark. In the second sample, 79% of the Head Start children tested on the PLS received identical results two years later on the Lee Clark; 7% scoring "average or better" on the PLS scored "below average" on the Lee Clark.

A group of 40 4 year olds testing "above average" on the PLS were retested two years later with the Burt Reading Test. A correlation of .24 was reported between scores obtained on the two measures.

Reliability

The authors report that "in various studies using the PLS, the two parts (auditory comprehension and verbal ability) have a median correlation of .68 (Zimmerman, Steiner, & Pond, 1979).

A group of Head Start children's PLS tests were scored twice (odd items only and even only), resulting in split-half reliability coefficients ranging from .75 to .92 with a median of .88.

Cautions and Comments by the Authors

Scores obtained through the use of the PLS are not directly comparable to intelligence test scores obtained from standardized measures.

The auditory comprehension category may be useful in program planning and screening children who would benefit from speech therapy.

Reviewer's Overall Judgment of the Measure

Assets. The scale is usable by a wide range of professionals from programs including children functioning in a broad age range. The inclusion of the Spanish version of the scale increases its usability for programs serving a bilingual community.

The directions for administering the tasks and scoring criteria are very extensive, thereby increasing the scale's usability and reliability. While the instrument has been designed to yield age and quotient scores, the lack of standardization would appear to reduce the value of these derived scores.

The format of the test booklet makes it easy to complete; it is

unlikely that scores will be recorded in the wrong place. The profile and checklist are useful in visually emphasizing the child's strengths and weaknesses.

The Picture Book is suitably designed for ease in its use and the illustrations are appropriate for the children's ages.

The small number of testing materials required for administering the scales makes it easily transportable to the testing site and reduces the time required to assemble the materials necessary for administering the scale. Costs for acquiring the necessary materials should be minimal.

Limitations. Referral criteria could be more specific. The instrument is ready for normative studies.

Bibliography

An extensive bibliography is included in the manual. References include language and general development scales (e.g., Test of Language Development, Bayley Scales of Infant Development) and developmental literature which includes work by Gesell and Slobin.

References

Lass, N. J., & Golden, S. S. A comparative study of children's performance on three tests for receptive language abilities. Journal of Auditory Research, 1975, 15, 177-182.

Simonek, L. Preschool Language Scale. In O. G. Johnson Tests and Measurements in Child Development: Handbook II (Volume 1). San Francisco: Jossey-Bass, 1976.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. Preschool Language Scale. Columbus: Charles E. Merrill, 1979.

The Revised Developmental Screening Inventory (RDSI)

Authors	Hilda Knobloch, Frances Stevens, and Anthony F. Malone
Publisher	Department of Pediatrics Albany Medical College of Union University Albany, New York 12208
Date of Publication	1980
Materials Available	Inventory Testing materials are not available from the publisher; specific descriptions of suggested materials are included on the inventory; most materials should be easily obtainable.

Purposes/Traits Assessed

The inventory was designed for screening infants and toddlers in the areas of adaptive, gross motor, fine motor, language, and personal-social skills.

Intended Test Population

The instrument may be used with children functioning in the one to thirty-six month range.

Individual or Group Administration and Time Required

The measure is individually administered. Time requirements will depend upon the number of questions answered by the parent, the age of the child, and the number of behaviors observed in spontaneous play. No estimate of the time required for testing is provided in the inventory.

Instrument Format

Tasks are arranged in four-week age intervals where every twelfth week is considered a "key age." Each category of skills contains from one to four tasks per age interval.

The response modes in the adaptive category include visual regard of stimuli, reaching for and securing stimuli, demonstrating problem solving skills, using pre-writing skills, and building with cubes. Head righting behaviors, sitting, and locomotor skills are included in the gross motor section. Babbling, laughing, location of sound, vocabulary, following verbal commands, pointing to stimuli, use of pronouns, and identification of colors are among the tasks in the language section. Grasping, stacking cubes, turning pages, and cutting tasks are included in the fine motor category. The personal-social behavior responses include visual regard, response to others, feeding skills, dressing skills, and dramatic play.

Administration Procedures

Special training in the use of the inventory is not required. Knobloch, Stevens, and Malone (1980b) suggest the inventory be completed by a trained child care worker, public health nurse, or primary care physician.

The examiner computes the child's chronological age, which is expressed in weeks through the fifty-sixth week, and in months after that. The evaluation begins at the age interval corresponding to the child's age (corrected for prematurity). If the child is failing at this age interval, tasks from the next lower age level are administered until the child is able to pass the items. The evaluation continues until the child is no longer passing any of the items.

Some items may be scored by observation and/or parental report, while others must be observed. For tasks evaluated by parental report, the items need to be stated exactly as they are on the score form. In the case of observations, the infant/toddler should be given the opportunity to display the most advance behavior first (e.g., naming body parts rather than pointing to body parts). The evaluation should cover one area of behavior at a time rather than completing all tasks in an age interval at one time.

The examiner is referred to A Manual of Developmental Diagnosis (Knobloch, Stevens, and Malone, 1980) for clarification of administration procedures. Notes are provided on the score form for administering a limited number of tasks.

Scoring Procedures

The inventory does not require special training for scoring the child's responses to the tasks. Criteria are provided for a few of the tasks, the examiner is referred to A Manual of Developmental Diagnosis (1980) for additional assistance in scoring.

Behaviors are scored as present, absent, infrequent, or unknown. Based on these scores, a "maturity level" expressed in weeks or months, as appropriate, is assigned for each of the five categories. The maturity level is the point where the child begins to exhibit more failures than passes. This maturity level is an approximate age score expressing the gradual achievement of skills.

A Developmental Quotient (DQ) is calculated for each category by dividing each maturity age level by the child's chronological age and multiplying by 100.

After DQ's are calculated, Diagnostic Categories are assigned to each area where a DQ of 75 or less is abnormal, a DQ of 76 to 85 is questionable, and a DQ of 86 or more is normal.

Space is provided on the score form for recording the maturity level, the DQ, and the Diagnostic Category for each area of behavior.

Interpretation Procedures

Guidelines are provided in the inventory for interpreting the significance of scores obtained through the use of this measure.

A Final Diagnostic Category is recorded based on the results of the DQ's. When adaptive or gross motor DQ's are 75 or less, an abnormal rating is assigned for that evaluation period. A questionable rating is obtained if adaptive or gross motor DQ's are 76 to 85, or fine motor only is 75 or less. If all areas are 86 or above, a normal identification is made.

A delay in the language section is not interpreted as significant if skills in the adaptive category are considered normal and the child appears to hear and understand what s/he hears.

Children identified as abnormal should be referred for a complete diagnostic evaluation. A questionable rating indicates the need for ongoing observation to determine whether extensive assessment is required.

Item Selection

This instrument is a revision of the Developmental Screening Inventory (1966) and is based on the Revised Gesell and Amatruda Developmental Schedules. No further information is provided in the inventory regarding the criteria or procedures used to select particular items.

Validity

The authors do not include any information concerning the validity of this instrument in the inventory.

Knobloch, Stevens, and Malone (1980b) report the results of evaluations of 125 normal and abnormal children (16 weeks to 36 months of age) on the RDSI and the Revised Gesell Developmental and Neurological Examination. The RDSI successfully detected all children rated as questionable or abnormal by the full examination. Overscreening with the RDSI occurred for 4 of 78 children or 5%. When scores are evaluated by neuromotor status and intellectual potential, the RDSI detected all abnormal or questionable evaluations for neuromotor status and overreferred in 5 of 79 cases or 6%; for intellectual potential, the RDSI missed 2 of 33 or 6% and overscreened 5 of 109 or 4.6%.

A comparison of screening results by the RDSI and the Denver Developmental Screening Test (DDST) was reported by Knobloch, Stevens, and Malone (1980b). For a sample of 237 children, use of the DDST failed to detect 9% of the children rated as abnormal by the Bayley or Stanford-Binet and 38% rated as questionable; in addition, 52% referred as abnormal or questionable were rated normal by the Bayley or Stanford-Binet. This was compared to the results of the RDSI (N=125) in which all children rated as abnormal or questionable by the Bayley or Stanford-Binet were detected and only 2% of those referred as abnormal or questionable were diagnosed as normal.

Reliability

Information concerning the reliability of this measure is not included in the inventory or in developmental literature.

Knobloch, Stevens, and Malone (1980b) report interobserver reliability for a sample of 48 children 16 weeks to 21 months of age. A total of 305 items were observed and 2,302 comparisons were made with an overall percentage of agreement of 93.7% (adaptive, 92%; gross motor, 96%; fine motor,

88%; language, 97%; personal-social, 95%).

Interrater reliability in assigning maturity age levels for 184 children is also reported by Knobloch, Stevens, and Malone (1980b). Correlations between the two raters ranged from .84 to .99.

Cautions and Comments by the Authors

The instrument should not be used for diagnostic purposes but as the basis for referral for a complete evaluation. It can not be used for predicting later development.

Parental information should be considered in determining age maturity levels, especially in the area of language development.

While a list of equipment is included in the inventory, these are only suggested materials and adaptations are allowed.

Reviewer's Overall Judgment of the Measure

Assets. The instrument may be used for screening purposes without special training and appears to require little administration time.

Limitations. Since the determination of maturity levels is subjective, the reliability of the Developmental Quotients and Diagnostic Category scores obtained will be questionable.

The small number of tasks per category for each age interval does not support the calculation of DQ's.

It also appears that the 1940 Gesell and Amatruda standards are used and may be out of date.

Directions for administering tasks and scoring criteria are vague for most of the items on the inventory.

Flexibility in testing materials can reduce the cost of using this instrument but may also reduce the reliability of scores obtained from the use of the measure.

Directions are unclear as to whether tasks may be observed to be passed or whether they may be passed by parental report.

Bibliography

No references are included in the inventory.

References

Knobloch, H., Stevens, F., & Malone, A. F. The Revised Developmental Screening Inventory. Albany: Albany Medical College, 1980a.

Knobloch, H., Stevens, F., & Malone, A. F. A Manual of Developmental Diagnosis: The Administration and Interpretation of the Revised Gesell and Amatruda Developmental Examination. Hagerstown, MD: Harper & Row, 1980b.

Vineland Social Maturity Scale

Author	Edgar A. Doll, Ph.D.
Publisher	American Guidance Service, Inc. Publishers Building Circle Pines, Minnesota 55014
Date of Publication	1965 edition
Materials Available	Condensed Manual of Directions Scoring Sheet No testing equipment is required. Also available: Doll, E. A. The Measurement of Social Competence. Minneapolis: Educational Test Bureau, 1953. (664 pages)

Purposes/Traits Assessed

The norm-referenced scale was designed to evaluate social competence ("self adequacy in a number of different kinds of situations" Doll, 1957) by measuring the subject's skills in the categories of self-help (general, dressing, eating), self-direction, locomotion, occupation, communication, and social relationships.

Intended Test Population

The scale may be used with normal or handicapped individuals from birth to adulthood.

Individual or Group Administration and Time Required

The administration procedure consists of an individually conducted, untimed interview of a person intimately familiar with the child, such as the primary caregiver or teacher.

Testing time is not specified in the manual but the format suggests that only a short period of time would be required for administering the scale. Goodwin and Driscoll (1980) estimate a 30-minute administration time.

Instrument Format

Items are organized in the manual and on the record form by categories and in developmental order.

The scale includes a total of 117 behaviors; tasks are not included for each category for each age interval. The number of items per age interval ranges from 3 to 17 and the younger age levels tend to have more items than the older age levels.

The record sheet is completed by interviewing someone familiar with

the child. Although tasks are included appropriate to and including adulthood, only those items designed for young children will be mentioned in this review. Tasks in the self-help category are divided into three sections: general (balancing the head, sitting, grasping, standing, overcoming simple obstacles, asks to go to the toilet, avoids simple hazards), eating (drinks from a cup; uses spoon, fork, and knife; discriminates edible substances, etc.), and dressing (removes socks, coat, or dress; washes and dries hands; buttons). Locomotion tasks evaluate the child's ability to move about his/her environment and include crawling, movement about the room, and in the home. Some of the tasks included in the occupation category are occupies self, pre-writing, cutting, and helping with household tasks. The communication section includes spontaneous cooing or laughing, imitation of sounds, following simple instructions, labelling familiar objects, and using short sentences. No self-direction items are included for children under four years of age. Tasks in the socialization section evaluate the child's interactions with others.

Administration Procedures

In the manual, Doll suggests the scale may be used by parents, teachers, clinicians, pediatricians, psychiatrists, and psychologists who follow administration procedures exactly and are skillful in interviewing. However, in Measurement of Social Competence (Doll, 1953) it is reported that training in the interview technique is essential. "The examiner must be broadly experienced in general techniques of clinical psychological casework or the equivalent in similar disciplines" (Doll, 1953, p. 266).

Tasks are to be presented to someone familiar with the child or they may be administered to a subject with a minimum mental age of five years. Items are grouped for convenience only and may be administered in any order.

Biographical data are obtained first and questioning begins below the anticipated final score for each category. Questions are phrased to determine what the subject does routinely rather than what the subject is capable of doing; general questions are to be followed up with detailed questions.

A basal level is established where there are at least two consecutive pluses within a category. Questioning should not be discontinued before at least two consecutive minus scores are recorded for each category.

Scoring Procedures

A fairly elaborate scoring procedure is outlined in the manual and additional scoring principles and guidelines are given by Doll (1953) in Measurement of Social Competence.

Consistent and habitual performance of a task, behaviors surpassed by a higher level of behavior, and tasks below the basal levels are scored as passing. A passing score (F+) is given if the child does not perform a task due to restraint or lack of opportunity but did perform the task when the restraint was removed or when the opportunity was available. A passing score (+N.O.) is also given if the child could presumably perform a task but has lacked the opportunity. A "plus-minus" score is given for inconsistent behaviors. A "minus" score is given for tasks the child is unable to perform or rarely performs. A score of "-N.O." is given when the child

has not had an opportunity to perform a task but would presumably fail the task.

The total score is obtained by adding the basal score to the sum of the credits. One credit is given for a "plus" or an "+F." One credit is given for an "+N.O." score which occurs among otherwise continuous plus scores; in all other cases one-half credit is given. One-half credit is also given for "plus-minus" scores. No credit is given for "minus" or "-N.O." scores.

The total score is converted to a social age score (SA) from a table in the manual. A social quotient (SQ) is derived by dividing the SA by the child's life age (LA) and multiplying by 100.

Interpretation Procedures

Training in interpretation of results is not specified in the manual.

Handicapping physical and social conditions are to be considered in interpreting the scores. The manual does not provide guidelines for making referrals based on results obtained from the use of the scale.

Item Selection

"Each item is conceived as representing a general growth in social responsibility which is expressed in some detailed performance as an overt expression of that responsibility" (Doll, 1965, pp. 1-2).

The scale was constructed on Binet and Simon's "age scale" principle where a "life age mean" value is assigned to each task. The life age mean reflects the age when the average child is able to perform the task (Handbook of Psychiatric Scales, 1981).

The standardization sample consisted of 620 "normal" subjects (20 subjects per year of age from birth to 30 years) equally divided by sex and living in southern New Jersey in 1935. The sample was controlled for social status.

The average age norms were used to arrange the tasks in developmental sequence.

Validity

The manual does not contain any reports of validity studies; however, a large number of studies are contained in Measurement of Social Competence (Doll, 1953). The focus of these studies includes the validity of this scale for use with young children, adults, retarded subjects, and the physically handicapped.

A more recent study of concurrent validity was conducted by Krasner and Silverstein (1976) using 27 multiply handicapped cerebral palsied children. The children were also classified as retarded and ranged in age from 1 year, 4 months to 4 years, 11 months. Each child was individually administered the Vineland Social Maturity Scale, the Preschool Attainment Record, and the

Gesell Developmental Schedules. Each instrument was administered in a separate session; a psychologist administered the VSMS and the Gesell, and one of six teachers administered the PAR. The VSMS social quotient scores ranged from 12 to 94 with a mean of 38.33, PAR age quotient scores ranged from 9.9 to 150 with a mean of 55.82, and the combined scores of the Gesell subtests ranged from 7 to 94 with a mean of 35.51. The VSMS correlated with the PAR at .88, the Gesell physical scale (motor and adaptive combined) at .98, the Gesell language scale at .94, and the Gesell personal-social scale at .99.

Goulet and Barclay (1963) conducted a study to determine the value of VSMS scores in estimating the Stanford-Binet Mental Age (MA) of non-institutionalized mentally retarded children. A group of 164 "mentally retarded" white children (96 males and 68 females; age range 26 to 161 months) were administered the VSMS and the Stanford-Binet. Social age scores from the VSMS had a correlation of .78 with Stanford-Binet Mental Age scores. The correlation between VSMS Basal Social Ages and MA was .71. A total of 87% of the subjects had social ages equal to or greater than their mental age scores. Goulet and Barclay concluded that social age scores could be used to predict mental age scores using the VSMS when it was impossible to obtain mental age scores.

A study by Fromme (1974) investigated the relationship between the VSMS and the Stanford-Binet (SB) and Wechsler Intelligence Scale for Children (WISC). Children (86 boys and 54 girls) with speech defects, ranging in age from 18 to 191 months, were tested by a staff psychologist on at least two of the instruments. The mean age at the time of testing for the 111 children evaluated on the VSMS was 4.6 years, and the mean test score was 95.6. For the 96 children completing the Stanford-Binet, the mean age at test time was 5.4 years, and the mean test score was 84.4. WISC scores were broken down by subtests (since the complete scale had not been administered to all subjects). A mean test score of 86.2 was obtained for 65 children (mean age at testing 8.5 years) on the Verbal IQ subtest; a mean test score of 90.5 was obtained for 84 children (mean age at testing 8.2 years) on the Performance IQ subtest; a mean score of 87.7 was obtained for 63 children (mean age at testing 8.4 years) on the Full Scale IQ. Correlations of .63 were obtained for the VSMS and the Stanford-Binet, .52 for the VSMS and the WISC Verbal IQ, .51 for the VSMS and WISC Performance IQ, and .45 for the VSMS and WISC Full Scale IQ. When the sample was divided into two groups: above-median-age (50 months) and median-age-or below, correlations between the VSMS and the Stanford-Binet for children older than 50 months (N=28) were .87, .61 between VSMS and WISC Verbal IQ (N=26), .55 between VSMS and WISC Performance IQ (N=36), and .52 between VSMS and WISC Full Scale IQ (N=26). Correlations decreased for children 50 months or younger; correlations for VSMS and Stanford-Binet were .40 (N=39), for VSMS and VIQ .40 (N=16), for VSMS and PIQ .40 (N=20), and for VSMS and FSIQ .31 (N=15). Fromme concluded that estimates of intelligence by use of the VSMS were appropriate only for those children over 50 months of age.

Reliability

The manual does not contain reports of any reliability studies. Doll

includes reliability studies in his Measurement of Social Competence (1953). Among these studies are test-retest data for a group of 250 subjects from the standardization population. Retesting occurred from 1.7 to 1.9 years after the original testing. A correlation of .976 was obtained between social age scores from the first and second tests. A third test session was completed for 196 of the 250 subjects 1.35 years after the second evaluation. The correlation between the second and third social age scores was .97 and between the first and third was .94.

Kaplan and Alatishe (1976) compared VSMS social quotient scores obtained from interviews with the mothers of 20 preschool children (8 girls and 12 boys) with social quotient scores obtained from scales completed by the children's teachers. The mean social quotient obtained from the mother's reports was 137.8 while a mean of 114.9 was obtained from the teachers. A Pearson correlation coefficient of .244 was obtained for the total sample.

Cautions and Comments by the Author

The scale is not a rating scale but can be used as a measurement of growth and change after re-evaluations.

The scale can be used to evaluate the influence of environment, physical handicaps, and culture.

The parent, as informant, may become more aware of the child's abilities, decrease his/her anxiety about the child's capabilities, and become more positive in his/her attitude toward the school (Doll, 1957).

Reviewer's Overall Judgment of the Measure

Assets. The manual and record form are of a convenient size for use and the format is easy to follow when administering the scale.

The conversion of total scores to age equivalent scores is greatly facilitated when the table in the manual is used.

Limitations. It appears that the VSMS requires the use of trained clinical psychologists for accurate administration and interpretation.

Scoring is heavily dependent on the subjective judgment of the examiner. Since the measure is conducted in an interview format, behaviors are never actually observed.

Although tasks were developed several decades ago, they are still representative of social maturity skills necessary in the 1980's. A revision of some of the age life norms to reflect current norms would be desirable. Some of the tasks appear to be culturally biased in favor of high socioeconomic children.

The manual does not contain any guidelines for referral based on scores obtained from the VSMS: therefore, referral must be based on the interpre-

ter's own judgment.

The scale does not include the same number of tasks at each age level. Seventeen items are included for each of the first two years but ages five to fifteen are evaluated by approximately one item per year. Therefore, the measure lacks consistency in its depth of coverage at different age levels.

The manual should include a summary of some of the major validity and reliability studies reported in Measurement of Social Competence (Doll, 1953).

Bibliography

The manual does not include a bibliography. An extensive bibliography is included in Measurement of Social Competence (Doll, 1953).

References

Doll, E. A. Measurement of Social Competence. Minneapolis: Educational Test Bureau, 1953.

Doll, E. A. Evaluating social maturity. Education, 1957, 77, 409-413.

Doll, E. A. Vineland Social Maturity Scale. Circle Pines, MN: American Guidance Service, 1965.

Fromme, D. K. On the use of the Vineland Social Maturity Scale as an estimate of intellectual functioning. Journal of Clinical Psychology, 1974, 30, 67-68.

Goodwin, W. L., & Driscoll, L. A. Handbook for Measurement and Evaluation in Early Childhood Education. San Francisco: Jossey-Bass, 1980.

Goulet, L. R., & Barclay, A. The Vineland Social Maturity Scale: Utility in assessment of Binet MA. American Journal of Mental Deficiency, 1963, 67, 916-921.

Handbook of Psychiatric Rating Scales. New York: Research and Education Association, 1981.

Kaplan, H. E., & Alatishe, M. Comparisons of ratings by mothers and teachers of preschool children using the Vineland Social Maturity Scale. Psychology in the Schools, 1976, 13, 27-28.

Krasner, P. R., & Silverstein, L. The Preschool Attainment Record: A concurrent validity study with cerebral palsied children. Educational and Psychological Measurement, 1976, 36, 1049-1054.

Instruments for which Validity and Reliability Data were not Found

Assessment-Programming Guide for Infants and Preschoolers

Author	Warren Umansky, Ph.D. University of Georgia Athens, Georgia	
Publisher	Developmental Services, Inc. P. O. Box 1023 Columbus, Indiana 47201 (812) 376-9404	
Date of Publication	1976 (Revised Edition)	
Materials Available	Guide Rating Packets	$2.50 $ .25 each

Most of the testing materials required are available in infant and preschool classrooms.

Purposes/Traits Assessed

The Assessment-Programming Guide for Infants and Preschoolers is designed to assess a child's needs through observation and task administration and after evaluation to guide and direct an individual program to meet these needs. Skills are assessed in six areas: motor, perceptual-motor, language, cognitive (24 to 72 months only), self-help, and social-personal. An assessment interview form for use with a parent or caregiver is included.

Intended Test Population

The guide can be used with both normal and handicapped children between the ages of 0 and 72 months. However, since severe handicaps which affect a particular area of development will bias the total score, the affected area is omitted when computing the final score. The omission is noted and a description of the child's performance in the affected area is included.

Individual or Group Administration and Time Required

It is recommended that this measure be individually administered to the child and used over a period of several weeks. Ongoing observation will probably be necessary as some items would be hard to structure into the typical one hour test situation (e.g., is protective of younger children). The child should be allowed to become familiar with the teacher and the classroom before the assessment is done. The child's history and the use of the "Assessment Interview" with the parents should provide a reasonable indication of where to begin the evaluation in the item sequence.

Instrument Format

Although no testing materials are provided, the guide contains a list

of materials which can be locally obtained. Some variation in appearance and content might be expected since specific descriptions of the materials are not provided. Some of the materials (balance beam, steps, large balls) are quite bulky and would be inconvenient to transport if testing is being done outside the classroom.

Color coded rating sheets and administration directions and developmentally sequenced items facilitate the use of the measure during the testing situation and should help reduce recording error.

Tasks are listed by category and are separated into 4-month age intervals; the number of tasks per interval and category range from three to twelve.

Response modes are varied by age group and area of behavior. Items assessed at the low end of the age range for each area (except cognitive tasks which begin at 24 months) involve observation of spontaneous motor and verbal behaviors. After 8 months of age the child is generally required to perform a skill. The 81 motor behaviors include locomotive skills, cube stacking, balancing, and throwing. Matching, cutting, pre-writing, and stringing skills are evaluated by administering the 76 perceptual-motor tasks. The language category (73 items) evaluates receptive and expressive language development through such tasks as naming body parts, using plurals, performing commands, and pointing to stimuli. Evaluation in the cognitive area (46 items) begins at 24 months of age and includes puzzle skills, shape identification, definition of objects by function, and pre-writing skills. Development in the category of self-help is assessed using 53 tasks which include feeding, grooming, and toileting skills. The final category, social-personal, contains 55 items; behaviors evaluated include adult and peer interactions, dramatic play, and demonstration of knowledge of social behaviors.

Administration Procedures

The teacher, after becoming familiar with the guide and rating sheets, is suggested as the most appropriate evaluator of the child. It is this author's opinion that reflexive testing, described in an appendix in the manual, be done by a physician or physical therapist, rather than by the teacher, as suggested by the author of the guide.

The child's performance is scored as s/he engages in a structured play situation; suggested test materials for some of these situations are listed beside the directions for administering the tasks. There is no suggested sequence for the order of assessment of the items; however, this is an accepted practice for an ongoing assessment device.

Scoring Procedures

Extensive training is not required for scoring the observed behaviors. The observer simply records whether the behavior being evaluated is absent, partially acquired, or has been mastered. Formal assistance in scoring is provided in the form of sample responses for some behaviors and examples of score computations. However, well-defined criteria for scoring items

are not included in the guide, and therefore scores are subjective.

A basal level is established where the child passes all tasks at a given age level. Behaviors are assessed until the child fails all items in two consecutive age ranges. Items between the basal and ceiling levels are assigned values which are then added to the base age. After all subscale scores have been computed and recorded on a one page evaluation form, these scores are totaled and then divided by six (i.e., number of subscales) to determine the average age of functioning. This is not an IQ score but an estimate of the child's developmental level.

Interpretation Procedures

No training is specified for interpreting results and no interpretive assistance is provided beyond the caution to not include low scores for developmental areas affected by a handicapping condition.

Determination of a delay is left to the user's individual interpretation since no guidelines are included in the manual. An individual program, including behavioral objectives and activities to develop the behavior, is written for the developmental area where a delay is observed. This program may be recorded on a "Prescriptive Program" form which is included in the guide.

Item Selection

This guide is a revision of an earlier (1972) version and was compiled to give an indication of the child's overall development. A literature review (including work by Bayley, Cattell, and Gesell and Amatruda) was the basis for choosing items for inclusion in the measure with further revision occurring through use in the classrooms of Developmental Services, Inc. of Columbus, Indiana (no data provided).

No information regarding the establishment of the norms is provided in the guide.

Validity

No validity studies are reported in the manual and a review of the literature did not reveal any information regarding the validity of this measure.

Items appear to measure the developmental areas as indicated and to be appropriate to the given age intervals. An adequate number of items are included for each age range to provide a good indication of the child's actual performance in each developmental area. There is some duplication of tasks between subscales (e.g., stacking blocks is on the motor and the perceptual-motor scales, cares for all toileting needs is included on the self-help and the social-personal scales).

Reliability

No reliability data were reported in the guide and a search of outside

literature failed to reveal any studies of reliability for this instrument.

A lack of criteria for scoring the child's performance and the lack of specific descriptions of test materials suggest reliability for the measure is not high.

Cautions and Comments by the Author

There are no cautions or comments other than those already described in earlier sections.

Reviewer's Overall Judgment of the Measure

Assets. The guide is designed so teachers familiar with the evaluation items can quickly learn to administer and score the measure. The wide range of behaviors, the broad age range, and the use of common classroom materials make the guide attractive to those early childhood programs serving a wide age range of children and operating on a restricted budget. Many items would be appropriate for a caretaker interview if evaluation time is short.

Limitations. The lack of support data and the lack of guidelines for interpreting results seriously compromise the use of this measure.

Bibliography

The twelve references cover a broad base in child development (e.g., cognition, language, motor, perceptual-motor) but are somewhat dated.

References

Umansky, W. Assessment-Programming Guide for Infants and Preschoolers. Columbus, Indiana: Developmental Services, 1976.

Birth to Three Developmental Scale

Authors	Tina E. Bangs and Susan Dodson	
Publisher	Teaching Resources Corporation 50 Pond Park Road Hingham, Massachusetts 02043	
Date of Publication	1979	
Materials Available	Manual Subscale Record Forms	$34.50
	Most materials required can be found in infant/ toddler classrooms.	

Purposes/Traits Assessed

The scale is a criterion-referenced instrument which can be used for screening. Skills evaluated by the instrument cover the areas of language (comprehension and production), problem solving, social-personal development, and motor development.

Intended Test Population

This instrument was designed for use with all children, regardless of handicapping conditions, functioning in the 0 to 36 month age range.

Individual or Group Administration and Time Required

Testing is done individually by observing or eliciting a response from the child with parental report allowed for behaviors not observed during the testing situation. No information is given regarding the time required to administer and score the scale, however, the small number of items included for each age range would suggest the scale can be completed quickly.

Instrument Format

The 85 scale items are listed in developmental sequence, by subscale, in the directions for administration and on the record forms. Each subscale includes 16 to 18 tasks, with 1 to 3 tasks per age interval.

Two lists of materials required for testing are included in the manual: the first is arranged by age range, and the second is arranged by subscale. While material descriptions are not detailed enough to permit duplication, enough information is available that test materials should be similar to those used in constructing the measure. Most materials will be familiar to the child and will already be available in an infant/toddler program.

Response modes are varied. The receptive language tasks have the

child responding to sounds and pointing to demonstrate understanding; vocalization of consonants and vowels, imitation of sounds, and vocabulary tasks are among those evaluated in the expressive language category. Problem solving skills are evaluated by tasks which include the child's demonstrating knowledge of object permanence, placing shapes in a form board, and copying shapes. In the category of social-personal, the child's peer and adult relationships and grooming skills are evaluated. Tasks in the motor area include an evaluation of sitting, throwing, and balancing skills.

Administration Procedures

Examiners should be able to "reliably observe" children's behavior and, in addition, should practice administering the scale to six normal children between the ages of birth and three years.

Subscale administration directions (in the manual) and record forms are color coded to reduce error. Administration instructions include the following information for each scale item: item number, description of the behavior to be observed, test materials required, procedure for administering the item, and the criteria for scoring the child's response. There is no set sequence for administering the items.

Scoring Procedures

Special training is not described in the manual for scoring the scale; an example of a completed record form is included in the manual. Scoring is made more objective by the inclusion of criteria for passing, failing, and emerging behaviors.

The sequence for scoring the scale is as follows: 1) total the number of passes on a subscale, 2) determine the basal age (highest age level at which all items are passed), 3) determine the categorical age by adding the number of items passed beyond the basal level to the basal score, 4) complete this sequence for each subscale used, and 5) complete the "Assessment Summary" form (a summary profile).

Interpretation Procedures

The child's subscale scores are plotted on the Summary profile along with his chronological age to graphically reveal any developmental delays. Information is not available in the manual regarding the interpretation of what constitutes a delay or the interpretation of the significance of a delay.

Item Selection

A review of developmental literature produced 244 tasks for consideration for inclusion on the scale. The final 123 scale items were selected according to seven criteria to represent the four behavioral categories at six month intervals. The criteria included definition of the developmental areas, inclusion of behaviors which generally appear within a six

month interval, elimination of items which must be performed within a specific time period, elimination of the need to specially train administrators, development of a reliable observation measure, and the requirement that 80% of the standardizing sample master the task.

A standardization sample of 357 "normal" children, ages 4½ to 36 months, was used to "validate" scale items. The children (from California, Tennessee, and Utah) were reported to be equally represented by age range, sex, and urban/rural residence. Tables in the manual illustrate the percentages of children passing the individual test items for each subscale; the criterion for inclusion on the scale was an 80% passing rate for children at the upper limit of that age range.

Some items appear to be taken from the Bayley Scales of Infant Development and the Denver Developmental Screening Test; however, no credit is given for the source of any of the tasks.

Validity

The manual contains no information concerning the validity of the scale; no information regarding its validity could be found in a review of the literature.

Reliability

To determine inter-rater reliability, three pairs of raters each scored 25 children (apparently equally divided among the six age intervals), with resulting reliability coefficients, for total and subscale scores, ranging from .88 to .99.

Additional information concerning the reliability of this scale was unavailable from any source.

Cautions and Comments by the Authors

Bangs and Dodson suggest the scale is also suitable as a basis for individual program planning, however, no guidelines are included with the measure.

Reviewer's Overall Judgment of the Measure

Assets. The scale is usable by a wide range of early childhood professionals with a minimal amount of time required for training. As a screening device, it contains an adequate number of items per subscale for most age ranges.

The color coding of the subscales facilitates the use of the manual and record forms. The list of testing materials, by age and subscale, will reduce the amount of preparation time necessary for administering the scale.

Limitations. Insufficient data are provided in the manual to

establish strong validity or reliability for the instrument. However, the inclusion of scoring criteria for each task should result in at least moderately reliable scores. Items taken from established instruments (e.g., the Bayley Scales of Infant Development) should be credited to their source.

Bibliography

A list of twelve references (published scales and research literature) is included in the manual.

References

Bangs, T. E., & Dodson, S. Birth to Three Developmental Scale. Hingham, MA: Teaching Resources Corporations, 1979.

The Boyd Developmental Progress Scale

Author	Robert D. Boyd, Ph.D. Crippled Children's Division (Psychology) & the Department of Medical Psychology University of Oregon Medical School Portland, Oregon
Publisher	Inland Counties Regional Center, Inc. P. O. Box 6127 San Bernardino, California 92408
Date of Publication	1974
Materials Available	Manual Record Sheet (package of 20) The small number of required testing materials are common to preschool classrooms with the exception of the Wide Range Achievement Test (WRAT).

Purposes/Traits Assessed

This is a criterion-referenced screening instrument designed to measure developmental functioning in the areas of motor skills, communication skills, and self-sufficiency skills.

Intended Test Population

Tasks are included to screen all children functioning in the developmental range of birth to eight years of age.

Individually or Group Administration and Time Required

The scale is individually administered through a combination of observation of the child's behavior and a parent interview. No information is available in the manual regarding the time required to administer and score the test. However, it has been reported that the test can be administered in ten to twenty minutes (Southworth, Burr, and Cox, 1980). Establishing a basal and ceiling level and using the parent as a source of information in the self-sufficiency section will reduce the time requirements.

Instrument Format

The 150 scale items are listed by age intervals for each of the three categories in the directions for administration and on the record form. The scale includes five items per age group for each developmental area.

A standardized reading test, such as the Wide Range Achievement Test

(WRAT), is required for the upper age levels of the communication area. Other testing materials are limited in number, inexpensive, and easily obtained.

The child is asked to respond to the items on the scale in a variety of ways. In the motor area, the performance of fine motor (e.g., grasping, block building, drawing) and gross motor (e.g., sitting, walking, climbing stairs, tricycle/bicycle riding) skills are required. Self-sufficiency tasks require the child to demonstrate feeding, grooming and toileting skills and mobility around the home and the neighborhood. In the areas of communication, the child is required to respond to sounds, vocalize, verbalize, point to stimuli, bring materials to the examiner, and exhibit reasoning skills.

Administration Procedures

The scale may be administered by anyone in the helping professions familiar with the scale items. The directions to the child or questions for the parent are included in the manual for each scale item.

Administration of items begins by establishing a base age (where the child passes all of the items in an age interval) and continues until a ceiling is reached (where the child fails all of the items in an age interval). There is no recommended sequence for administering scale items, although, the suggestion is made to administer all items using the same test materials at one time and to administer gross motor items last.

Scoring Procedures

Special training in scoring the scale is not required. Criteria for passing each item are included in the administration directions.

Items are scored as passed (X) or not present/inconsistent (-) by drawing through the blocked space listing the task on the record form. Space is provided on the record form to record the child's verbal response to items in the communication area.

Interpretation Procedures

The scale is designed to indicate a general level of functioning for the child. Vague guidelines are provided for interpreting screening results but *no* criteria are designated for referring the child for further diagnosis; referral is essentially left to the screener's judgment.

Item Selection

Tasks included on the scale were required to meet two criteria: 1) research evidence that 60 to 70% of the children in that age range had passed the item, and 2) the item must have "survival value" (lead to more efficient living).

References for item selection are not included in the manual. Several

of the items appear to be taken from the Denver Developmental Screening Test and the Progress Assessment Chart.

Standardization data are unavailable for the placement of tasks on this scale.

Validity

Data concerning the validity of this scale are not reported in the manual or in the literature.

Items in the motor and self-sufficiency sections appear to evaluate skills in these areas and appear to be placed in the appropriate age interval. Some of the items in the communication area might be more appropriately placed in a personal-social category or in the self-sufficiency category (e.g., engages in parallel play, separates without a fuss).

Reliability

Reliability data are not provided in the manual. In addition, reliability studies have not been published in current developmental literature.

The inclusion of criteria for scoring tasks increases the reliability of scores obtained through the use of the instrument.

Cautions and Comments by the Author

Boyd does not make any additional statements regarding the use of this instrument which would restrict the usability of the scale.

Reviewer's Overall Judgment of the Measure

Assets. The measure is usable by anyone knowledgeable about young children and for programs serving a wide range of children.

Directions for administering scale items and criteria for scoring tasks are well-defined as are the examples for establishing rapport and motivating children.

Limitations. The manual fails to credit sources for tasks, report standardization data, or provide any information which sould indicate strong validity or reliability for the scale. Items at the upper age levels of the self-sufficiency area appear to be biased in favor of urban-residence children over rural-residence children.

While there does not appear to be a logical sequencing of tasks within an age interval in the administration of directions (e.g., fine motor skills interspersed with gross motor, feeding with toileting skills), similar tasks have been grouped horizontally on the record form (vertically grouped by age interval).

A list of required materials, by age range or by skill area, would be helpful in administering the scale.

The score sheet is simple to use but visually overcrowded. The manual is bound with a hardback which increases its durability but a binder type of manual would be easier to use during a testing situation.

Bibliography

There is no bibliography included with the manual.

References

Boyd, R. D. The Boyd Developmental Progress Scale. San Bernardino: Inland Counties Regional Center, 1974.

Southworth, L. E., Burr, R. L., & Cox, A. E. Screening and Evaluating the Young Child. Springfield: Charles C. Thomas, 1980.

Brigance Diagnostic Inventory of Early Development (Birth to Seven Years)

Author	Albert H. Brigance
Publisher	Curriculum Associates, Inc. 5 Esquire Road North Billerica, Massachusetts 01862
Date of Publication	1978
Materials Available	Inventory of Early Development $49.95 Developmental Record Books (10) A preassembled kit of testing materials is not available from the publisher.

Purposes/Traits Assessed

This is a criterion-referenced measure designed for assessment and program planning. The inventory includes the areas of psychomotor skills (preambulatory, gross, and fine motor), self-help, speech and language (also pre-speech), general knowledge and comprehension, and early academics (readiness, basic reading skills, manuscript writing, and math).

Intended Test Population

The measure may be used to assess those children functioning between birth and seven years of age.

Individual or Group Administration and Time Required

Children are evaluated on an individual basis or in a small group situation.

Administration of the entire inventory is not recommended. The time required for administering the instrument is dependent on the number of skills assessed, the number of skills adapted for group administration, the skill range, the child's developmental level, and the existence of a handicap.

The examiner may choose to observe the child, elicit a response from the child, or interview the parent during the assessment procedure; the recommended method is included for each task.

Instrument Format

Within each area, the inventory items are developmentally sequenced and arranged by skill sequence. Developmental areas are separated by colored dividers in the manual to provide quick access to the desired category.

The manual is bound in loose leaf style with a durable plastic cover. It is, therefore, possible to remove only those pages required for assessing particular skills rather than using the entire inventory which is quite large. (Care needs to be taken to ensure all pages are promptly and correctly returned to the manual if the instrument is used in this manner).

The "child pages" included in the manual are attractively designed; illustrations are of adequate size and are generally uncrowded.

A broad spectrum of responses are elicited from the child in the administration of this inventory. Pre-ambulatory skills (birth to one year) evaluate the child's response to 35 tasks while in the supine position (e.g., thrusts legs, shakes rattle), 14 tasks in the prone position (e.g., lifts head 90°, creeps), 11 in the sitting position (e.g., head erect, pivots), and 16 in the standing position (e.g., supports weight, walks alone). Gross motor skills cover the age range of 1 to 6 years and behaviors include standing (18 tasks), walking (19), stairs and climbing (16), running (8), jumping (5), hopping (2), kicking (4), balance board (11), catching (8), rolling and throwing (6), ball bouncing (4), rhythm (6), and wheel toys (8). The items evaluating fine motor skills include ages 7 months to 7 years for general eye/finger/hand manipulation (33), block tower building (4), puzzles (3), pre-handwriting (14), draw-a-person, designs, cutting with scissors (13), painting with a brush (9), and using clay (9). Self-help skills (birth to 7 years) evaluate feeding/eating (35), undressing (10), dressing (13), unfastening (8), fastening (12), front and back, inside-out, toileting (12), bathing (13), grooming (9), and household chores (9). Pre-speech behaviors (birth to 1 year, 3 months) evaluate receptive language (11 tasks), gestures (9), and vocalization (14). Speech and language skills (6 months to 7 years) include syntax (22), length of sentences, personal data response (e.g., gives name, age, address), social speech (16), following verbal directions, pointing to pictures, naming pictures, articulation of sounds, repeating number sequences, sentence memory, and singing (10). Tasks in the general knowledge and comprehension area (18 months to 7 years) include pointing to and naming body parts; matching, pointing to, and naming colors; matching, pointing to, and naming designs; time, quantity, and direction concepts; sorting and classifying; knowledge of use of objects, function of community helpers, what to do in different situations, and where to go for services. Readiness skills (18 months to 7 years) include the child's experience with books, visual discrimination, and recognition of letters. Basic reading skills (5 years, 3 months to 7 years, 9 months) include auditory discrimination, identification of vowels and consonants, and reading level. Printing skills are evaluated in the manuscript writing section (5 years, 3 months to 7 years). The math section (18 months to 7 years) includes counting, numerical comprehension, ordinal position, addition, subtraction, recognition of money, and time concepts.

Administration Procedures

Brigance suggests trained child development specialists or early child-

hood educators supervise the use of the inventory, but no special training procedures are described.

The manual contains extensive information for administering the inventory items. The general directions for each skill sequence include: the skill to be assessed, the developmental age notation, the corresponding page number of the "Developmental Record Book," the assessment method, materials required, the criteria for discontinuing testing on the task, the time allowance, scoring procedure, the references used for sequencing skills and validating developmental ages, directions for administering individual items within each skill sequence (e.g., painting with a brush), and helpful notes.

Since the examiner administers only those tasks targeted for acquisition (after the initial assessment), no sequence is suggested for administering the inventory.

Scoring Procedures

Special training is not required for scoring the inventory items; however, knowledge of child development would be beneficial since scoring criteria are not provided for every inventory item.

Scores are recorded in the individual "Developmental Record Book" which also has space provided for biographical information about the child and for observations about the child's learning style. The record book may be used for consecutive assessments by color coding each evaluation. Test items are scored as mastered, not achieved, or omitted; the assessment method (parent interview, observation, or elicited response) is noted for each item. A horizontal bar graph is included for some skill sequences for the option of graphing the child's developmental age in that skill sequence.

Interpretation Procedures

Special training in interpretation is not required.

After mastered skills have been identified, the skills listed next in the developmental sequence can be targeted in the development of an individual program for the child (suggested objectives are listed for each skill sequence).

After several assessments, the record book will illustrate the child's original performance, the progress made by the child, and the targeted objectives for the next instructional period.

The developmental age graphs provide a quick illustration of the child's performance for his chronological age. No guidelines are included for referral for more extensive evaluation when delays are observed.

Item Selection

An extensive literature review was completed in selecting items and

"validating" skill sequences and developmental ages (references included in the manual).

References for age norms are reported for the individual tasks in the administration directions for the skill sequences. Field testing of the inventory was conducted in programs in sixteen states (program names and addresses are provided but no data are reported in the manual).

Sequences and developmental ages for academic skills were established and validated from curriculum practices and current tests (data not reported in the manual).

Validity

No information concerning the validity of this measure is reported in the Inventory or in the developmental literature.

Items included in each skill sequence and developmental area appear to adequately assess behaviors appropriate to the six categories; furthermore, age "norms" are within currently accepted ranges.

Reliability

Reliability data are not reported in the manual or in the literature.

The lack of scoring criteria for some inventory items suggests low to moderate reliability for scores obtained for these tasks. Reliability is potentially further weakened when administration procedures are adapted for a handicapping condition.

Cautions and Comments by the Author

Because individual children develop at different rates and patterns, the developmental ages listed for test items should be considered as guidelines.

In order to obtain the child's highest level of performance, the examiner may adapt the format or specific instructions to the child.

A list of nineteen recommendations for using the inventory is included in the manual (e.g., demonstrate motor skills, when helpful, while giving directions; don't base evaluation of motor skills on the child's ability to hear and follow directions).

Reviewer's Overall Judgment of the Measure

Assets. The measure can be used, without specially trained examiners, for assessing and planning for a wide age range of children.

Administration directions are extensive; the developmental age notations are valuable in deciding where to begin testing and provide cues as to the point where the child should begin to fail tasks; the objectives listed for

the skill sequence should be useful in planning for the child.

Limitations. The examiner may find the manual and record book awkward to use in a classroom situation. In addition, administration of all tasks appropriate to an individual child's age level would be time consuming. Teachers may therefore, find the use of this instrument more appropriate in developing Individualized Education Programs than for a regular preschool classroom.

It appears that the Inventory has been designed to provide a more complete evaluation of children in the upper age levels than in the younger age levels. Also, performance on academic and school-oriented skills will probably be heavily dependent on local school programs.

Adequate space is provided in the "Developmental Record Book" for making notes about the child's performance; however, space is limited for noting the method of assessment of the task, indicating omitted items, and indicating items assessed but not mastered, as suggested by the author.

Bibliography

A total of 53 instruments and texts were utilized in developing this measure. Included among these are the Slosson Intelligence Test, the Vineland Social Maturity Scale, Bower's Development in Infancy (1974), and Lavatelli's Piaget's Theory Applied to an Early Childhood Curriculum (1970).

References

Brigance, A. H. Brigance Diagnostic Inventory of Early Development (Birth to Seven Years. North Billerica, MA: Curriculum Associates, 1978.

The Bzoch-League Receptive-Expressive Emergent Language Scale for the Measurement of Language Skills in Infancy (REEL)

Authors	Kenneth R. Bzoch and Richard League	
Publisher	University Park Press 300 North Charles Street Baltimore, Maryland 21201	
Date of Publication	1971	
Materials Available	Assessing Language Skills in Infancy (handbook)	$16.95
	Test forms	$12.95

Purposes/Traits Assessed

The REEL scale is a norm-referenced assessment tool designed for use during "the critical linguistic imprinting period" (Bzoch & League, 1971). The scale evaluates the child's receptive (decoding and understanding) and expressive (encoding of meaning for communication purposes) language skills.

Intended Test Population

Children in the birth to 3 years developmental age range may be evaluated through the use of this instrument.

Individual or Group Administration and Time Required

Parental interview (with concurrent observation of the child) is the recommended technique for using the scale. Although no information is available in the manual regarding the length of time the interview will require, the competent interviewer, viewing the parent as a reliable source, should be able to complete the scale in a short period of time.

Instrument Format

Testing materials are not required when the scale is scored from the parent interview; in cases where the parent is unable to report on a skill or provides unreliable information, a very few easily obtained materials will be needed.

Items on the record form are listed in developmental order and divided into receptive and expressive categories.

Tasks in the receptive category include the child's response to sound, ability to understand nonverbal cues, ability to follow directions, and demonstration of understanding by pointing to named objects.

Items evaluated in the expressive language area include cooing and babbling, the use of consonants, playing with sounds, imitating new sounds or words, and using sentences.

"The greatest number of observable new language behavior items occurs in the earliest periods of infancy" (Bzoch & League, 1971, p. 18). Therefore, the age intervals are not equal for each year but are progressively larger: 1-month intervals for the first year, 2-month intervals for the second year, and 3-month intervals for the third year. There are three items within each age interval for the receptive and expressive language sections respectively.

In addition to administration and scoring instructions for the scale, the manual includes: the theoretical model on which the scale is based, a yearly developmental description of language skills for the first 3 years of life, the hypothetical basis for interpreting patterns of language development as related to chronological age, and a glossary of terms used in the record forms and the handbook.

Administration Procedures

Special training in the use of the scale is not required; familiarization with the handbook (especially the chapter on language development during the first 3 years and the glossary) should be sufficient to administer the scale.

The scale is suggested as suitable for use by public health nurses, pediatricians, child psychiatrists, educators, clinicians, and speech pathologists.

The interview begins with receptive language items at the child's chronological age level, proceeds in either direction to obtain a basal level (the age level where the child passes two items) and moves forward until a ceiling is reached (two failures at an age level).

A Receptive Language Age (RLA) is determined (the highest listed age for the highest age interval passed), and then expressive language items are administered beginning at the level of the RLA (regardless of chronological age). The same procedure is used to determine the Expressive Language Age (ELA), as was used to determine the RLA.

Scoring Procedures

Special training in scoring is not specified in the manual; criteria for scoring are not included but the glossary and the chapter on language development may be helpful in the determination of scores.

Behaviors are scored as typical, emerging or only partially exhibited, or not observed. If the child passes two of the three items in an age interval, s/he is given credit for the entire age interval. An emerging score counts as a pass if the other two items in that age interval are passes or if two or three items are passed in the next higher age interval; otherwise, the emerging skill is interpreted as a failure. Receptive language credits are computed first and then expressive language credits.

The RLA and ELA are recorded on the face of the record form. The

Combined Language Age (CLA) may be determined from a table in the handbook, as are the Receptive Quotient (RQ), the Expressive Quotient (EQ), and the Language Quotient (LQ).

Interpretation Procedures

The manual does not specify training for interpreting scores obtained through the use of the scale. Guidelines for interpreting results are nonspecific. The scale is based on the hypothesis of language development occurring during a critical period and in a general pattern for all children. A review of the child's scores on specific tasks, the child's family history, and the child's medical history can be used for the examiner's interpretation of the basis of any delays.

Item Selection

Scale items were partially obtained from researching developmental literature (29 sources are included in the bibliography) and existing developmental scales and were reported as confirmed in laboratory testing.

A sample of 50 "normal" (free of sensory or organic disabilities) infants were selected to represent "the probable norm of environmentally language-advantaged Causasian infants" (Bzoch & League, 1971, p. 19) in determining developmental milestones. No age or demographic information on the sample was provided in the handbook.

Validity

The authors claim "inherent validity" for the scale based on developmental literature and reported sample support.

While positive correlations with intelligence tests, language tests, and social maturity scores were reported for the scale in dissertations and theses at the University of Florida, published data from these studies will not be available until 1984 (Bzoch, Note 1).

Fox, Lynch, and Brookshire (1978) compared 24 cleft palate and 24 control children on the REEL, Denver Developmental Screening Test, and the experimental version of the Birth-3 Scale. They concluded that the expressive language section of the REEL was the single most powerful subtest in predicting the classification of children into cleft palate or non-cleft categories, solely by developmental tests.

Reliability

"Using the criteria of test-retest agreement within plus or minus one age interval" (3 week testing interval), for a sample of 28 "normal" infants (demographic data unavailable) from linguistically stimulating environments, "agreement between different administrators" ranged from 90 to 100%, and yielded an overall Language Quotient (LQ) correlation of .71 (Bzoch & League, 1971, pp. 19-20).

The lack of specific criteria for scoring individual tasks on the

scale will probably reduce the reliability of scores obtained through the use of this instrument.

Cautions and Comments by the Authors

Due to the innate capacity for developing language, "environmental deprivation and/or organic disorders either must be extremely severe, or they must occur at critical imprinting or developmental periods to markedly impair emergent language development" (Bzoch & League, 1971, p. 17). The REEL was designed to differentiate between environmental and organic causes of language delay.

Achievement of receptive language skills appears to be related to the development of cognitive skills in children.

Reviewer's Overall Judgment of the Measure

Assets. The scale has been developed for use by a wide range of professionals without special training, administration time should be short, and no testing materials are required. These factors will keep the costs of using this scale at a low level.

Limitations. The scale appears to have limitations in item selection, theoretical support, norm sample size, and validation. However, it may be clinically useful to describe early vocal, symbolic, and communicative behaviors. The scale may be more appropriately used as a screening device rather than a diagnostic tool. In additon, the inclusion of only three tasks in each age interval per category is, in this author's opinion, inadequate for computing language age scores.

Bibliography

The 29 references cover a wide range of language literature including theory and research.

References

Bzoch, K. R., & League, R. Assessing Language in Infancy. Gainesville, FL: The Tree of Life Press, 1971.

Fox, D., Lynch, J., & Brookshire, B. Selected developmental factors of cleft palate children between two and thirty-three months of age. Cleft Palate Journal, 1978, 15, 239-245.

Reference Notes

Bzoch, K. R. Personal communication, June 1983.

Child Development Assessment Form

Authors	Marsha Kaufman and T. Thomas McMurrain
Publisher	Humanics Limited P. O. Box 7447 Atlanta, Georgia 30309
Date of Publication	1982
Materials Available	Manual Score forms (package of 5) $13.95 Testing materials are not available from the publisher; the manual contains a list of suggested materials; all equipment should be easily obtainable.

Purposes/Traits Assessed

This instrument has been designed as a "developmental checklist" covering the areas of social-emotional, language, cognitive, gross motor, and fine motor development. The manual may also be used for parent training, program planning, developing individualized educational programs, and developing a center-wide assessment program.

Intended Test Population

The checklist may be used with children functioning between the developmental ages of birth and 3 years.

Individual or Group Administration and Time Required

The authors suggest the checklist be completed for the individual child from informal observations over an unspecified time period, but recommend evaluation three time per year.

Instrument Format

Behaviors are listed in "a generally progressive developmental sequence" (Caballero & Whordley, 1981, p. 9), and include 18 tasks for each of the five areas for a total of 90 items on the checklist.

Responses to tasks in the social-emotional area include response to sound and visual stimuli, interaction with others, and expressing needs and emotions. Among the behaviors evaluated on the language subscale are babbling, imitation, pointing to indicate understanding, and construction of sentences. Cognitive tasks include manipulation of objects, display of object permanence behaviors, and discrimination of shape, quantity, and size. Skills on the gross motor section include reaching, crawling, balancing, standing, and walking behaviors. Fine

motor tasks evaluate the child's ability to grasp and manipulate objects, feed him or herself, and pre-writing skills.

The first part of the manual includes an overview of assessment, development of young children, nutritional needs of young children, a list of suggested testing equipment, and recommendations for the use of the checklist. The second part of the manual contains the directions for evaluating the child's behaviors. Tasks are listed by category and include scoring criteria, developmental significance of the behavior, task description, sample objectives for program planning, and suggested activities for eliciting or teaching the desired behavior. The final section of the manual makes suggestions for using the results of the evaluation.

The 12-page assessment form includes space for recording biographical data on the child and the results of three evaluations. The last page of the form is the summary profile which illustrates tasks performed by the child consistently and occasionally.

Administration Procedures

The checklist was designed for use by teachers and parents who have an understanding of the development of young children and have been trained in observing young children.

The manual contains suggested activities to elicit the desired behavior for each task.

The authors do not suggest any guidelines for beginning and discontinuing the observation procedure in terms of when the individual behaviors might be expected to occur.

Scoring Procedures

Observation skills and knowledge of the development of young children are recommended for scoring the checklist.

The checklist is completed by making a check in the appropriate column to indicate whether the child performs a task occasionally or consistently and dating the item to indicate when the task was evaluated. If the child is not observed performing a behavior, then no check mark is made.

An example of a completed checklist is included in the manual. The summary profile is completed by circling the numbers of the tasks performed consistently in one color and the numbers of the tasks performed occasionally in another color.

Interpretation Procedures

The authors do not include any training procedures in the manual for the interpretation of results from the use of this instrument. The final chapter of the manual includes a completed score form and an interpretation of the results of the use of the form with one child.

The authors suggest that results from the evaluation be used in planning for the child.

Item Selection

No information is provided regarding the criteria used in selecting tasks for inclusion on this instrument. A statement is included in the manual that the skills and behaviors may be observed during the first three years of life.

Validity

The authors do not include any validity data in the manual and no validity studies were available in current literature.

While the tasks appear to be representative of the appropriate developmental categories, no age guidelines are provided in the manual or on the score form.

Reliability

Reliability data are not included in the manual and no reliability studies were found in current literature.

The lack of standardized procedures for eliciting the desired behaviors would appear to reduce the reliability of scores obtained through the use of this instrument.

Cautions and Comments by the Authors

The assessment form is not norm-referenced and should therefore not be used to compare the performance of children or for diagnostic evaluation.

Reviewer's Overall Judgment of the Measure

Assets. The assessment form appears to be usable by teachers in the classroom without formal training in its use. Activities are included for eliciting or teaching the desired behaviors which should be useful for planning purposes.

The provision of space for re-evaluation and the summary profile appear to be useful in monitering the progress of children in a program.

Limitations. Although designed in a checklist format, the instrument is severely restricted in the number of behaviors being evaluated. Each developmental area only contains 18 tasks, which if divided equally among the three years would result in only six items per year interval. This number appears to be inadequate for the purposes suggested by the authors.

A list of "suspect" behaviors and possible indicators of handicapping conditions are included for the user's information but no

guidelines are included for referring children for formal diagnostic evaluation.

Age guidelines for the behaviors would be useful in determining where to begin and end the evaluation and for program planning purposes.

The authors suggest that the instrument is useful for a variety of purposes including staff and parent training but do not include any information regarding the use of the manual in this form.

Bibliography

The manual contains references which may be used by teachers and parents for planning activities for young children. It also contains information on nutrition for young children.

The chapter which provides an overview of child development includes references to the work of Bereiter and Engelman, Erikson, Hunt, and Piaget.

References

Cabellero, J., & Whordley, D. Orientation to Infant-Toddler Assessment. Atlanta: Humanics Limited, 1981.

Circle Preschool Child Assessment List

Authors	S. Hering, A. Fazio, and J. Hailey	
Publisher	Circle Preschool 9 Lake Avenue Piedmont, California 94611	
Date of Publication	1975	
Materials Available	Page of General Directions Score Form	$1.00
	Most testing materials will already be present in an early childhood program; a large number of materials are used.	
	Also available: Classroom Screening for 2½ to 5 year olds.	$1.00

Purposes/Traits Assessed

The Child Assessment List is a criterion-referenced assessment instrument designed to assist early childhood educators in individualized program planning and curriculum development. Six developmental categories are included in the assessment: gross motor, fine motor, self help, socio-emotional, cognitive, and language.

Intended Test Population

The list was designed for use with children functioning between the ages of 12 months and 6 years (a few tasks assess earlier behaviors), who are enrolled in an early childhood program.

Individual or Group Administration and Time Required

While performance is individually evaluated, many of the items can be set up in the classroom as group activities. The instrument is designed to be used in the classroom over a period of time (on-going classroom assessment) and, therefore, it is difficult to estimate the time required for administration.

Instrument Format

Tasks included on the list are developmentally sequenced and developmental categories are broken down into subscales: gross motor items, for example, are divided into walking, jumping, running, balancing, hopping, skipping, climbing, throwing, kicking, catching, and play activities.

The test materials required for administering the tasks are extensive,

no descriptions are provided for duplicating test materials, and no list of materials is included which will require the examiner to spend extra time in preparation for assessing a child.

The 67 items in the gross motor category evaluate the child's skills in walking, running, balancing, jumping, hopping, skipping, climbing, throwing, kicking, catching, and playing. The fine motor skills (62 tasks) include manipulation of objects, eye-hand coordination, pre-writing skills, and visual perception. Feeding, toileting, dressing, personal hygiene, and safety skills are evaluated by the 60 self-help tasks. The 52 items in the social-emotional area include interaction with adults and children, play patterns, self-control, achievement, and motivation. The cognitive tasks (81) require the child to point to objects, verbalize, classify, seriate, count, and engage in dramatic play. (The cognitive section is heavily influenced by Piaget's work). Response to sound, imitation of sounds, pointing to or bringing objects, following directions, and verbalizing to demonstrate understanding are included in the evaluation of the 108 language tasks.

Administration Procedures

Specialized training in the use of the Child Assessment List is not required and the instructions included for administering the list are very limited.

Assessment begins at or below the child's chronological age and then, if necessary, the tester works backwards until the child passes three consecutive tasks in each subscale area. No information is available for discontinuing assessment in each category.

Scoring Procedures

Knowledge of child development will be beneficial in scoring the list but training in scoring is not specified in the manual. Scoring is subjective since criteria for passing or failing tasks are not specified.

Except in the area of social-emotional development, items are scored to indicate whether the skill is present or absent and then dated. Additional space is provided on the score form for recording the goal date for acquisition of the skill, the date the skill was acquired, and the date and result of re-evaluations.

Social-emotional behaviors are scored on the basis of frequency (often, sometimes, never) and dated. The desired acquisition date, the actual acquisition date, and the date and results of re-evaluations are also recorded for social-emotional items.

Interpretation Procedures

Diagnostic referral guidelines are not included in the instructions; it is suggested that assessment results be used for planning by utilizing the child's strengths to improve the areas of weakness.

The breakdown of tasks by subscale allows the examiner to determine whether a delay is in a specific skill or in an entire developmental area.

Item Selection

Skills on the list were compiled from a review of assessment instruments (unspecified), chosen for their appropriateness for use as activities in an early childhood program and written as observable behaviors.

No information is provided by the authors regarding a confirmation of the "developmental sequences" established in this instrument.

Validity

No validity data are provided in the manual nor are validity studies available in current literature.

A wide variety of tasks are assessed for each category, tasks appear to be indicative of functioning in those developmental areas, and items appear to be age appropriate.

Reliability

The manual does not contain any report of reliability data and no studies of reliability for this instrument were found in developmental literature.

The lack of test material descriptions, directions for administering items, and criteria for scoring items would suggest low to moderate reliability for this instrument.

Cautions and Comments by the Authors

There are no cautions or comments included in the manual concerning the usability of this measure.

Reviewer's Overall Judgment of the Measure

Assets. The instrument may be administered without special training, a wide variety of skills are assessed, and it may be used in programs serving a wide age range of children.

The format of the checklist (developmental sequence and skill sections) appears to facilitate the process of planning for the individual child. Many of the skills can be easily incorporated into the daily activities of an infant/preschool program.

There is some repetition of tasks in different skill sections; however, the items are written so that the teacher can evaluate whether the child has generalized the skill or whether the skill is limited to a specific behavior (e.g., social knowledge: recognizes daily routines in the classroom, and time: relates clock time to daily schedule).

Limitations. With a targeted population of children enrolled in an early childhood program, this instrument may not be appropriate for assessing those children not enrolled in such a program.

The categories vary in the age range covered; some areas begin at 6 months, while others begin at 10, 12, 18, or 30 months. Likewise, some categories have items up to the 60-month age level, while others have items up to the 74-month level.

Bibliography

A bibliography is not included for the Child Assessment List (references for 18 instruments are included in the Classroom Screening measure).

References

Hering, S., Fazio, A., & Hailey, J. Circle Preschool Child Assessment List. Piedmont, CA: Circle Preschool, 1975.

Hering, S., Fazio, A., & Hailey, J. Circle Preschool Classroom Screening. Piedmont, CA: Circle Preschool, 1975.

Communication Evaluation Chart from Infancy to Five Years

Authors	Ruth M. Anderson, Madeline Miles, and Patricia A. Matheny
Publisher	Educators Publishing Service, Inc. 75 Moulton Street Cambridge, MA 02238
Date of Publication	1963
Materials Available	Four page chart containing general directions and score sheet. Test materials are not available in a preassembled kit but most are present in an infant/preschool classroom.

Purposes/Traits Assessed

This is a criterion-referenced screening instrument for "early detection of childhood communicative disabilities." The language section screens for intactness of the speech musculature, hearing acuity and auditory perception, vowel/consonant acquisition, and receptive and expressive language. In addition, physical growth, motor coordination, and perceptual skills are evaluated in the performance category.

Intended Test Population

The chart can be used to evaluate children functioning at levels in the range of 3 months to 5 years.

Individual or Group Administration and Time Required

Described as a "quick appraisal" of the individual child, no further information about the time requirements for administering and scoring the chart is available.

Instrument Format

Items are arranged by age interval and are divided into two categories: language and general growth and development. The number of items per age level varies from 14 to 30; the age intervals included on the instrument are 3 months, 6 months, 9 months, 1 year, 1½ years, 2 years, 3 years, 4 years, and 5 years.

Parental input is necessary to score some of the items in the 3 to 6 month age interval. The remaining items can be scored from observation of the child. Items listed in the language sections of the chart evaluate the child's response to sound, ability to drink and eat, understanding through pointing or following directions, and verbal ability. Items evaluated in

the general growth and development section require the child to demonstrate eye-hand coordination, locomotor skills, and pre-writing skills.

Administration Procedures

Extensive training is not required in administering the chart; it is suggested for use by child specialists. Directions are not provided for administering the tasks, and there are no guidelines in determining where to begin and end the screening process.

Scoring Procedures

Training in scoring the chart is not specified and no criteria for scoring the tasks are included.

Skills are scored as present, not present, or fluctuating and then a performance and a language age level are calculated (instructions are provided) and recorded on the front of the score sheet.

Interpretation Procedures

Training procedures are not included in the manual for interpreting the screening results and no criteria are available for referring the child for further diagnosis. It is implied that a discrepancy in scores between the two categories and the presence of more failures than passas at the child's chronological age level, continuing over time, should lead to referral for more extensive evaluation.

Item Selection

Some tasks were taken from other assessment instruments (e.g., Gesell, Binet, Cattell) but credit for these references is not provided for individual items. Other tasks were chosen by the authors as "diagnostically significant." The chart was organized and tested over a four year period, apparently in a clinical or classroom setting.

Validity

Items included in the chart appear to evaluate the designed categories and to be in the appropriate developmental age ranges.

Evidence of validity for the chart was not available in the administration directions nor in the child development literature.

Reliability

Reliability data for this instrument were not reported in the manual or in developmental literature.

Low reliability for the chart is suggested by the lack of scoring criteria, referral criteria, directions for administering the items, and descriptions for any testing materials.

Cautions and Comments by the Authors

The chart does not contain any cautions or comments about its use.

Reviewer's Overall Judgment of the Measure

Assets. The chart may be used by a wide variety of professionals, requires little time to administer, covers children in a wide age range, and contains enough items per age interval to adequately screen development in the designated categories.

Items in the language category are not restricted to the child's demonstration of receptive and expressive skills but also provide some information about organic development which may be useful in determining the cause of language delays.

Limitations. This instrument is limited by its lack of validation and reliability data.

Bibliography

There are no references listed on the chart.

References

Anderson, R. M., Miles, M., & Matheny, P. A. Communication Evaluation Chart from Infancy to Five Years. Cambridge: Educators Publishing Service, 1963.

Delco-Elfman Developmental Achievement Test

Author	Rose Marks Elfman, Ph.D. Delaware County Intermediate Unit 6th and Olive Streets Media, Pennsylvania 19063
Publisher	Same as above.
Date of Publication	1974 (Clinical and Research Edition)
Materials Available	Manual Summary and Profile Sheets A preassembled kit of test materials is not available from the publisher; the small number of required materials should be locally available or are usually present in infant/preschool programs.

Purposes/Traits Assessed

The achievement test may be used for assessment of current level of functioning and as a basis for individual program planning, according to Elfman. Areas of development assessed by the instrument are physical (mobility and visual motor), social (oral communication, self-help, interpersonal), and intellectual (information, cognition, oral comprehension/memory).

Intended Test Population

Children functioning between the ages of 6 months and 6 years may be evaluated by this test.

Individual or Group Administration and Time Required

The test is individually administered to the child when s/he is alert and cooperative. If more than one testing session is required, the sessions should be scheduled within a one-week period and more more than a two-week period.

Instrument Format

Tasks are developmentally listed by area and subscale on the record form and in the administration directions; only one task per subscale is included for each 6 month age interval.

A "child's work sheet" is included on the back of each record form. The child is asked to copy a circle, a square, and a rectangle and complete a drawing of a man in a limited amount of space. Two stimulus pictures are included in the manual; one is composed of eight pictures for the child to verbally identify, the other is the figure of a boy to be used by the child

for pointing to body parts.

The child responds to tasks in the physical area by performing gross motor (e.g., walking, climbing, jumping) and fine motor (e.g., grasping, block building, drawing) skills. In the social area, the child demonstrates self-help skills (e.g., eating, grooming, toileting), peer interaction skills, and verbalization. Behaviors required in the intellectual area are responding to sound, pointing, matching, following directions, and verbalizing.

Administration Procedures

Specialized training in the use of this test is not required for those professionals familiar with children in this developmental age range. This group may include psychologists, educators, teachers, social workers, and guidance counselors.

Administration procedures and directions to be given to the child are included in the manual for individual tasks. Information is not provided for determining where to begin testing; testing is discontinued after four consecutive failures in a subscale.

Scoring Procedures

There are no special instructions for learning to score the test and only limited scoring criteria are included.

Tasks are scored as passed or failed; recorded scores include a raw score for each category, an Age Achievement Score for each category, a Total Score, a DEDAT Age, and a DEDAT Quotient.

The raw score for each category is the sum of the number of items mastered; each category Age Achievement Score is obtained by multiplying the raw score by 5. The Total Score is the sum of all passes for all areas; the Total Score, divided by 16, results in the DEDAT Age. The DEDAT Age, divided by the child's chronological age, yields the DEDAT Quotient.

Interpretation Procedures

There are no special training requirements and no guidelines for interpreting results or developing an educational program in the manual.

When a handicapping condition results in the omission of a subscale (e.g., mobility) in the evaluation procedure, the method used for calculating the DEDAT Age Score is modified to reflect that omission. For example, when one subscale is omitted the Total Score is divided by 14 rather than 16 (the number of tasks per year).

Item Selection

The criteria for choosing the tasks for inclusion on the achievement

test are not discussed in the manual.

Validity

Validity studies for this measure are not reported in the manual; no studies were available in a review of the child development literature.

Tasks included in each subscale appear to evaluate development in that category. Age classifications may be slightly low for some items at the upper age limits (e.g., define three out of four words).

Reliability

Reliability studies are not reported in the manual or in the literature for the achievement test.

The inclusion of scoring criteria for the tasks would appear to increase the reliability of scores obtained by the measure.

Cautions and Comments by the Author

The achievement test should not be administered by anyone unfamiliar with young children's development.

Concrete illustrations of the child's behavior should be included for interpreting test results.

Evaluation which takes place over a period of time greater than two weeks will be inaccurate due to a change in figuring the child's chronological age.

Reviewer's Overall Judgment of the Measure

Assets. The instrument may be used in programs serving a wide age range of developmentally young children and may be administered by anyone familiar with the development of young children.

The score sheet is simple to use, scoring is objective, and scores are simple to calculate.

Limitations. The lack of standardization of norms reduces the value of Age Achievement Scores, DEDAT Ages, and DEDAT Quotient.

The limited number of tasks per area would appear to make this instrument of more value as a screening test rather than as an achievement test and a guide for program planning.

Criteria for interpreting results and guidelines for using the results for planning purposes would be useful. A list of required materials by subscale and/or age interval would reduce examiner preparation time.

Bibliography

There are no references included in the manual.

References

Elfman, R. M. Delco-Elfman Developmental Achievement Test. Media, PA: Rose Marks Elfman, 1974.

The Early Learning Accomplishment Profile for Developmentally Young Children (Birth to 36 Months)

Authors	M. Elayne Glover, Jodi L. Preminger, and Anne R. Sanford
Publisher	Kaplan Press 600 Jamestown Road Winston-Salem, North Carolina 27103
Date of Publication	1978
Materials Available	Manual Scoring Forms Some testing materials (e.g., picture cards and manipulatives) are available upon request from the publisher. The remainder of the materials are familiar to most children and should be locally obtainable if desired.

Purposes/Traits Assessed

The criterion-referenced EARLY-LAP is an assessment instrument which evaluates development in the gross motor, fine motor, cognitive, language, self-help, and social-emotional areas.

Intended Test Population

The profile was designed for use with children who are functioning in the birth to 3 year developmental range, although the self-help category contains tasks designed to evaluate children up to 5 years of age.

Individual or Group Administration and Time Required

Administration and scoring time guidelines are not provided in the manual for this individually administered instrument.

Instrument Format

The profile consists of administration directions, references, a materials list organized by subscales, a glossary, the score form, forms for the individual education program, and the "Accomplishment Profile."

The score form includes the age at which the child is expected to successfully perform each task, the procedure for administering the item, the criteria for passing the item, and space to record the results of the pre-evaluation, the date the behavior was mastered, the result of the post-evaluation, and any comments on the child's performance of that task.

Reflexive behaviors, sitting skills, locomotor skills, and balancing

skills are representative of the 88 tasks on the gross motor subscale. Grasping ability, solving object permanence tasks, completion of a form-board, and pre-writing tasks are among the 73 items included in the area of fine motor development. Development in the category of cognition (105 tasks) is evaluated by such tasks as responding to sound, following directions, overcoming obstacles to obtain materials, and naming objects. Expressive and receptive language development (59 items) is evaluated by behaviors which include pointing to or naming objects, using pronouns, and reciting nursery rhymes. Development in the self-help area (49 items) is assessed through feeding, dressing, and toileting behaviors. The 38 tasks included on the social-emotional subscale include awareness of others, understanding of the word "no," response to his/her own name, and playing simple games.

Administration Procedures

Specialized training in the use of the EARLY-LAP is not required, although knowledge of child development would be beneficial; the manual includes a glossary of child development terms (e.g., palmar grasp).

Directions for presenting the tasks to the child are included in the manual but are usually written in general terms rather than using standardized directions.

There is not a suggested sequence for administering the profile items. Observation of the child's present behaviors is the suggested basis for determining where to begin scoring tasks on the profile, except for the handicapped child who should be evaluated beginning at half the child's age.

Scoring Procedures

There is no specified training procedure to learning how to score the items; criteria are included for each task to determine whether the child passes the item, fails the item, or exhibits an emerging skill.

Behaviors in the category of social-emotional are scored on the basis of being present or absent.

The basal level is established by eight consecutive passes (in each area of development) and assessment continues until a ceiling level is determined (three failures in five consecutive subscale tasks). Scores are graphed on the "Early Learning Accomplishment Profile" to indicate the child's current developmental status and to moniter progress from pre-to post-evaluation.

Interpretation Procedures

The profile does not contain any instructions for interpreting the results of the assessment.

Strengths and weaknesses for each category of development may be

recorded on a form included in the profile. Additional forms are provided for developing individualized education programs and recording the provision of special services.

Item Selection

Tasks were chosen to represent developmental milestones in the birth to 3 year age interval in a format similar to the Learning Accomplishment Profile and are a revision of the Learning Accomplishment Profile for Infants (1974-75).

Tasks included on the profile were chosen from norm-referenced assessment instruments which are in wide use and generally regarded as valid and reliable by most early childhood specialists (e.g., the Vineland Social Maturity Scale, the Slosson Intelligence Test). Credit is given for the source of each task included on the profile.

Validity

There are no validity studies reported in the manual and none were found in the developmental literature.

Reliability

Studies of reliability for the profile are not reported in the manual nor was information found to be available in current literature.

The lack of specific procedures for administering all of the tasks may reduce the reliability of the instrument.

Cautions and Comments by the Authors

Assessment should initially take place soon after the child's entry into the program, there should be ongoing assessment, and a final evaluation should be made at the end of the school year.

Reviewer's Overall Judgment of the Measure

Assets. The EARLY-LAP may be used, without extensive training, by early childhood professionals in the ongoing assessment of developmentally young children.

Limitations. The convenience of the measure could be increased by the color coding of skill areas and division of the materials list by age range (materials are currently listed by category only). For example, the list of materials for the fine motor section contains 39 objects. By dividing the list by area and age interval, the examiner could reduce the amount of time required for assembling testing materials prior to the assessment session.

Bibliography

The bibliography contains 26 references which were used in the

selection of tasks for inclusion on the profile. Examples of these references include the Merrill-Palmer Scale, the Bayley Scales of Infant Development, and the Hawaii Early Learning Guide.

References

Glover, M. E., Preminger, J. L., & Sanford, A. R. The Early Learning Accomplishment Profile for Developmentally Young Children. Winston-Salem: Kaplan Press, 1978.

EMI Assessment Scale

Author	Wanda B. Elder, Ed.D. Education for Multihandicapped Infants Project Department of Pediatrics University of Virginia Medical Center
Publisher	Department of Pediatrics University of Virginia Medical Center Box 232 Charlottesville, Virginia 22901
Date of Publication	1975 (Field Test Version)
Materials Available	EMI Assessment Scale $1.25 No testing equipment is available from the publisher but should be locally available.

Purposes/Traits Assessed

The EMI Assessment Scale is a "norm-referenced" checklist to be used for diagnosis and program planning in the following areas: gross motor skills, fine motor skills, socialization, cognition, and language.

Intended Test Population

The checklist is intended for use with multihandicapped infants functioning in the developmental age range from birth to 24 months.

Individual or Group Administration and Time Required

The checklist is individually administered. Time guidelines are not available for administering the scale; however, the checklist format suggests a short time requirement.

Instrument Format

The checklist is composed of 360 tasks with three behavior items in each developmental category for each month of age.

Gross motor tasks include reflexive responses and locomotor, catching, throwing, balancing, and climbing skills. The fine motor category includes visual tracking behaviors, manipulation of small objects, puzzle skills, and pre-writing skills. The items in the socialization section include feeding skills, interaction with adults, displaying emotions, dressing skills, and dramatic play behaviors. Cognitive tasks include problem solving behaviors, response to a verbal stimulus, verbalization, and matching objects. Tasks in the language section include locating sounds, vocalization, imitation of sounds, pointing to body parts, use of pronouns, and labelling familiar objects.

Administration Procedures

The checklist does not contain any information concerning training in administering nor any directions for administering the items.

Scoring Procedures

Training in scoring the scale is not specified; criteria for scoring the tasks are not included in the checklist.

Interpretation Procedures

No information is provided concerning the interpretation of scores obtained through the use of the scale.

Item Selection

Tasks were chosen for inclusion on the scale based on instruments in print and from 2 years of observations by the EMI staff of multihandicapped infants.

Validity

The scale does not contain any information concerning validity nor was any information found in published literature.

Reliability

Reliability information is not available from the information provided with the checklist nor discovered in developmental literature.

Reliability would appear to be low due to the lack of administration directions, scoring criteria, and description of testing equipment.

Cautions and Comments by the Author

The author suggests the instrument may be used for program planning and comparisons of individuals or groups of infants. Guidelines are not included for these suggested purposes.

Reviewer's Overall Judgment of the Measure

Assets. Age placement of tasks appears to correspond with other current instruments.

Limitations. The scale is seriously limited in its value as a diagnostic instrument by its lack of administration and scoring directions, and validity and reliability data. It is described as a norm-referenced test, but normative data are not included in the manual.

Although a bibliography of current evaluative instruments is included to indicate the sources of the tasks on the scale, individual items are not referenced as to their source.

Bibliography

A total of 15 instruments are listed in the bibliography. Some of these instruments include the Bayley Scales of Infant Development, the Vineland Social Maturity Scale, and the Denver Developmental Screening Test.

References

Elder, W. B. EMI Assessment Scale. Charlottesville, VA: University of Virginia Medical Center, 1975.

KNI Developmental Scale

Author	Eldene Woellhof
Publisher	Psychology Department Kansas Neurological Institute 3107 West 21st Street Topeka, Kansas 66604
Date of Publication	1965
Materials Available	Manual Profile Graph Score Sheets Testing materials are not available from the publisher; the few materials required should be locally available.

Purposes/Traits Assessed

This criterion-referenced screening instrument is to be used in determining the level of functioning of the "retarded" child in comparison with the expected development of "normal" children.

The scale is composed of 80 items which are intended to evaluate the child's functioning in the areas of socialization, communication, physical development, and self-help skills.

Intended Test Population

The targeted group for evaluation with the scale are institutionalized retarded children functioning in the birth to 5 year age range. This scale may also be appropriate for use with physically handicapped and emotionally disturbed children, according to the author.

Individual or Group Administration and Time Required

The scale was designed for evaluating the child on an individual basis. There is no information available for determining the amount of time required to administer the instrument.

Instrument Format

Tasks have been listed by developmental age intervals using a 3-month time span for the first 2 years and a 6-month time span for the third to the fifth years. For each 3-month interval, there is one task per category; two items per category are included for the 6-month intervals.

The child responds to the various tasks by exhibiting reflexive behaviors, vocalizing, performing gross motor skills (e.g., walking, kicking, climbing), demonstrating an understanding of verbal requests, producing language, and showing an interest in others.

The two page scoring form lists tasks in developmental sequence and by category. The circular "Profile Graph" allows scores to be plotted by developmental age levels and may be used for successive evaluations.

Administration Procedures

Special training in the use of the measure is not required; cues are provided for the examiner's actions and directions to the child.

If the interview technique is used with the scale, the examiner is referred to the procedure used on the Vineland Social Maturity Scale (begin below the chronological age of the child, administer the items by groups rather than by individual items, try to obtain information about the child's routine performance rather than his/her optimum performance, and follow up general questions with detailed questions).

There are no suggestions for sequencing the categories for administration/interviewing. Information is not provided for determining a basal level or for discontinuing testing prior to the final listed task.

Scoring Procedures

While knowledge of children's development would appear to be beneficial in scoring the scale, training procedures in its use are not specified by the author.

Individual tasks are scored as passed, inconsistent, or absent; criteria are provided for scoring each item.

An estimated developmental age for each area may be calculated by multiplying the number of passes in each of the four categories by 3 months (example provided in the manual). A total score may be obtained by multiplying the total number of items passed by 3/4 month (example provided in the manual).

Interpretation Procedures

Since guidelines are not provided in interpreting evaluation results, referral for additional diagnosis is left to the examiner's own interpretation.

The Profile Graph may be used to visually illustrate the items passed by the child (circled on the graph) and any "deficiencies in the progressive steps of development" (Woellhof, 1965, p. 14). Results may be plotted over an extended period of time.

Item Selection

Tasks included on the scale were selected from five standardized assessment instruments: Cattell Infant Intelligence Scale, Gesell Developmental Schedules, Merrill-Palmer Scale of Mental Tests, Rosenzweig Behavior Profile, and Vineland Social Maturity Scale. Individual credit is not provided in the manual for the source of each item.

Validity

The manual does not provide any information regarding the validity of this measure and no information appeared to be available from any other published source.

The selection of items from well known standardized instruments would appear to support the validity of the scale. Tasks appear to be categorized appropriately for area and age level.

Reliability

Reliability data were not provided in the manual nor were data found from any other source.

Cautions and Comments by the Author

Standardized developmental age levels can not be derived for various areas since items were pulled from various sources and norms have not been established for the instrument.

Scores should be based on observable behaviors rather than on caregiver opinions.

Any special handicapping conditions should be noted on the record sheet.

Reviewer's Overal Judgment of the Measure

Assets. The record sheet provides sufficient cues to the examiner who is familiar with the manual and test items so that the manual does not need to be used during the evaluation session.

Limitations. This instrument lacks any validation and reliability data. It would appear that reliability of the measure could be increased by the inclusion of standardized administration procedures, a complete description of testing materials, and an indication of those items which can be passed by parental report and which must be passed by observation (if any).

Bibliography

In addition to the five instruments mentioned earlier, the bibliography contains an additional five references.

References

Doll, E. A. Vineland Social Maturity Scale. Circle Pines, MN: American Guidance Service, Inc., 1965.

Woellhof, E. KNI Developmental Scale. Topeka: Kansas Neurological Institute, 1965.

The Lexington Developmental Scale (Short Form)

Authors	Staff of United Cerebral Palsy of the Bluegrass Child Development Centers, Inc.
Publisher	The Child Development Centers of the U.C.P.B. P.O. Box 8003 465 Springhill Drive Lexington, Kentucky 40503
Date of Publication	1974
Materials Available	Manual Score Forms The manual contains the work sheets for use in testing; other materials should be commonly available in infant/preschool programs.

Purposes/Traits Assessed

The scale was designed for three primary functions: 1) to evaluate four areas of development (motor, language, personal-social, and cognitive); 2) as a basis for intervention; and 3) as a screening device for referral purposes.

Intended Test Population

The scale is intended for use with those children functioning in the age range of birth to 6 years.

Individual or Group Administration and Time Required

The scale is suitable for testing the individual child in a clinic or home setting. The usual amount of time required for administering the tasks is 30 minutes.

Instrument Format

Tasks are developmentally sequenced and categorized by developmental area in the administration directions and on the scoring forms.

Items are placed 6-month intervals for the first 2 years and in 1-year intervals for the third through the fifth years. The motor category contains seven tasks per interval, the language category contains six, and the cognitive and personal-social areas each contain five, for a total of 1984 tasks.

Responses to these 184 tasks are varied. Tasks in the motor category include grasping, throwing, kicking, drawing, and skipping. Skills in the language area are demonstrated through response to sound, babbling, pointing to objects, repeating digits, and following directions. Behaviors representing items on the cognitive subscale are manipulation of objects,

identifying letters in the alphabet, and naming colors. Dressing skills, peer interaction, and dramatic play behaviors are evaluated in the personal-social area.

Administration Procedures

It is suggested that the scale is appropriate for use by nurses, health educators, social workers, social service aides, and homemakers without any special training in its use.

Biographical data and the child's chronological age are recorded on the front page of the score form. Gross motor items are evaluated first, followed by fine motor, language, cognitive, and personal-social tasks, respectively. The child is allowed two trials to complete each task.

Scoring Procedures

The manual does not specify any training for learning to score the scale.

The score form has short descriptions of all the tasks, presented in "age blocks." Failed tasks are scored with an "X" at the left of the age block for that item, in the middle of the block to indicate partial success, and at the right of the block to indicate mastery of the task. Results are graphed by connecting the X's with straight lines, an example is provided in the manual. Different colors should be used to indicate successive evaluations.

Criteria for mastery of the tasks are included for some of the items. Examples of passes and failures are included in the manual for scoring the child's ability to copy shapes.

Interpretation Procedures

There is no special training required for interpreting the results of the evaluation. Referral for more diagnostic assessment is recommended if the child shows an "erratic pattern of development" or if an average lag of 6 months to 1 year of development is observed (except for the premature infant).

A follow up parent conference should be conducted upon completion of the evaluation to explain results and make any necessary recommendations.

Item Selection

References for developmental norms include the Bayley Scales of Infant Development, the Preschool Attainment Record, the Denver Developmental Screening Test, and the Slosson Intelligence Test. However, credit is not given for the source of individual tasks.

Testing of the scale was also carried out by the staff of the Child Development Center, although descriptions of the studies are not reported in the manual.

Validity

The authors claim that validity of the measure is supported by current literature and results of use by the center's staff. It is also reported that language and cognitive scores on the Lexington Developmental Scale correlated with scores on the Stanford-Binet Intelligence Scale. Unfortunately, data are not included in the manual to support these claims.

Placement of tasks, by category and age, on the scale appear to be appropriate according to current developmental literature.

Studies concerning the validity of this instrument apparently have not been reported in current developmental literature.

Reliability

Information regarding the reliability of this instrument was unavailable in the manual and was not found in other published sources.

The lack of specific administration directions, scoring criteria, and the lack of standardization of testing materials would appear to reduce the reliability of the results obtained through the use of this scale.

Cautions and Comments by the Authors

The child's scores should indicate the performance level at which s/he does not require any assistance on tasks. The child should not be passed on an age interval until all items in the preceding level have been mastered. The child should not be tested on any items s/he could not be expected to pass due to a handicapping condition.

If scores are high in most areas, referral may not be necessary, and parents may be encouraged to work with the child on a few weak skills. Reevaluation should take place in six to nine months.

Reviewer's Overall Judgment of the Measure

Assets. The scale should be quickly and easily administered by anyone familiar with the development of young children. The format of the score form appears to make it convenient to use. The length of time for administration is appropriate for the age of the children being evaluated.

Graphing scores on the score form provides a good visual indication of the child's strengths and weaknesses.

Testing materials will be familiar to most children, are readily available for purchase, and easily trasportable to a testing site.

Limitations. Evidence of high validity or reliability has not been established for this instrument.

Bibliography

A total of 29 references are included in the bibliography. Included

among these are the work of Gesell, Piaget, Bowlby, and Smart and Smart.

References

The Lexington Developmental Scale. Lexington, KY: The Child Development Centers of the U.C.P.B., 1974.

MEMPHIS Comprehensive Developmental Scale

Authors	Alton D. Quick, Thomas L. Little, and A. Ann Campbell
Publisher	Fearon-Pitman Publishers, Inc. 6 Davis Drive Belmont, California 94002
Date of Publication	1974
Materials Available	Programming Guide $1.00 Instruments for Individual Program Planning and Evaluation (10 each of diagnostic, prescriptive, and evaluative forms) $7.95 Enhancing Developmental Progress in Preschool Exceptional Children (160 page handbook explaining the history and goals of the MEMPHIS program, and how to set up a similar program) $7.50 Lesson Plans for Enhancing Preschool Developmental Progress (This 544 page book containing 260 sample lesson plans is currently out of print and is in the process of being revised). A preassembled kit of materials is not available from the publisher, however, required materials are usually found in most infant/preschool programs.

Purposes/Traits Assessed

The scale is the first step in a screening and programming model used to determine functioning in five developmental areas: personal-social, gross motor, fine motor, language, and perceptuo-cognitive.

Intended Test Population

The targeted population are those children functioning between birth and 5 years of age with suspected or confirmed handicaps.

Individual or Group Administration and Time Required

The scale was designed for "quick" administration to the individual child using an observation or parental interview technique.

Stimulus Items

Items on the score form are arranged by categories in developmental sequence. The format of each area is designed to provide biographical information about the child, age norms for tasks, a scoring column, the name of the task, and the child's progress as indicated by later evaluation.

The mode of response is varied for the scale. The personal-social area contains three tasks for each three month age range; skills include feeding, toileting, and inter-personal relationships. Behaviors evaluated in the gross motor category (2 tasks per age range) include sitting, walking, balancing, climbing and catching. The fine motor area (2 tasks per age range) assesses grasping, block building, drawing, and cutting behaviors. Language skills (3 items per age range) include response to sound, pointing to and labelling objects, use of plurals, singing, and relating stories. Problem solving skills, labelling, counting, sorting, matching, and puzzle skills are evaluated in the perceptuo-cognitive category (3 tasks per age range).

Administration Procedure

The scale is intended for use by teachers without additional training; the only assistance provided for administering the scale comes from the cues provided by the item name.

Initial evaluation begins with the administration of the earliest developmental item for each category and continues until the child fails six or more consecutive tasks in each area.

Scoring Procedures

Extensive training is not required for scoring the scale; criteria for scoring the tasks are not included in the manual and appear to be the examiner's responsibility to establish.

Individual tasks are scored as passed or failed; the total number of passes for each area is the raw score for that category. The item number corresponding with the total number of passes (the raw score) is then found on the score form for each category; the age interval in which this task occurs is recorded as the child's developmental age (e.g., a raw score of 35 is obtained on the personal-social subscale, the 35th task falls at the 3.00 age level or 36 months). Developmental ages are computed for each subscale and are then plotted on the "Profile of Developmental Status" on the cover of the scale, along with the child's chronological age.

Interpretation Procedure

Additional training is not required for interpreting screening results.

When the developmental age line falls below the chronological age line on the "Profile of Developmental Status," delay is indicated for those developmental areas. The child is developing "normally" in those areas where the developmental age line is at or above the chronological age line.

Screening results are "rough estimates" of the child's functioning and should be the basis for individual program planning.

Item Selection

Skills were chosen that were viewed as important to later school learning

(criteria are not reported in the manual).

Validity

Data regarding the validity of the scale is unavailable in the manual, in the handbook, or in the child development literature.

Reliability

Neither the manual or the handbook contains any information concerning the reliability of the scale; in addition, no studies of reliability have been reported in the early childhood literature.

Cautions and Comments by the Authors

Project MEMPHIS was developed as a prototype and may be used by other programs in part or in totality.

Screening information is to be used "for educational program planning and not for psychological or developmental diagnosis" (Quick, Little, & Campbell, 1974, p. 36).

The scale is the first of three steps of the MEMPHIS Model. It is used for the initial evaluation, the basis for program planning (step 2), and for re-evaluation (step 3) at the end of the instructional period.

Reviewer's Overall Judgment of the Measure

Assets. Project MEMPHIS was designed by the Department of Special Education and Retardation at Memphis State University for use by programs serving exceptional preschool children. Enhancing Developmental Progress in Preschool Exceptional Children (1974) describes the process of designing a program, training staff and parents, using the Comprehensive Development Scale, and provides sample lesson plans.

Limitations. The Scale is supposed to be used for screening, program development, and re-evaluation, which would in reality be an ineffective use of a screening instrument.

The manual does not contain any standardization criteria or data necessary for judging the appropriateness of the instrument for other groups of children or to provide support for determining developmental ages.

Bibliography

A bibliography is not included in the manual/programming guide. Enhancing Developmental Progress in Preschool Exceptional Children (1974) includes a bibliography of selected readings in the areas of theory, curriculum, parent training, development, administration, and program evaluation.

References

Quick, A. D., Little, T. L., & Campbell, A. A. Enhancing Developmental Progress in Preschool Exceptional Children. Belmont, CA: Fearon Pitman, 1974a.

Quick, A. D., Little, T. L., & Campbell, A. A. Guide to Programming. Belmont, CA: Fearon Pitman, 1974b.

The Milani-Comparetti Motor Development Screening Test

Author	P. Pearson, L. Rice, and J. Trembath	
Publisher	Meyer Children's Rehabilitation Institute University of Nebraska Medical Center Omaha, Nebraska	
Date of Publication	1977	
Materials Available	Manual	$6.00
	Videotape describing the test and demonstrating the administration procedures.	$82.00 Purchase $21.00 Rental
	Scoring chart	
	No testing materials are required.	

Purposes/Traits Assessed

The norm referenced screening test was designed to assess motor development during the first two years of life to determine and identify neuromotor delay or deficit.

Intended Test Population

The test is appropriate for use with children functioning in the birth to 2 year age range. The authors suggest the following as "key ages" for evaluation: 1½ to 2 months, 3½ to 4 months, 5½ to 6 months, 9 to 10 months, 12 months, 15 months, 18 months, 21 months, and 24 months.

Individual or Group Administration and Time Required

The test is individually administered and should require only four to eight minutes to complete.

Instrument Format

Reflexes and the child's degree of motor control are evaluated by assessing control over the head, sitting posture, parachute reactions, standing, locomotion, hand and foot grasp, Landau response, symmetric and asymmetric tonic neck reflexes, body derotation and rotation, Moro reflex, and tilting reactions. There is a total of 28 items on the test.

Administration Procedures

The manual appears to contain enough information to learn to use the test, although training should include viewing the videotape demonstrating the administration procedures. Suggested users of the test include physicians, nurses, therapists, and physicians' assistants. A thorough knowledge of reflexive behavior is important in observing and judging responses.

The manual includes directions for eliciting responses, illustrations of the appropriate responses, and age guidelines for each task; examples of advanced or delayed responses are also provided.

Testing is to begin "at the level of motor development that the child has attained" (Pearson, Rice, and Trembath, 1977, p. 16). Testing may generally be discontinued at the 1 month to 1½ month age interval above the infant's chronological age.

Scoring Procedures

The manual and videotape provide the necessary information for learning to score the test.

Responses are scored as present or absent except when the test item is indicated to be an emerging reaction. The general state of the child is noted at the top of the chart. A sample scoring chart is included in the manual. The chart is divided by vertical lines into half month increments from birth to 12 months and in one month increments from 12 to 24 months. Shaded areas in the chart indicate when a reflex or reaction can be expected to be observed and tapered areas indicate a reaction which is gradually acquired.

The child's age, in months, is used to score each item and is placed on the vertical line corresponding to the child's performance on the test item (e.g., if a 9-month old passes a 12-month item, a 9 is recorded on the 12-month line). Examples are provided in the manual. Deviation (delayed or advanced) are marked with asterisks; examples of deviant responses are provided with a description of how the item is to be scored.

If a behavior can not be observed, then the mother may be asked about the child's response; the mother's response is scored on the chart with the notation "mother report."

Interpretation Procedures

The manual does not specify training requirements for interpreting the results of the test.

While prematurity is not compensated for when determining chronological age, it is considered when interpreting the significance of delays.

"An infant who is suspected of showing deviations should be tested several times if possible. The younger the baby, the more difficult it is to draw conclusions about significant deviations. Over subsequent test periods, delays or asymmetrical development becomes more obvious" (Pearson et al., 1977, p. 22).

"Babies who show delays and/or asymmetrical motor development in the early months of life may 'catch up' or 'even out' near the end of the first year. Any child between 9 and 12 months of age with significant delay and/or asymmetry should also be given a neurological evaluation and other developmental assessments" (Pearson et al., 1977, p. 22).

Item Selection

The original chart was developed by Milani-Comparetti and Gidoni (1967) for the examination of children's reflexes and early motor patterns to assess the attainment of developmental milestones. Modifications were made based on five years of use in a child welfare clinic.

The chart included for use with the screening test was modified by the staff of the Meyer Children's Rehabilitation Institute for more "efficient" use but the original administration and scoring procedures were unchanged.

Although this screening test provides "norms" for particular behaviors, no data are reported in the manual concerning the establishment of these age levels.

Validity

Information concerning the validity of this screening test is unavailable in the manual found in, or the literature.

The age ranges for the behaviors included on the instrument are comparable to age ranges for these behaviors for such instruments as The Revised Developmental Screening Inventory, the Assessment-Programming Guide for Infants and Preschoolers, the Brigance Diagnostic Inventory of Early Development, the Denver Developmental Screening Test, and the Bayley Scales of Infant Development.

Reliability

The authors do not report any reliability data in the manual for this instrument nor was any information found in the literature.

Cautions and Comments by the Authors

The authors suggest screening children several times during the first two years of life and recommend "key ages" for this screening. Where consistent delays, asymmetry, or abnormal development is exhibited in repeated testing, more frequent testing is indicated with motor assessment and neurological examination to be considered.

The child's "general state" should be noted on the profile sheet because of its influence on his/her response to the tasks.

Reviewer's Overall Judgment of the Measure

Assets. The instrument may be quickly administered by medical personnel or physical therapists without requiring any testing materials. The shaded areas on the score sheet make it possible to quickly determine any delays in behaviours.

Limitations. It would be desireable to have a report of the data used to establish "norms" or age guidelines.

Bibliography

References for this instrument include the work of Gesell and Amatruda and Bayley. However, the major influence in the development of this instrument was the work by Milani-Comparetti and Gidoni.

References

Milani-Comparetti, A., & Gidoni, E. A. Pattern analysis of motor development and its disorders. Developmental Medicine and Child Neurology, 1967, 9, 625-630.

Milani-Comparetti, A., & Gidoni, E. A. Routine developmental examination in normal and retarded children. Developmental Medicine and Child Neurology, 1967, 9, 631-638.

Pearson, P., Rice, L., & Trembath, J. The Milani-Comparetti Motor Development Screening Test. Omaha: Meyer Children's Rehabilitation Institute, 1977.

Portage Guide to Early Education (PGEE)

Authors	Susan M. Bluma, Marsha S. Shearer, Alma H. Frohman, and Jean M. Hilliard
Publisher	Portage Guide to Early Education The Portage Project Cooperative Educational Service Agency 12 626 East Slifer Street Portage, Wisconsin 53901
Date of Publication	1976 (Revised Edition)
Materials Available	Manual Checklist of behaviors Card file (Methods for teaching behaviors) Carrying case Test materials are generally found in the classroom.

Purposes/Traits Assessed

The Portage Guide was designed to evaluate a child's current level of functioning in order to plan an educational program. The five developmental areas covered are: self-help, motor, socialization, language, and cognitive. In addition, there is a section on infant stimulation.

Intended Test Population

The guide may be used with normal or handicapped children functioning within the birth to 6 year mental age range and with older children and adults functioning at the preschool level.

Individual or Group Administration and Time Required

The checklist may be completed by using results from standardized tests, from a parent interview, or by observing the child. The initial evaluation of the individual child is made at the time the child enters the program. Re-evaluation may be made individually or in a group situation. Time requirements will vary depending on the method of evaluation.

Instrument Format

Tasks are developmenatlly sequenced in each category. Items are divided into 1 year age intervals with each interval containing from 8 to 45 tasks. The infant stimulation section contains 45 items to be assessed. Totally, there are 580 behaviors listed.

Directions for completing the checklist and using the card file are included in the manual. The color-coded checklist has space on the front cover for biographical information. Each task is identified and the age level is provided; space is provided to score the behavior as present or

absent, to record the date the skill was achieved, and to record any comments about the child's performance. Cards in the file are color-coded, and include information on the developmental area and the the number of the behavior, the age level of the task, the title of the behavior (stated as a behavioral objective), and teaching suggestions.

The teacher has flexibility in choosing materials from those available in the classroom for assessing the child's level of functioning and for teaching the behavior as part of the child's educational program.

The child's functioning in the self-help category is evaluated by eating (e.g., reaches for bottle, uses spoon and cup with some spilling, feeds self entire meal), toileting (e.g., uses words or gestures indicating need to go to bathroom), and dressing tasks (e.g., holds out arms and legs while being dressed). Tasks in the language section include items evaluating both receptive and expressive abilities. The socialization area includes items assessing interaction with people, manipulation of objects, and motoric responses to requests. The cognitive category requires verbal and pointing responses to indicate understanding of classification, number, and pre-academic skills. The motor area covers fine and gross motor development and includes eye-hand coordination, pre-writing, balancing, and catching skills.

Administration Procedures

The measure is suggested as appropriate for use by teachers, aides, nurses, and parents without the need for special training in its use.

Initial assessment (in each category) for "normal" children should begin at one age interval below the child's chronological age. If delay is suspected for any area, begin testing at two age levels below the chronological age. When the child is unable to perform 10 to 15 behaviors at this level, the examiner should administer tasks at a 10 point lower level. Testing is discontinued when the child misses 10 to 15 sequential items in a category.

Information is provided for writing behavioral objectives, task analysis, implementing tasks, and correcting and reinforcing behaviors.

Scoring Procedures

Training in scoring the checklist is not required.

Scores are subjective due to a lack of criteria for passing the tasks. Behaviors present upon entry into the program are scored on the checklist with a checkmark. Behaviors which the child is no longer performing because s/he is using a developmentally superior behavior are credited as being present. Tasks which are administered to the child are scored as present or absent; the date the behavior is achieved is also recorded.

Interpretation Procedures

The guide does not require additional training for interpretation. The completed checklist is used by the teacher to target the skills to be taught.

Targeted skills should be the lowest level skills the child was unable to perform. "The decision would be based upon which skill the child was closest to being able to perform, which would be most functional for he child and which skill, when learned, would serve as a prerequisite to the greatest number of additional skills" (Bluma, Shearer, Frohman, & Hilliard, 1976, p. 24).

After the skill was targeted for learning, the teacher would find the corresponding card in the card file "read over the suggestions given and choose the one which would be most effective in teaching the child. This decision would be made based upon her knowledge of the child. His present level of funtioning and the type of aid which best facilitates his learning should be considered" (Bluma, Shearer, Frohman, & Hilliard, 1976, p. 25).

Item Selection

This measure is based primarily on the 1972 experimental edition. Criteria for developing the guide were developmental theory, encompassing several developmental areas, a method for recording behaviors achieved by the child, and methods for teaching these skills (bibliography included in the manual). Data to confirm this "developmental" sequence are unavailable in the manual or in current literature.

Validity

Information concerning the validity of this instrument is not reported in the manual or in any available literature.

Reliability

The authors do not include any data concerning the reliability of this measure in the manual and no reports were found in the child development literature.

The lack of a standardized format for administering the tasks and the subjectivity of the scoring process would suggest that the reliability of scores obtained for this instrument would not be very high.

Cautions and Comments by the Authors

While the checklist may be used to determine the child's present level of functioning, the teacher should know the child's learning style in order to make best use of suggestions for teaching new skills.

Receptive language behaviors have been included in the cognitive category due to overlap between the two areas; the teacher should plan to teach these behaviors in coordination.

The sections on teaching aids, behavioral objectives, and task analysis have been included to help the teacher plan for the individual child.

Reviewer's Overall Judgment of the Measure

Assets. Adequate information is provided for teachers to use the checklist with developmentally young children. Suggestions for developing and implementing individual program planning are excellent for this type of measure (e.g., developmental sequencing for targeting skills, specific to general teaching suggestions, task analysis for attaining complex behaviors).

Limitations. Color-coding of the checklist and the cards in the file and the size of the cards should be helpful in actual use of the guide.

Bibliography

The manual contains 26 references of assessment instruments (e.g., Bayley Scales of Infant Development) and learning theory. A recommended reading list of ten texts is also included for those interested.

References

Bluma, S.M., Shearer, M. S., Frohman, A. H., & Hilliard, J. M. Portage Guide to Early Education. Portage, WI: The Portage Project, 1976.

Uniform Performance Assessment System (UPAS)

Authors	Owen R. White, Eugene Edgar, and Norris G. Haring
Publisher	Charles E. Merrill Publishing Company 1300 Alum Creek Drive Box 508 Columbus, Ohio 43216
Date of Publication	1978
Materials Available	UPAS Program Kit: Tester's Manual, Items Manual, Stimulus Cards and vinyl carrying case. $70.00 UPAS Student Record Books (package of 12) $ 9.95 Testing equipment is not available from the publisher but may be ordered from school supply companies.

Purposes/Traits Assessed

The UPAS is a "curriculum-referenced" instrument which covers five curriculum areas: 1) preacademic/fine motor (fine motor development, visual perception and discrimination, pre-reading, and pre-math), 2) communication (receptive and expressive), 3) social/self-help (feeding, dressing, toileting, grooming, and play skills), 4) gross motor ("infant behaviors," locomotion, balance, and "large-muscle play skills"), and 5) behavior management (checklist of inappropriate verbal and physical behaviors).

Intended Test Population

The instrument was designed for use with handicapped children functioning in the birth to 6 years age range.

Individual or Group Administration and Time Required

Evaluation is made during large or small (two to three children) group activities and is completed over a 1 to 2 week period.

Instrument Format

Tasks are arranged by curriculum areas and in developmental order within skill sequences. Directions for administering tasks and record sheets are color coded by category. Items considered basic skills are indicated on the record sheet by circling the coded category; uncircled items are not considered high priority items. A red flag designation indicates those skills which, if failed, may suggest a handicapping condition. Items designated as indicators are those which assess a sample of a larger class of behavior (e.g., response to a two-action command).

Response modes to the tasks in the four categories are varied. Included on the pre-academic scale are visual perception, reaching, grasping, and releasing, perceptual motor skills (placing pegs in holes, hammering pegs, cutting, matching parquetry blocks), sequencing, completing puzzles, writing (scribbling, copying designs, printing letters), visual discrimination (matching, sorting, naming), pre-reading (rhyming words, discriminating vowels and consonants), and pre-math skills (rote counting, ordering numerals, matching quantities). The communication scale includes receptive (response to sound, following directions, discriminating between textures, and responding to prepositions) and expressive behaviors (production of vowels and consonants; labelling objects; use of verbs, pronouns, prepositions and adjectives; describing the functions of objects; and relating stories). Feeding dressing, toileting, grooming, play behaviors, giving personal information, and classroom work skills are among the tasks on the social/self-help scale. Tasks on the gross motor scale include head righting, sitting, crawling, standing, locomotion, stair climbing, jumping, balanceing, use of playground equipment, playing with a ball, and rhythm skills. A section is also included for evaluating the ability of a wheel chair-confined child to move about his/her environment.

The behavior checklist is a list of inappropriate verbal (talking out, crying, insults, talking to self) and physical behaviors (failure to follow teacher directions, self-abuse, breaking objects, tantrums).

A list of test materials (by curriculum area) is provided in the manual. Descriptions of some materials are specific (e.g., set of DLM large parquetry blocks and designs P114-3, P114-10, and P114-12).

Worksheets, included with the instrument, are illustrated with relatively small black and white stylized figures.

Administration Procedures

No formal training in the use of the UPAS is required to administer the instrument. Familiarity with handicapped preschoolers is desirable; the authors imply that evaluation is usually conducted by the classroom teacher.

Directions for each item include the skill description, the equipment needed, and the test/observation procedure.

Evaluation of the individual child occurs within a group situation. Examples are provided in the manual to illustrate how test items can be incorporated into the daily schedule or how an activity can be planned to combine several of the tasks.

Examples of adaptations to accommodate handicapping conditions are provided in the manual.

There is no suggested sequence for administering the tasks.

Scoring Procedures

No training beyond reading the tester's manual is required; familiarity with child development terms would be helpful.

Criteria for passing or failing each task are included for all areas except behavior management. Tasks are scored with a "yes" (passed), "no" (failed), or a number (indicating the correct number of responses or the number of seconds in which the task was performed).

Scoring instructions suggest additional information may be recorded for any given task. For example: 1) tasks may be noted as inappropriate due to a handicapping condition or the pupil's age, 2) instruction was provided to the child for mastery of this skill between evaluations, 3) modification of the testing procedure or the child's response was made due to a handicapping condition, 4) a prosthetic device (other than glasses or a hearing aid) was necessary to perform the task, 5) the task was omitted due to the absence of the child or the evaluator, 6) the lack of equipment made the item untestable, and 7) the skill is/is not an objective on the child's Individualized Education Program (IEP).

The behavior management section is scored through the use of a checklist. Behaviors are scored as absent, occasionally exhibited but not interfering, or frequent enough to interfere with other programs. The goal for items on the checklist is that none of the behaviors be exhibited by the child.

Interpretation Procedures

Training in the interpretazion of results is not specified in the manual.

Some items are designated as "red flags" indicating important skills. Failure on these skills at th- appropriate age indicates the child should be referred for diagnostic evaluation.

When testing motor items, the child's muscle tone should be noticed with anything unusual being referred to a therapist.

Information gathered from the test may be used in placement and program decisions.

After several evaluations the child's progress may be plotted on a graph on the back of each record sheet. The completed graph allows comparison of the child's progress to the expected development of a non-handicapped child.

Item Selection

Criteria for inclusion of items on the instrument included 1) appropriate for handicapped children, 2) appear in developmental order, 3) be acquired by most children in about the same amount of time, and 4) appear during birth and 6 years of age.

At the time of publication of this manual, three major revisions had been made based on computer analysis and tester feedback from early childhood handicapped programs in Washington state.

Validity

The authors did not include any data concerning the validity of this instrument in the manual nor are validity studies available in current literature.

Reliability

Reliability data for this measure are unavailable in the manual or in developmental literature.

Cautions and Comments by the Authors

Due to the the child's age some nonbasic skills (not necessary for independent living) may be inappropriate and should not be tested.

Reviewer's Overall Judgement of the Measure

Assets. The measure may be helpful in making placements (children with similar strengths and weaknesses could be placed together for instructional purposes), writing IEP's, developing curriculum, and judging the child's progress.

The format of the administration directions, color coding of categories of development, and the sequencing of skills in developmental order should facilitate the evaluation process.

Limitations. The lack of normative and support data makes it difficult to judge the appropriateness of this instrument for use with normal children if the measure is being used in a mainstreamed program.

The worksheets included for use with the manual are rather unattractive and some of the illustrations are quite small which may be a problem for children with some handicapping conditions.

References

White, O. R., Edgar, E., & Haring, N. G. Uniform Performance Assessment System. Columbus: Charles E. Merrill, 1978.

Appendix A

A Glossary of Terms

AGE EQUIVALENT SCORE: A score which is derived from a raw score and represents the average performance for a chronological age group based on normative data. (Goodwin & Driscoll, 1980; Salvia & Ysseldyke, 1978)

CONSTRUCT VALIDITY: A judgment of how well an instrument measures the theoretical construct or variable (e.g., creativity) it was designed to measure. This may be indicated from a correlation between test scores and the performance of behaviors indicative of that construct. (Goodwin & Driscoll, 1980)

CONTENT VALIDITY: A judgment about the relationship among test items and test objectives, the process in which tasks measure the content of the subject domain, and the completeness of the item sample. (Goodwin & Driscoll, 1980; Salvia & Ysseldyke, 1978)

CORRELATION: A statistical procedure for determining the degree of relationship between two variables for a group of individuals or the relationship between pairs of individuals on one variable. (Goodwin & Driscoll, 1980; Salvia & Ysseldyke, 1978)

CORRELATION COEFFICIENT: A numerical expression of the relationship between two variables. This index ranges from +1.00 (perfect positive relationship) through .00 (no relationship) to -1.00 (perfect negative relationsOip). The Pearson product-moment correlation coefficient (r) is commonly used as an index of these relationships for equal-interval scales. (Goodwin & Driscoll, 1980; Salvia & Ysseldyke, 1978)

CRITERION-REFERENCED TEST: A test designed to indicate whether a child has mastered particular skills. Performance is compared with predetermined criteria rather than with the performance of a peer group. (Goodwin & Driscoll, 1980; Salvia & Ysseldyke, 1978)

CURRICULUM-REFERENCED TEST: A test which measures performance of skills which are taught in the curriculum of a given program. (Brigance, 1978; White, Edgar, & Haring, 1978)

DEVELOPMENTAL QUOTIENT: The relationship between an individual's chronological age and developmental age (e.g., IQ score). (Salvia & Ysseldyke, 1978)

DEVELOPMENTAL SCORE: A method of transforming raw scores for the purpose of comparison with the performance of a norm group; two common types are age equivalents and grade equivalents. (Salvia & Ysseldyke, 1978)

INTERRATER RELIABILITY: The correspondence or agreement between two or more testers in their scoring of a set of tests. (Goodwin & Driscoll, 1980)

INTEROBSERVER RELIABILITY: A measurement of the degree to which independent observers agree in their recordings of a set of behaviors. (Goodwin & Driscoll, 1980)

MEAN: The arithmetical average, calculated by summing all of the scores and dividing by the total number of scores. (Goodwin & Driscoll, 1980)

MEDIAN: The point (score) that divides the group into two equal parts; half of the scores are higher and half of the scores are lower than this point. (Goodwin & Driscoll, 1980; Salvia & Ysseldyke, 1978)

NORM: A numberical description of the performance of a defined or reference group. (Goodwin & Driscoll, 1980)

NORM-REFERENCED TEST: A test designed to compare an individual's performance with that of his/her peers, where the test results obtain meaning from their relationship to the scores of others, with emphasis being placed on the rank ordering of individuals according to the number of skills mastered. (Goodwin & Driscoll, 1980; Salvia & Ysseldyke, 1978)

RANGE: A measure of variability of a group of scores; the distance between a set of scores including the extremes for th8t set. (Goodwin & Driscoll, 1980; Salvia & Ysseldyke, 1978)

RELIABILITY: A measure of the accuracy and consistency of scores or classifications based on the use of the instrument. (Goodwin & Driscoll, 1980; Salvia & Ysseldyke, 1978)

SCREENING TEST: A brief assessment procedure, designed to identify children who may need further evaluation and educational intervention.

SPLIT-HALF RELIABILITY: A technique used to determine internal consistency where items on a tesz are divided into two equal parts and the resulting scores from the two parts are correlated. (Goodwin & Driscoll, 1980)

STANDARD DEVIATION: A unit of measurement which can be used to express the average variability of scores around the mean in a normal distribution. (Hardyck & Petrinovich, 1976; Salvia & Ysseldyke, 1978)

STANDARDIZATION: Administration of a test to a large, representative sample of people under standard conditions for the purpose of establishing norms. (Sattler, 1982)

STANDARD SCORES: The transformation of raw scores on an equal-interval or ratio scale into a set of scores which have the same mean and standard deviation. (Salvia & Ysseldyke, 1978)

TEST-RETEST RELIABILITY: A measure of the stability of scores over time, determined by administering a test to the same group of people on two different occasions and computing the correlation between their scores. (Sattler, 1982)

VALIDITY: A judgment of how well a test measures the content it was designed to measure and its appropriateness for the targeted group. (Goodwin & Driscoll, 1980; Salvia & Ysseldyke, 1978)

References

Brigance, A. H. Brigance Inventory of Early Development. North Billerica, MA: Curriculum Associates, 1978.

Goodwin, W.L., & Driscoll, L. A. Handbook for Measurement and Evaluation in Early Childhood Education. San Francisco: Jossey-Bass, 1980.

Hardyck, C. D., & Petrinovich, L. F. Introduction to Statistics for the Behavioral Sciences. Philadelphia: W. B. Saunders Company, 1976.

Salvia, J., & Ysseldyke, J. E. Assessment in Special and Remedial Education. Boston: Houghton Mifflin Company, 1978.

Sattler, J. M. Assessment of Children's Intelligence and Special Abilities. Boston: Allyn & Bacon, Inc., 1982.

White, O. R., Edgar, E., & Haring, N. G. Uniform Performance Assessment System. Columbus: Charles E. Merrill Publishing Co., 1978.

Appendix B

A Description of Instruments by Referencing, Type, Scoring Procedure, Areas Evaluated, Age Range, Validity, and Reliability

Name of Instrument	Referencing				Type				Scoring Procedure			Areas Evaluated								Age Range														Validity	Reliability
Variable	Norm-referenced	Criterion-referenced	Curriculum-referenced		Assessment	Program Planning	Screening		Observation	Parent Interview		Gross Motor	Fine Motor	Language	Cognitive	Self-help	Behavior Management	Social-Emotional		Birth to .5 year	.5 to 1.0 year	1.0 to 1.5 years	1.5 to 2.0 years	2.0 to 2.5 years	2.5 to 3.0 years	3.0 to 3.5 years	3.5 to 4.0 years	4.0 to 5.0 years	5.0 to 6.0 years	6.0 to 7.0 years	7.0 to 8.0 years	8.0 to 9.0 years	9.0 to Adult			
Albert Einstein Scales					•				•				•		•					•	•	•	•											•	•
Assessment Programming Guide					•	•			•	•			•		•					•	•	•	•											•	•
Bayley Scales	•				•				•			•	•		•					•	•	•	•											•	•
Birth to 3 Developmental Scale		•					•		•	•		•	•		•			•		•	•	•	•												•
Boyd		•					•		•	•		•	•		•					•	•	•	•	•		•	•	•		•					
Brigance		•	•		•	•			•	•		•	•		•		•			•	•	•	•			•	•	•	•						
Bzoch-League (REEL)	•				•					•					•					•	•	•	•												•
Child Development Assessment Form					•	•			•			•	•		•			•		•	•		•	•	•										
Circle Preschool Child Assessment List		•			•	•			•			•	•		•			•		•	•	•	•	•	•	•	•								
Communicative Evaluation Chart		•					•		•	•		•	•		•					•		•	•	•	•	•	•	•	•						
Delco-Elfman (DEDAT)					•	•			•			•	•		•	•		•			•	•	•	•	•	•	•	•	•						
Denver Developmental Screening Test	•						•		•	•		•	•		•			•		•	•	•	•	•	•	•	•	•	•					•	•
Developmental Profile	•						•		•	•		•	•		•	•		•		•	•	•	•	•	•	•	•	•	•	•				•	•
Developmental Programming for Infants		•			•	•			•			•	•		•	•		•		•	•	•	•	•	•									•	•
E-LAP		•			•				•			•	•		•	•		•		•	•	•	•	•										•	•
EMI Assessment Scale	•				•	•			•			•	•		•	•		•		•	•	•	•												
Infant Evaluation Scale		•					•		•	•		•	•		•	•		•		•		•	•	•	•	•	•	•	•					•	•
KNI Developmental Scale		•					•		•	•		•	•		•		•	•		•	•	•	•	•	•	•	•	•							
Lexington (Short Form)		•					•		•	•		•	•		•	•		•		•	•	•	•	•	•	•	•	•							
MEMPHIS						•	•		•	•		•	•		•	•		•		•	•	•	•	•	•	•	•	•							
Milani-Comparetti	•						•		•	•		•	•		•	•		•		•	•	•	•	•	•	•	•	•							
Ordinal Scales of Psychological Development					•				•	•		•	•		•	•		•		•	•	•	•	•	•	•	•	•							
Portage Guide to Early Education (PGEE)					•	•			•	•		•	•		•	•		•		•	•	•	•	•	•	•	•	•							
Preschool Attainment Record (PAR)					•	•			•	•		•	•		•	•		•		•	•	•	•	•	•	•	•	•							
Preschool Language Scale (PLS)						•	•		•			•	•		•	•		•		•	•	•	•	•	•	•	•	•							
Revised Developmental Screening Inventory							•		•	•		•	•		•	•		•		•	•	•	•	•	•	•	•	•							
Uniform Performance Assessment System			•		•				•			•	•		•	•		•		•	•	•	•	•	•	•	•	•							
Vineland Social Maturity Scale	•				•				•	•		•	•		•	•		•		•	•	•	•	•	•	•	•	•							

A Handbook of Evaluation Instruments for Use With Children from Birth to Three Years

by

SARA ANN SNYDER

B. S., Kansas State University, 1980

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Family and Child Development

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1984

ABSTRACT

This handbook was designed for use by early childhood professionals to aid them in selecting an evaluation instrument. A total of 28 instruments were reviewed for their usability in programs having children within the birth to three year age range. The factors which were used to determine usability include: 1) the purpose and format of the instrument; 2) administration, scoring, and interpretation procedures; 3) validity of the instrument; and 4) reliability of the instrument. The evaluative format utilized in the handbook is based on one outlined by Goodwin and Driscoll (1980). The majority of the instruments evaluated were found to cover most developmental areas and to be administrable by the classroom teacher. However, many of the instruments have limited standardization, validity, reliability, and depth. In addition, early childhood assessment tools tend to be based on items from earlier existing instruments, resulting in the regeneration of old instruments in new forms.