

COMMONLY USED ASSESSMENT and SCREENING INSTRUMENTS

February 2004

Prepared for HIPPY USA

By
Marsha M. Black, Ph.D.
University of South Florida
Dept. of Child and Family Studies

In collaboration with
Dr. Diane Powell, Ph.D.
University of South Florida
Dept. of Child and Family Studies

Table of Contents

Introduction	1
Assessment of Young Children	2
Issues in Assessing Young Children	3
Quality of Preschool Assessment Instruments	5
Developmental Screening	5
Developmental Domains	6
Measuring School Readiness	7
Statistical and Measurement Concepts	9
Standardized Test	9
Norm-Referenced Tests	9
Criterion-Referenced Tests	10
Full Battery Assessment	10
Locally Constructed Tests	11
Psychometric Properties of Tests	11
Reliability	11
Validity	12
Test Scores	13
Raw Scores	13
Derived Scores	14
Normative Scores	14
Standard Scores	14
Stanine Scores	15
Normal-Curve Equivalent Scores	15
Mental Age Scores	16
Age-Equivalent Scores	16
Grade-Equivalent Scores	16
Percentile Rank	17
Guidelines for Using Individually Administered Tests	17
Selecting an Instrument From the Instrument Document	18
Criteria for Inclusion of Instruments	19

The Assessment and Screening Instruments Document	21
Detailed Descriptions of Instruments	23
Instrument Document	25
Cognitive	25
Boehm Test of Basic Concepts - Preschool Version	25
Bracken Basic Concept Scale (BBCS-R)	25
Brigance Diagnostic Inventory of Early Development	26
Detroit Test of Learning Aptitude - Primary Second Edition	27
Kaufman Survey of Early Academic and Language Skills (K-SEALS)	27
McCarthy Scales of Children=s Abilities (MSCA)	28
Woodcock-Johnson III	28
Communication	29
Comprehensive Receptive and Expressive Vocabulary Test - Second Edition	29
Expressive One-Word Picture Vocabulary Test - 2000 Edition	29
Illinois Test of Psycholinguistic Abilities - Third Edition	30
Oral and Written Language Scales Listening Comprehension And Oral Expression (OWLS)	30
Peabody Picture Vocabulary Test- III	31
Preschool Language Scale - Third Edition (PLS-3)	31
Receptive One-Word Picture Vocabulary Test (ROWPVT)	32
Tests of Early Language Development - Third Edition (TELD-III)	32
Social-Emotional	33
Ages & Stages Questionnaire: Social-Emotional	33
Devereux Early Childhood Assessment (DECA)	33
Social Skills Rating System	34
Vineland Adaptive Behavior Scales	34
Vineland Social-Emotional Early Childhood Scales	35
Developmental Screening	36
Ages & Stages Questionnaire	36
AGS Early Screening Profiles	36
Battelle Developmental Inventory Screening Test	37
Brigance Screens	38
Denver II	38
Developmental Indicators for the Assessment of Learning - Third Edition (DIAL-4)	39
Early Screening Inventory-Revised (ESI-R)	39

Parenting Knowledge and Skills	40
The Adult-Adolescent Parenting Inventory (AAPI-2)	40
Parent-Child Relationship Inventory.....	41
Parenting Stress Index- Third Edition.....	42
Home Environment.....	43
Home Observation for Measurement of the Environment (HOME)	43
Descriptions of Assessment Instruments.....	44
Cognitive	44
Boehm Test of Basic Concepts - Preschool Version	44
Bracken Basic Concept Scale (BBCS-R)	45
Brigance Diagnostic Inventory of Early Development	46
Detroit Test of Learning Aptitude - Primary Second Edition	47
Kaufman Survey of Early Academic and Language Skills (K-SEALS).....	48
McCarthy Scales of Children=s Abilities (MSCA)	49
Woodcock-Johnson III.....	50
Communication	51
Comprehensive Receptive and Expressive Vocabulary Test - Second Edition	51
Expressive One-Word Picture Vocabulary Test - 2000 Edition	52
Illinois Test of Psycholinguistic Abilities - Third Edition	53
Oral and Written Language Scales Listening Comprehension And Oral Expression (OWLS)	54
Peabody Picture Vocabulary Test- III	55
Preschool Language Scale - Third Edition (PLS-3)	56
Receptive One-Word Picture Vocabulary Test (ROWPVT)	57
Tests of Early Language Development - Third Edition (TELD-III)	58
Social-Emotional	59
Ages & Stages Questionnaire: Social-Emotional	59
Devereux Early Childhood Assessment (DECA).....	60
Social Skills Rating System	61
Vineland Social-Emotional Early Childhood Scales	62
Vineland Adaptive Behavior Scales	63
Developmental Screening	64
Ages & Stages Questionnaire	64
AGS Early Screening Profiles.....	65
Battelle Developmental Inventory Screening Test.....	66
Brigance Screens.....	67
Denver II	68

Developmental Indicators for the Assessment of Learning - Third Edition (DIAL-3)	69
Early Screening Inventory-Revised (ESI-R)	70
Parenting Knowledge and Skills	71
The Adult-Adolescent Parenting Inventory (AAPI-2)	71
Parent-Child Relationship Inventory.....	72
Parenting Stress Index- Third Edition.....	73
Home Environment.....	74
Home Observation for Measurement of the Environment (HOME)	74
Glossary	75
References	80

HIPPY USA=s commitment to research

Since the HIPPY program model was piloted in the early 1970s, evaluation has been an integral component of the program's development. Studies from around the world have shown how and to what extent the HIPPY program impacts children, parents and whole communities. (Westheimer, 2002)

Data collection and analysis are a crucial part of successful program implementation. This is why each local HIPPY program is required to collect client-specific and service-utilization data through the HIPPY MIS (Management Information System) and program process and implementation data through the HIPPY SAVI (Self-Assessment and Validation Instrument). Analysis of these data helps HIPPY USA (and the local program) determine the quality of program implementation and helps guide the local agency towards improvement.

HIPPY USA is committed to supporting and improving research efforts at the local, state and national levels, both to demonstrate the effectiveness of the HIPPY program and to identify areas in need of further development. The HIPPY Early Learning Goals were developed in April 2003, to assist local programs with evaluation efforts. And, the theme for program year 2003-04 is *Research: The Measure of Our Success*, with several technical assistance tools being developed under this theme. The University of South Florida (USF), acting as the research arm for HIPPY USA developed this document, **Commonly Used Assessment and Screening Instruments**, in direct response to requests from local HIPPY programs.

There are currently many pressures on local HIPPY programs (and other early childhood initiatives) to evaluate the impact their program has on children=s early learning. Many funders, implementing agencies, and other stakeholders specifically require pre- and post-testing of children who participate in HIPPY. Sometimes, they even prescribe which assessment tool to use for this testing activity. HIPPY USA decided to develop this source to help local HIPPY programs make informed decisions about testing, as well as help these programs better communicate with funders and others who are making evaluation requests.

This document does not constitute an endorsement by HIPPY USA or USF, either of the individual tools described here or of this way of testing children in general. Instead, the information provided here should be useful for

activity. HIPPY USA decided to develop this source to help local HIPPY programs make informed decisions about testing, as well as help these programs better communicate with funders and others who are making evaluation requests.

This document does not constitute an endorsement by HIPPY USA or USF, either of the individual tools described here or of this way of testing children in general. Instead, the information provided here should be useful for understanding the strengths, limitations, and appropriate uses of the instruments described. It can also be used to understand issues involved in the assessment and testing of young children in general and to educate stakeholders about these issues so that programs are not forced to use child assessment instruments in inappropriate ways.

Assessment of Young Children

3 Numerous rationales exist for assessing young children.

Support learning and instruction

Identify children for additional services

Evaluate programs and monitor trends

Assessment procedures are employed for various reasons including screening, classification/placement, and program planning. Examples include:

- 1) Individualizing the curriculum to capitalize on a child=s strengths to help address his/her special needs;
- 2) Screening to identify children needing further assessment to determine the need for health or other special services or supports;
- 3) Determining eligibility for special education and related services;
- 4) Planning an intervention program, or monitoring a child=s progress; and
- 5) Evaluating program effects.

It should be noted that most of the child assessment instruments with good psychometric properties were developed for purposes of assessing and tracking the developmental skills of children for instructional purposes, rather than for program evaluation purposes.

Issues in Assessing Young Children

<p><i>Young children's inability to read, the episodic nature of their learning, and their stress in unfamiliar settings with unfamiliar people all contribute to the special challenges facing those concerned about the assessment of young children.</i> @</p> <p>Assessing the State of State Assessments: SERVE</p>	<p>Many early childhood professionals have multiple concerns about the assessment of preschool children including what is being measured, the quality of the measurement tools, the conditions under which children are being assessed, and how the assessment results are being used. Some oppose the wide-spread use of standardized testing of preschool children because they feel the trend toward escalating assessment practices creates unrealistic academic demands for young children, their teachers, and families. It is felt that much of school readiness testing focuses primarily on measuring skills in academic domains with less emphasis on skills in the adaptive and socio-emotional domains.</p>
<p>3 The assessment process should never focus exclusively on a test score or number.</p>	<p>Tests are a standard for evaluating the extent to which children have learned the basic cognitive, academic, and social skills necessary for functioning successfully in our culture. It is important to remember that a test score represents a sample of behavior in a structured testing situation and should not be viewed as the sole determinant of a child's current competencies or future achievement. Preschool children's development is rapid and uneven, and their development is greatly impacted by environmental factors such as the care they have received and the learning environments they have experienced. Test results should always be interpreted in light of a child's cultural background, primary language, and any handicapping conditions.</p>
<p>3 Formal testing settings may not capture the full scope and depth of knowledge of what children know and can do.</p>	<p>Assessment of preschool children differs from testing older children because the standardized paper-and-pencil tests used in later grades are not appropriate for young children. Test scores and other performance measures may be adversely affected by temporary states of fatigue, anxiety, or stress. Additionally, test scores depend on a child's cooperation and motivation. Building and maintaining rapport with children is a continuous process that must be interwoven throughout the testing process. Testing young</p>

	children is challenging because an examiner must
	successfully carry out a multitude of tasks during the testing situation such as establishing rapport, administering the items according to instructions, keeping the materials ready, responding appropriately to the child, precisely recording the child=s responses, keeping the child engaged, and scoring the child=s responses. With practice, many of these procedures become routine, but even the most experienced test administrator must be thoroughly familiar with the test manual.
<p>3 The content and procedures of young children=s assessments must be different than that used with older children.</p>	<p>Just as all children do not learn in the same ways, some children do not perform well on standardized tests. Performance assessment, also known as authentic assessment, is viewed by many early childhood educators as an alternative approach to using standardized tests in assessing young children. This method of assessment involves collecting information in children=s natural settings while they are engaged in their typical daily activities as opposed to testing children in an artificial, decontextualized setting.</p> <p>In authentic assessment, a child=s performance can be evaluated as he completes a task or a product resulting from the performance, such as a portfolio. Though authentic assessments can be very useful to an early childhood educator, this approach does have some drawbacks such as the difficulty of obtaining representative samples for all of the content covered during a specific time period.</p>
<p>3 NAEYC website</p>	<p>Additional information on the established guidelines for assessing young children is available from the website of The National Association for the Education of Young Children: naeyc.org.</p>

Quality of Preschool Assessment Instruments

There are no perfect, off-the-shelf, easy-to-do assessments
...@SERVE, 2003

There is tremendous variability in the quality of assessment instruments that are available and used with young children. Early childhood educators and researchers agree that there is no perfect test. Each measure has its strengths and weaknesses. Some instruments lack the technical and psychometric qualities that assure reliable and valid test scores. Others are not appropriate for the diversity of children in a particular community, do not measure all the dimensions of children's early development and learning, and/or do not measure the full range of abilities in the domains they cover. Despite the limitations of some assessment instruments, there is a sufficient number of psychometrically sound instruments available. The challenge is selecting an instrument that will be most appropriate and useful for a specific purpose.

Developmental Screening

3 The purpose of developmental screening is to identify children who should be referred for further assessment.

Many preschool and kindergarten programs have mandated formal developmental screenings for all children they serve in order to identify children with developmental delays at an early age. Screening instruments are designed to identify children who should be referred for further assessment to determine the need for special services or supports. Because they are designed for administration to large numbers of children as the first stage in a program of assessment, they contain a limited number of items and can be administered quickly. These few items do not measure the entire range of achievement, and thus these instruments are of limited usefulness in measuring child progress over time.

3 Developmental screening instruments have limited usefulness for measuring child progress.

Each developmental screening instrument includes instructions on how to interpret raw scores. For some instruments, total raw scores in each domain are compared to pre-established cutoff points. Scores above the cutoff point mean the child is progressing as expected for his/her

developmental age. Scores below the cutoff point mean a child may need further diagnostic assessment. If raw scores are compared to cutoff points, the only information the scores provide is the number and percent of children above

and below the cutoff scores at each point in time. Other developmental screening instruments provide instructions for converting raw scores to age-equivalent scores. Converting raw scores to age-equivalent scores for two groups of children provides a comparison of the difference in average age-equivalent scores for the two groups at each point in time. This provides a more meaningful interpretation of raw scores than simply the percentage of children above and below the preestablished cutoff points.

Developmental Domains

Most developmental screening instruments have items that are clustered within five domains:

3 Personal-Social Domain

Personal-Social Domain - those abilities and characteristics that facilitate children engaging in positive and meaningful social interactions. The behaviors measured include adult interaction, expression of feelings/affect, self-concept, peer interaction, coping, and social role.

3 Adaptive Domain

Adaptive Domain - self-help skills and task-related skills. Self-help skills are those behaviors that enable the child to become increasingly more independent in daily living skills such as feeding, dressing, and personal toileting needs. Task-related skills involve the child's ability to pay attention to specific stimuli for increasingly longer periods of time, to assume personal responsibility for his or her actions, and to initiate purposeful activity and follow through appropriately to completion. Behaviors measured include attention, eating, dressing, personal responsibility, and toileting.

3 Motor Domain

Motor Domain - gross motor development (large muscle movement and control) and fine motor development (hand and finger skills; and hand-eye coordination). Behaviors measured include muscle control, body coordination,

<p>3 Communication Domain</p>	<p>locomotion, fine muscle, and perceptual motor.</p> <p>Communication Domain - Understanding and using language to communicate for various purposes. Behaviors measured include a child=s reception and expression of information, thoughts, and ideas through verbal and nonverbal means.</p>
<p>3 Cognitive Domain</p>	<p>Cognitive Domain - skills and abilities that are conceptual in nature. Abilities measured include perceptual discrimination, memory, and reasoning. Tasks include comparison among objects based on physical features (color, shape, size) and properties (weight); sequencing events; putting together parts of a whole; grouping and sorting similar objects and identifying similarities and differences among objects based on common characteristics.</p>
<p>Measuring School Readiness</p>	
<p>3 By the year 2000, all children in America will start school ready to learn.@ National Education Goals Panel</p>	<p>The National Education Goals Panel identified three components of school readiness: 1) readiness in the child; 2) schools= readiness for children; and 3) family and community supports and services that contribute to children=s readiness.</p> <p>There are five dimensions of early development and learning considered critical to a child=s readiness for school (the National Education Goals Panel: Kagan, Moore, & Bredekamp, 1995):</p> <ul style="list-style-type: none"> § Health and physical development § Emotional well-being and social competence § Approaches to learning § Language development § Cognition and general knowledge <p>In addition, the NEGP identified three supporting conditions that contribute to preparing a child to enter school: having access to quality preschool programs, parents as children=s first teachers, and appropriate nutrition and health care.</p>

<p>3 Mandated formal school readiness assessments have become common practice.</p>	<p>Defining and measuring school readiness is important because many states have enacted legislation mandating the measurement of school readiness. Mandated formal assessments for every child entering kindergarten have become the norm in many local school districts. In addition, the determination of program effectiveness for many preschool programs is increasingly being linked to the degree of children=s preparedness to enter school.</p>
<p>3 There is a lack of consensus among early childhood educators and policymakers regarding the specific definition and measurement of school readiness.</p>	<p>Underscoring the general problem of quality in early childhood assessment is the fact that though a number of assessment instruments exist that claim to measure school readiness and are in substantial use by schools around the country, there is a lack of consensus among early childhood educators and policymakers regarding the specific definition and measurement of school readiness. Some of the instruments marketed as school readiness instruments have limited item selection and subscale development, poor reliability and validity (especially predictive validity), and a lack of rationale for determining cutoff scores. This is particularly true of the self-developed tests that have been created by some states and school districts. Assessment instruments that produce totally inconsistent results over different periods of time, over different samples of questions, or over different raters cannot provide valid information about the performance being measured.</p>
<p>3 AThe principal difference between readiness tests and achievement tests is temporal..@ Samuel J. Meisels</p>	<p>Children=s readiness has typically been assessed by tests that are variations of achievement tests (Meisels,1998). Tests of critical skill mastery monitor progress in the development of critical skills that have been identified as necessary for school readiness such as receptive language skills, expressive language skills, and reading comprehension. Readiness tests are administered at the outset of the school year and achievement tests are usually given at the end, however, the content of both tests is essentially the same: a measure of skill achievement (Meisels, 1998). School readiness has also been assessed using developmental screening instruments that assess children in a variety of domains such as speech, language, gross and fine motor skills to provide a global index of a</p>

child=s developmental status. However, the limitations of these instruments as described in the previous section make their use for this purpose questionable.

Statistical and Measurement Concepts

3 Standardized tests

A **standardized test** is a test for which procedures have been developed to ensure consistency in administration and scoring across all testing situations.

The strengths of standardized instruments outweigh their weaknesses.

Standardized tests have several advantages: 1) the items are well written and have been tested for clarity; 2) standard conditions of administration and scoring have been established; and 3) tables of norms are provided. These advantages are offset by some drawbacks. Guessing, response sets (giving the same type of answer to all of items), or random or careless answers can distort the scores. As most standardized tests have a restricted time limit, they may not accurately reflect the characteristics of children who are much slower, more deliberate, or more thoughtful in responding than their peers. Additionally, scores on standardized tests do not reflect the unique experiences of different types of individuals. However, overall, the strengths of standardized instruments outweigh their weaknesses.

3 Norm-referenced tests

Norm-referenced tests are designed to provide a measure of performance that is interpretable in terms of an individual=s relative standing in some known group.

A child=s performance is compared with the performance of a specific group of children.

Norms are useful because knowing a child=s raw score does not provide information on how other children performed on the same test. In norm-referenced testing, a child=s performance is compared with the performance of a specific group of children. A norm provides an indication of average or typical performance of a specified group. Norm-referenced tests also provide valuable information about a child=s level of functioning in the areas covered by the tests - information that would be unavailable to even the most skilled observer who did not use tests. They also provide an index for evaluating change in many different aspects of the child such as developmental growth and the effects of an intervention. However, a limitation of norm-referenced tests

	is that they provide only limited information about the ways children learn.
3 Criterion-referenced tests	Criterion-referenced tests are designed to provide a measure of performance that is interpretable in terms of a clearly defined and delimited domain of learning tasks.
<input type="checkbox"/> Criterion-referenced measures are used to describe a child=s performance with respect to an established standard.	<p>While norm-referenced testing is used to evaluate a child=s performance in relation to the performance of other children on the same measure, criterion-referenced measures are used to describe a child=s performance with respect to an established standard; it measures level of mastery.</p> <p>There are two subtypes of criterion-referenced measurement: domain-referenced measurement involves selection of a random sample of items drawn from an item pool that is representative of all possible test items for a well-defined content area. The second is objective-referenced measurement sometimes referred to as curriculum-based measurement that provides information on how well the student performs on items measuring attainment of specific instructional objectives.</p>
3 Full battery assessment <input type="checkbox"/> Most full battery assessments include subtests that measure basic skills in reading, language, and mathematics.	<p>Full Battery Assessment is a series of individualized tests all standardized on the same representative group of students.</p> <p>Full battery assessments provide for a comparison of test scores on the separate tests and thus determine a child=s relative strengths and weaknesses in the different areas covered by the tests. The basic skills covered in a full battery assessment are measured by a number of subtests. Although the names of the subtests may vary somewhat from one test publisher to another and the batteries at the primary level usually contain fewer subtests, there is considerable uniformity in the outcomes measured by the various basic skill batteries. Most full battery assessments include subtests that measure basic skills in reading, language, and mathematics. One limitation of full battery assessments is that all parts of the battery are usually not equally appropriate for measuring all instructional objectives. Batteries are constructed based on the objectives and content considered important by the</p>

	specialists developing the test.
3 Locally constructed tests	A locally constructed test is developed by teachers for use with their own students or by program coordinators who want a measure of program effectiveness such as a parent satisfaction measure.
<input type="checkbox"/> Locally constructed tests may lack adequate psychometric properties.	Locally constructed tests can aid in determining 1) the appropriateness and attainability of the instructional objectives; 2) the usefulness of the instructional materials, and 3) the effectiveness of the instructional methods. In some cases, states or local school districts may construct tests for specific purposes such as measuring the extent that students are ready for the next learning experience. However, if constructed by individuals without sufficient knowledge of the principles of test construction, these instruments lack adequate psychometric properties which are explained in the next section.

Psychometric Properties of Tests

3 Reliability refers to the consistency of measurement.	Test results need to be dependable. High reliabilities are especially needed for tests used for individual assessment. Reliability provides the consistency that makes validity possible and indicates how much confidence can be placed in the evaluation results. Suppose, for instance, that an achievement test is administered to students in a classroom during the 10 th week of instruction. How similar would the students' scores have been if they had been tested during the 9 th or 11 th week of instruction. Unless the scores can be reasonably consistent (that is, generalizable) over several occasions, little confidence can be placed in the results. For most tests of cognitive and special abilities, a reliability coefficient of .80 or higher is generally considered to be acceptable. There are many factors that affect the reliability of a test including:
<input type="checkbox"/> Ways to improve reliability	<p>§ Test length</p> <p>The more items there are on a test and the more homogeneous they are, the greater the reliability is likely to be;</p>

	<p>§ Test-retest interval The smaller the time interval between administration of two tests, the smaller the chance of change;</p> <p>§ Guessing The less guessing that occurs on a test or the less examinees respond to items randomly, the higher the reliability is likely to be;</p> <p>§ Variation within the test situation Factors such as misleading or misunderstood instructions and scoring errors introduce error into the testing procedures.</p>
3 Validity refers to the extent to which a test measures what it is supposed to measure.	Validity is always specific to some particular use or interpretation of test scores. Unless a test is valid for the purpose for which it is being used, the results cannot be used with any degree of confidence. It is important to remember that no test is valid in general; tests are valid only for a specific purpose. Validity is not a matter of all or nothing; it is a matter of degree. It is the examiner who is responsible for the valid use of test results.
<input type="checkbox"/> Ways to improve validity	Interpretations of test scores are likely to have greater validity when the test examiner understands 1) the test content and the specifications it was derived from; 2) the relation of the test scores to other significant measures, and 3) the nature of the characteristic(s) being measured.
<input type="checkbox"/> Poorly constructed test items lower the validity of scores.	<p>A careful examination of the test items will indicate whether the test appears to measure the subject-matter content that is of interest. However, any of the following factors in the test itself can prevent the test items from measuring what they were supposed to measure, and thereby lowering the validity of the interpretations that can be made from the test scores:</p> <p>§ Unclear directions</p> <p>§ Reading vocabulary and sentence structure too difficult</p>
	§ Inappropriate level of difficulty of the test

	<p>items</p> <ul style="list-style-type: none"> § Poorly constructed test items (test items that unintentionally provide clues to the answer may be measuring the student=s alertness in detecting clues) § Ambiguous statements in test items § Inadequate time limits § Test too short § Improper arrangement of items within the test (placing more difficult items early in the test)
<p>3 Test Scores</p>	<p>Each instrument listed in the Commonly Used Assessment and Screening Instruments document includes a listing of the type of scores reported in the instrument=s test manual. A description of the different types of scores is included below.</p>
<p>3 Raw Scores</p>	<p>A raw score is simply the number of points received on a test when the test has been scored according to the directions.</p>
<p>□ A raw score provides limited information about a child=s performance.</p>	<p>It does not make any difference whether each item is counted as one point, whether the items are weighted in some way, or whether a correction for guessing is applied. The resulting score is still a raw score. Although the ranking of raw scores may be useful for reporting the results of a test, ranking scores is of limited value beyond the immediate situation because the meaning of a given rank depends on the number of children taking the test.</p> <p>Although a raw score is a numerical summary of a student=s test performance, it is not very meaningful without some further information. Meaning is provided to a raw score by either converting it into a description of the specific tasks that a child can do (i.e., identifies 80% of the alphabet) or converting it into some type of derived score that indicates the child=s relative position in a clearly defined reference group (child=s score was above the average obtained by the 5th graders in the norming sample).</p>
	<p>A derived score is a transformation of a raw score (e.g.</p>

3 Derived Scores	age-equivalents) to reveal the individual=s performance relative to the norming group.
<p>Derived scores aid in the interpretation of raw scores by providing a quantitative measure of each child=s performance relative to a comparison group. For example, if a child answered 35 items correctly on a math test, and therefore, has a raw score of 35, what does this mean? Is this a good score? Is this the average score we would expect of a child the same age? Answers to these questions are necessary to interpret a score of 35 and provide meaningful information.</p>	
3 Normative Scores	A normative score is a derived score that can be used to compare a child=s performance on a test to a representative population (children of the same age, socio-economic status, ethnicity, etc.), to children in the child=s school, or with his/her scores on other tests.
<input type="checkbox"/> Compare a child=s performance with that of other children.	<p>Normative scores are calculated by converting the child=s raw score into some relative measure. Most test publishers of standardized tests have performed the conversions and reported these in the table of norms in the test manual. There are various types of derived scores and the different types of scores are merely different expressions of a child=s performance. The most common types of derived scores are standard scores, age equivalents, grade equivalents, and percentile ranks.</p>
3 Standard Scores	A standard score is a raw score that has been transformed to have a given mean and standard deviation (typically, mean=100; standard deviation=15).
<input type="checkbox"/> Mean <input type="checkbox"/> Standard Deviation	<p>Standard scores indicate a child=s relative position in a group by showing how far the raw score is above or below the average score on that test. Standard scores express test performance in terms of standard deviation units from the mean. The mean (M) is the arithmetical average, which is determined by adding all of the scores together and dividing by the number of scores. The standard deviation (SD) is</p>

	<p>the measure of the spread of scores in a group. Commonly used standard scores on many standardized tests are AZ-scores[®] and AT-scores[®].</p>
3 Stanine Scores	<p>Stanine scores are single-digit standard scores using a nine-point scale in which 9 is high, 1 is low, and 5 is average. Stanines provide test norms on a nine-point scale of equal units.</p>
<p>□ Compare a child=s performance with that of other children.</p>	<p>Stanines are computed like percentile ranks but are expressed in standard score form. They have a mean of 5 and a standard deviation of 2. Stanines are widely used for local norms because of the ease with which they can be computed and interpreted. Parents can readily visualize where test performance falls (e.g., 7) on a scale of 1 to 9. The primary limitation of stanine scores is that growth cannot be shown from one year to the next. If a child=s performance matches that of the norm group in one year, the same position in the group will be retained and the same stanine assigned in the second year if his performance matches the norm. This shortcoming is also true of other standard scores that are used to indicate relative position in a group.</p>
3 Normal-Curve Equivalent Scores	<p>Normal-Curve Equivalent Scores are normalized standard scores with a mean of 50 and a standard deviation of 21.06 so that NCE scores range from 1 to 99.</p> <p>Because Federally supported compensatory education programs often require normal-curve equivalent scores, publishers of major standardized achievement tests have added this type of score to their list of reporting options.</p>
3 Mental Age Scores	<p>Mental Age Scores indicate the degree of general mental ability of an average child of a specific chronological age.</p> <p>Mental age can be thought of as a developmental measure that indicates a level of cognitive functioning. It can also be conceived of as a level of achievement that may indicate a child=s readiness to learn. Mental age (MA) divided by chronological age (CA) and multiplied by 100 produces the ratio IQ (Intelligent Quotient), which is a reported score on some cognitive batteries. The MA score provides an age-equivalent for the child=s raw score, whereas the IQ</p>

	indicates a child=s performance relative to that of children who are at his/her own chronological age.
3 Age-Equivalent Scores	An age-equivalent score is a type of derived score that indicates the age at which a specific raw score is average.
3 Grade-Equivalent Scores	A grade-equivalent score is a type of derived score that identifies the grade level at which the typical child obtains that raw score.
	Age-equivalent and grade-equivalent scores are derived by determining the average score obtained on a test by children of various ages or grade placements. Age equivalents may be expressed in months or years. For example, if the average score of 10-year old children on a test is 15 items out of 25, then any child scoring 15 is said to have an age-equivalent score of 10. Grade equivalents are expressed by two numbers; the first indicates the year and the second, the month. Grade equivalents for the fifth grade, for example, range from 5.0 to 5.9 with September=0 and June=9.
<input type="checkbox"/> Grade equivalents may be misinterpreted.	Though age-equivalent and grade-equivalent scores are easy to understand and interpret, they are also subject to misinterpretation based on unjustified assumptions about the meaning of the scores. These inappropriate assumptions are 1) assuming that norms are standards of what should be; 2) assuming that grade-equivalents indicate the appropriate grade placement for a child; 3) assuming that all pupils
	should be expected to grow one grade equivalent unit per year; 4) assuming that all units are equal throughout the score range, 5) assuming that grade equivalents for different tests are comparable; and 6) assuming that the scores that are well above or below the test level represent the actual performance of children at these levels.
3 Percentile Ranks	A percentile rank is a point in the distribution at or below which the scores of a given percentage of children fall. Percentile ranks are derived scores that indicate a child=s position relative to the standardization sample. If 63% of the scores fall below a given score, then that score is at the 63 rd percentile meaning that 63% of the children in the

sample had lower scores.

Guidelines for Using Individually Administered Tests

3 Follow standardized procedures.

Tests and other assessments are powerful tools, but their usefulness depends on the test administrator's skills in testing young children, familiarity with the test manual, and ability to follow precisely the instructions for administering the test. For example, it is important to use the exact words of the question, the specific test materials, and to adhere to the specified time limits.

3 Test manuals

Well-designed manuals present a clear rationale and overview of the test (how it was developed, what it is intended to measure, and the limitations of the test) along with eligibility criteria, examiner qualifications, testing time, scoring guidelines, and score interpretation (conversion of raw scores to derived scores such as percentile ranks, age equivalents, grade equivalents, and standard scores). Technical information such as norming procedures and psychometrics (reliability and validity indices) is also included in the test manual

3 Assessments should be non-intrusive, non-threatening and take as little of a child's time as possible.

- § **Be attentive to a child's needs and concerns.**
Know your task well enough so that the test flows almost automatically, leaving you maximally free to observe all aspects of the child's behavior.
- § **Conduct several trial test administrations of an instrument.**
This is one way a test examiner becomes knowledgeable about the instrument and comfortable in the testing situation. Many publishers offer technical support to programs and also have training videos available for purchase.
- § **Establish and maintain rapport with the parents** and enlist their support in the testing situation.
- § **Develop a plan for sharing information with parents** and acting on this information should test results suggest the need for a follow-up assessment.

<p>3 Disclosing personal and/or sensitive information.</p>	<p>As is true of all testing procedures, it is important to remember that when administering parenting measures, some parents may be concerned about disclosure of personal and sensitive information on surveys or during an interview. Parents must be comfortable that their information will remain confidential. The evaluation process should be conducted legally, ethically, and with due regard for the welfare and privacy of those involved in the evaluation, as well as those affected by its results.</p>
<p>Selecting an Instrument From the Instrument Document</p>	
<p><input type="checkbox"/> Several individuals may have the responsibility of selecting an instrument.</p>	<p>For some programs, the implementing agency and/or funder may require child outcomes be measured by a particular assessment instrument. In other programs, the program coordinator may have the responsibility of administering, interpreting, and reporting results from an instrument she/he selects. Or, the coordinator may select an instrument and contract with another professional to administer, score, and interpret the results.</p>
<p><input type="checkbox"/> The important question is, To what extent will the interpretation of the scores be appropriate, meaningful, and useful?</p>	<p>When using this guide to select an assessment or screening instrument, compare and contrast the information provided on each instrument. Though sharing many commonalities, the instruments may differ from each other in several respects such as administration time, ease of administration, ease of score interpretation, and amount of training required to reliably administer test.</p>
<p><input type="checkbox"/> Tests differ with regard to the level of examiner skills necessary to administer the instrument.</p>	<p>Some instruments require the test examiner to establish basal and ceiling points, a process that could prove difficult for a novice test administrator. Other tests require test examiners to be very familiar with the child being tested, which would require observing the child on different occasions and in different situations (with parents at home, playing with other children, speaking with adults).</p>
<p>The more psychometrically sound instruments require a considerable amount of time and clinical expertise on the part of the examiner, which adds to the complexities of selecting appropriate instruments for program evaluation.</p>	
<p>Selecting an appropriate instrument usually involves trade</p>	

	<p>offs between the resources available and the ideal instrument. Factors to consider include:</p> <ul style="list-style-type: none"> § Comprehensiveness of the test. § Fit with purpose of the test. § Cost of the instrument. § Qualifications and training required for properly administering the test. § Time needed to administer. § Ease of scoring and interpreting scores.
<p>Criteria for Inclusion of Instruments</p>	
	<p>Every effort was made to identify a range of commercially available instruments, but this document is not an exhaustive list of all available instruments. Other instruments that might exist were not intentionally excluded. Likewise, inclusion of an instrument does not imply endorsement by the University of South Florida HIPPY Research Collaborative or HIPPY USA.</p>
<p>3 Sources of information about the instruments that are included in this document.</p>	<p>The selection of instruments for the review was drawn from several sources including a literature review of research in emergent child and family literacy; national program evaluations (Head Start, Family and Child Experiences Survey); Project SERVE=s Compendium of Assessment Instruments; Preschool Evaluation and Assessment Project: Preschool Test Compendium, Florida Department of Education; and information gathered from surveys distributed at HIPPY workshops.</p>
	<p>Information compiled about the instruments was obtained from publisher=s websites and the Buros Mental Measurements Yearbook, a resource that critically evaluates commercially available testing instruments. Evaluative comments on the narrative pages were drawn from comments in the test reviews published in the Buros Mental Measurements Yearbook.</p>
<p>3 Parenting measures</p>	<p>The majority of the published parenting measures that were reviewed are designed for use with families undergoing court-ordered custody mediation, child abuse allegations, or those experiencing a high degree of personal or family</p>

	stress such as the lack of finances, divorce, custody dispute, and potential domestic violence.
	The purpose of many parenting instruments is to provide information on parenting practices and beliefs that would be useful in a clinical setting such as in professional counseling or for a court-ordered evaluation. These instruments would not be appropriate for use with HIPPY parents.
<input type="checkbox"/> Additional criteria for selecting parenting measures.	<p>In addition to the criteria described below used for the review of all instruments, there were some additional considerations for the inclusion of parenting measures in the Annotated Bibliography of Commonly Used Assessment Instruments.</p> <ul style="list-style-type: none"> § Sensitivity of Item Content § Test items asking a parent to disclose information about the quality of their relationship with their spouse and children, family history, and current discipline practices and beliefs can be viewed as intrusive. § Inclusion of a built-in scale for assessing social desirability responses sets and response inconsistency.
3 Standards for selecting an instrument	<p>Instruments listed in the Annotated Bibliography of Assessments and Screening Instruments met the following criteria:</p> <ul style="list-style-type: none"> § Measures child outcomes identified in the HIPPY Early Learning Goals developed by HIPPY USA. § Standardized, norm-referenced instruments. § Appropriate for the ages of HIPPY children.
	<ul style="list-style-type: none"> § Sound psychometric properties (reliability and validity indices). § Standardization sample of adequate size and representative of children from low socio-economic status. § Easily interpretable instructions for raw score conversions. § Manual is well-organized, written clearly, and contains the information necessary for administering the test and interpreting the results.

	<p>§ Instruments are in print, available and affordable.</p> <p>§ An instrument does not duplicate information measured by another more well-known and widely-used instrument.</p>
<p>The Assessment and Screening Instruments Document</p>	
<p>3 This document provides information about assessment instruments available for use with young children.</p>	<p>The instrument chart provides brief overview information of child and parent outcome measures that could be used to evaluate child progress and change and screening instruments that measure a child=s developmental status. Instruments are arranged alphabetically within categories (<i>numbers in parenthesis are the number of instruments in the category</i>) and offers the reader a quick glance at all instruments included in the document.</p>
	<p>The following categories included are:</p> <p>§ Cognitive (multiple domains) (7)</p> <p>§ Communication (language and literacy) (8)</p> <p>§ Social-Emotional (5)</p> <p>§ Developmental Screeners (9)</p> <p>§ Parenting Knowledge and Skills (3)</p> <p>§ Home Environment (1)</p>
<p>3 Information provided in the instrument chart</p> <p><input type="checkbox"/> Instrument name</p>	<p>The following information is provided for each instrument:</p> <p>This is the official name of the instrument and its authors. The publisher, publication date, and publisher=s website are included for the purpose of acquiring information regarding test purchase and test purchaser qualifications. Some publishers require test purchasers meet certain qualifications to purchase the test to demonstrate they are qualified to use the test.</p>
<p><input type="checkbox"/> Purpose</p>	<p>The purpose refers to the goals or objectives identified by the test developers in the test manual.</p>
<p><input type="checkbox"/> Description</p>	<p>The test description includes the type of instrument (e.g., full battery assessment, criterion-referenced instrument, parent report questionnaire, developmental screening instrument), ages (the age range for which the instrument was developed) and domains measured.</p>

<input type="checkbox"/> Administration	<p>Information in this column indicates the perceived ease with which the instrument can be administered and includes method of administration (whether it is group or individually administered), time (average time it takes a child to complete all of the items), and examiner=s qualifications as identified by the test reviewer and/or as stated on the publisher=s website. Administration time is calculated based on the assumption that the test will be given by an individual who is experienced in individual test administration and familiar with the instructions for administration and scoring provided in the test manual.</p>
<input type="checkbox"/> Scores	<p>Types of scores reported in the test manual such as standard scores, percentiles, age-equivalent scores, etc.</p>
<input type="checkbox"/> Use and limitations	<p>Use refers to the specific ways in which the instrument can be used for evaluation purposes such as measuring child progress and change over time. Spanish version indicates the availability of a Spanish translated version available for purchase from the publisher. Limitations refer to concerns that should be considered when utilizing the instrument to measure HIPPIY goals.</p>

Detailed Descriptions of Instruments

3 Information provided in the detailed descriptions of the instruments.	<p>Each of the tests listed in the instrument chart is profiled in greater detail. Instruments are listed alphabetically by domain and the following information is provided for each:</p>
<input type="checkbox"/> Publisher=s price	<p>The most up-to-date information from the publisher=s website is provided. However, the current publisher=s price for the instrument is subject to change. All of the publishers have toll free numbers and welcome questions regarding the use of their instruments.</p>
<input type="checkbox"/> Description of the Instrument	<p>Test administration procedures (what the examiner and examinee actually do during the test) and a description of materials included in the test kit provides the reader more in-depth information on using the instrument. For example, is the child expected to look at and name objects, draw a figure, provide definitions to words, or construct a figure?</p>

	Does the test examiner work from an easel or a test booklet?
<input type="checkbox"/> Strengths	<p>The advantages of using the instrument for the purposes of HIPPY evaluation activities are presented. For example, ease of administration (does instrument require a high level of skill to be able to follow precise instructions), difficulty level of scoring and score interpretation (are raw score conversion tables provided; is there a computer scoring program available), and length of time it takes to administer the instrument. If a HIPPY program has used the instrument, this is also listed as a strength because program coordinators who have experience with the instrument would serve as resource for information about the test. If the instrument has been used in national program evaluations, this is also indicated as a strength because programs may be able to compare their results with national data.</p>
<input type="checkbox"/> Limitations	<p>These are considerations to keep in mind when using the instrument such as length of time to administer the test and test directions that requiring subjective judgment on the part of the examiner. Also, raw score conversions that require using several different tables in the test manual=s appendix may involve substantive time.</p>
Glossary	
<p>The glossary of measurement terms provides definitions of commonly used statistical and measurement terms that are included in the chart and the detailed descriptions of the instruments.</p>	
References	
<p>An alphabetical list of resources utilized in the development of the instrument chart.</p>	

Commonly Used ASSESSMENT and SCREENING INSTRUMENTS

January 2004

Cognitive					
Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>Boehm Test of Basic Concepts -III Preschool Version</p> <p>Psychological Corporation - 1984-86</p> <p>Test purchase restricted to qualified professionals. Publisher=s website: http://www.harcourt.com/index.html</p>	<p>Measure a child=s knowledge of various relational concepts (such as direction, amount, and time) that are thought necessary for achievement in the first few years of school.</p>	<p>Criterion-referenced Assessment</p> <p>Ages: 3.0 to 5-11 years</p> <p>Domains: 26 relational concepts including: Space Quantity Size Direction Time Different Matching</p>	<p>Individual administration</p> <p>Time: 20-30 minutes</p> <p>Examiner Qualifications: Master=s degree in psychology or education or training in assessment</p>	<p>Percentiles by 6-month age bands</p>	<p>Use: An indicator of school readiness. Can be used as a pre/post measure.</p> <p>Spanish Version: Yes</p> <p>Limitations: Does not measure the expressive use of concepts.</p>
<p>Bracken Basic Concept Scale (BBCS-R)</p> <p>Psychological Corporation - 1998</p> <p>Test purchase restricted to qualified professionals. Publisher=s website: http://www.harcourt.com/index.html</p>	<p>Measure basic concept acquisition and receptive language skills in children.</p>	<p>Full Battery Assessment</p> <p>Ages: 2 2 to 8 years</p> <p>Domains: Color Letters Numbers Comparisons Shapes Direction, Social/emotional connotations Size Texture Quantity Time</p>	<p>Individual administration</p> <p>Time: 30 minutes.</p> <p>Examiner Qualifications: Master=s degree in psychology or education or training in assessment</p>	<p>Standard scores Age-equivalent scores Percentile ranks</p>	<p>Use: Screening instrument, a measure of basic concept development, and a school readiness assessment.</p> <p>Spanish version: Yes</p> <p>Limitations: Can be administered by paraprofessionals.</p>

Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>The Brigance Diagnostic Inventory of Early Development</p> <p>Curriculum Associates, Inc. 1978-1991</p> <p>Test purchase restricted to qualified professionals. Publisher website http://www.curricassoc.com</p>	<p>Assessing and tracking the developmental skills of children.</p>	<p>Cognitive Screener</p> <p>Ages: Birth to 7 years</p> <p>Domains: Preambulatory Motor Skills and Behaviors Gross-Motor Skills and Behaviors Fine-Motor Skills and Behaviors Self-Help Skills Speech and Language Skills General Knowledge and Comprehension Social and Emotional Development Readiness Basic Reading Skills Manuscript Writing Basic Math.</p>	<p>Individual administration, parent interviews, and observation</p> <p>Time: Not provided by publisher.</p> <p>Special considerations: Due to the extensiveness of the inventory, it is not possible to administer the entire test at one time. Examiners are encouraged to use their judgment in determining which skills to evaluate in a single sitting.</p> <p>Examiner Qualifications: Master=s degree in psychology or education or extensive training in assessment.</p>	<p>Each basic skill is regarded as distinct, and there is no cumulative score associated with a skill area or subarea. Responses are recorded in the Developmental Record Book, an intact booklet consisting of an ordered set of all basic skills and behaviors. Recommended coding for each skill is (a) not assessed, (b) assessed and set as an objective, (c) introduced by not achieved, and (d) skill has been achieved.</p>	<p>Use: Assessing and tracking developmental skills.</p> <p>Spanish Version: No</p> <p>Limitations: No significant limitations noted in test reviews.</p>

Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>Detroit Test of Learning Aptitude - Primary Second Edition</p> <p>PRO-ED, Inc. - 1986-1991</p> <p>No test purchase restrictions. Publisher=s website: http://www.proedinc.com/</p>	<p>Identify a child=s strengths and weaknesses in a variety of developmental domains such as language, attention, and motor abilities</p>	<p>Screening Instrument</p> <p>Ages: 3-0 to 9-11 years</p> <p>Domains: Linguistic Attention Motoric</p>	<p>Individual administration</p> <p>Time: 15 to 40 minutes</p> <p>Examiner=s Qualifications: Examiners should be thoroughly familiar with the administration directions in the manual. No special assessment training or experience is required.</p>	<p>Standard scores Percentile ranks Age-equivalent scores</p>	<p>Use: Develop a profile of child=s strengths and weaknesses or serve as a screener to identify children who may need further assessment.</p> <p>Spanish version: No</p> <p>Limitations: Scores should not be used in making predictions about future academic performance (aptitude).</p>
<p>Kaufman Survey of Early Academic and Language Skills (K-SEALS)</p> <p>American Guidance Service - 1993</p> <p>Test purchase restricted to qualified professionals. Publisher=s website: http://www.agsnet.com/</p>	<p>Measure children=s language skills (expressive and receptive vocabulary), pre-academic skills, and articulation.</p>	<p>Full Battery Assessment</p> <p>Ages: 3-0 to 6-11 years</p> <p>Domains: Expressive Skills Receptive Skills Number Skills Letter and Word Skills</p>	<p>Individual administration</p> <p>Time: 15-25 minutes</p> <p>Examiner=s Qualifications: Master=s degree in psychology or education or training in assessment</p>	<p>Standard scores Percentile ranks Age-equivalent scores</p>	<p>Use: A screening instrument or as part of a comprehensive language battery</p> <p>Spanish Version: No</p> <p>Limitations: Overall composite score is the most useful score. Use other measures for expressive and receptive skills.</p>

Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>McCarthy Scales of Children=s Abilities (MSCA)</p> <p>The Psychological Corporation - 1970-78</p> <p>Test purchase restricted to qualified professionals. Publisher=s website: http://www.harcourt.com/index.html</p>	<p>Measure cognitive and motor development in children.</p>	<p>Full Battery Assessment</p> <p>Ages: 2-4 to 8-7 years</p> <p>Domains: Verbal Perceptual-Performance Quantitative General Cognitive Memory Motor</p>	<p>Individual administration</p> <p>Time: 45 minutes for children under 5; 1 hour for older children</p> <p>Examiner Qualifications: Master=s degree in psychology or education or training in assessment. .</p>	<p>Standard scores Mental Age scores</p> <p>An overall General Cognitive Score which indicates mental competence but is not synonymous with an IQ score.</p>	<p>Use: Assess child skill level of cognitive ability and, to a lesser extent, the motor abilities of young children.</p> <p>Spanish Version: No</p> <p>Limitations: Does not measure social-emotional functioning such as maturity and judgment.</p>
<p>Woodcock-Johnson III</p> <p>Riverside Publishing 1977-2001</p> <p>Test purchase restricted to qualified professionals. Publisher=s website: http://www.riverpub.com</p>	<p>Measure general and specific cognitive abilities, scholastic aptitude, oral language, and academic achievement in two distinct, co-normed batteries: The Tests of Cognitive Abilities and the Tests of Achievement.</p>	<p>Full Battery Assessment</p> <p>Ages: 2 to 90 years.</p> <p>Domains- Cognitive Ability: Picture Vocabulary Spatial Relations Memory for sentences Visual-auditory Learning Blending Quantitative Concepts</p> <p>Domains-Achievement Letter-Word Identification Passage Comprehension Applied Problems Story Recall Story Recall Delayed Understanding Directions Spelling</p>	<p>Individual administration</p> <p>Time: Approximately five minutes per test; 35-115 minutes for each full battery:</p> <p>Examiner Qualifications: Master=s degree in psychology or education or training in assessment.</p>	<p>Grade-equivalent scores Age-equivalent scores Percentile ranks</p>	<p>Use: Pre/post measure or a measure of skill acquisition in the Cognitive Abilities Domain.</p> <p>Spanish Version: Yes</p> <p>Limitations: Test items must all be answered (no skipping) in order for scores to be valid for preschool children.</p>

COMMUNICATION					
Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>Comprehensive Receptive and Expressive Vocabulary Test - Second Edition</p> <p>PRO-ED, Inc. 1994-2002</p> <p>No test purchase restrictions.</p> <p>Publisher=s website: http://www.proedinc.com</p>	<p>Identify deficiencies in oral vocabulary, identify discrepancies between receptive and expressive vocabulary; document progress in instructional programs and serve as a research tool.</p>	<p>Test of Oral Language Skills</p> <p>Ages: Receptive test: 4 - 89 years; Expressive test, age 5 - 89 years.</p> <p>Domains: Receptive Vocabulary Expressive Vocabulary</p>	<p>Individual administration</p> <p>Time: 20 to 30 minutes.</p> <p>Examiner Qualifications: Must be reasonably competent in the administration of tests</p>	<p>Standard scores Percentile ranks Age-equivalent scores Percentile ranks and Age-equivalent scores are provided for each subtest and standard scores from the two subtests are combined and converted into a General Vocabulary Composite score.</p>	<p>Use: Pre/Post test measure.</p> <p>Spanish Version: No</p> <p>Limitations: May not adequately distinguish children with language abilities that are either above or below average (gifted or language impaired children).</p>
<p>Expressive One-Word Picture Vocabulary Test - 2000 Edition</p> <p>Academic Therapy Publications, 1990</p> <p>Test purchase restricted to qualified professionals.</p> <p>Publisher=s website: http://www.academictherapy.com</p>	<p>Assess a child=s English speaking vocabulary. The Spanish/Bilingual edition assesses the total acquired vocabulary of individuals who are bilingual in Spanish and English.</p>	<p>Test of Expressive Vocabulary</p> <p>Ages: 2 through 18 years.</p> <p>Domains: Expressive Vocabulary</p>	<p>Individual administration</p> <p>Time: 10 - 15 minutes</p> <p>Examiner Qualifications: Master=s degree in psychology or education or training in assessment.</p>	<p>Standard scores Percentile ranks Age-equivalent scores</p>	<p>Use: Pre/Post test</p> <p>Spanish Version: Yes</p> <p>Limitations: Scope of language assessment is rather narrow. It is recommended the test be used with the Receptive One-Word Picture Vocabulary Test for a comparative analysis of the child=s speaking and writing abilities.</p>

Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>Illinois Test of Psycholinguistic Abilities - Third Edition</p> <p>PRO-ED, Inc. 1961-2001</p> <p>No test purchase restrictions. Publisher=s website: http://www.proedinc.com/</p>	<p>Assess a child=s verbal and nonverbal psycholinguistic ability, i.e., to understand, organize, and express information in both an auditory-vocal and a visual-motor channel.</p>	<p>Full Battery Assessment</p> <p>Ages: 5 through 12 years.</p> <p>Domains: Auditory and Visual Reception Auditory and Visual Association Verbal and Manual Expression. Patterns of Retention and Retrieval of Language.</p>	<p>Individual administration</p> <p>Time: 45 to 60 minutes</p> <p>Examiner Qualifications: Some specific knowledge regarding the evaluation of cognitive and linguistic abilities as well as general test administration experience.</p>	<p>Standard scores Percentile scores Age-equivalent scores Grade-equivalent scores</p>	<p>Use: Pre/post test in a longitudinal study of language growth of children after entering kindergarten since the age range covered by test begins with age 5.</p> <p>Spanish Version: No</p> <p>Limitations: Accurate interpretation of the results requires knowledge and competence in language and psycholinguistics.</p>
<p>Oral and Written Language Scales Listening Comprehension and Oral Expression (OWLS)</p> <p>American Guidance Service, Inc.- 1995</p> <p>Test purchase restricted to qualified professionals. Publisher=s website: http://www.agsnet.com/</p>	<p>Provide a broad measure of a child=s competence in listening comprehension and oral expression.</p>	<p>Test of Oral and Written Language Skills</p> <p>Ages: 3 - 21 years.</p> <p>Domains: Written Expression Oral Expression Listening Comprehension</p>	<p>Individual administration</p> <p>Time: Depending on the age of the child, administration time for the Listening Comprehension Scale is 5 to 15 minutes and for the Oral Expression Scale, 10 to 25 minutes.</p> <p>Examiner Qualifications: Master=s degree in psychology or education or training in assessment.</p>	<p>Standard scores Percentile ranks Sanines Age-equivalent scores</p>	<p>Use: Assess general listening and speaking skills.</p> <p>Spanish Version: No</p> <p>Limitations: Use the Listening Comprehension subtest as a screening device only for children aged 6 to 9 years.</p>

Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>Peabody Picture Vocabulary Test-III</p> <p>American Guidance Service, 1959 - 1997</p> <p>Test purchase restricted to qualified professionals. Publisher=s website: http://www.agsnet.com/</p>	<p>Measure receptive vocabulary for standard English and a screening test of verbal ability.</p>	<p>Test of Hearing Vocabulary</p> <p>Ages: 2 2 to 90 years.</p> <p>Domain: Receptive Vocabulary</p>	<p>Individual administration</p> <p>Time: 10 to 12 minutes</p> <p>Examiner=s Qualifications: Does not require specialized or highly technical knowledge to administer and score.</p>	<p>Standard scores Age-equivalent scores Percentile ranks Normal-curve equivalent scores Stanines</p>	<p>Use: A screening instrument of a child=s verbal ability, or as one element in a full battery assessment of cognitive processes.</p> <p>Spanish version: Yes</p> <p>Limitations: Measures only listening vocabulary which is a more restricted aspect of overall linguistic and cognitive functioning.</p>
<p>Preschool Language Scale-Fourth Edition (PLS-4)</p> <p>The Psychological Corporation - 1992</p> <p>Test purchase restricted to qualified professionals. Publisher=s website: http://www.harcourt.com/index.html</p>	<p>Measures receptive and expressive language.</p>	<p>Test of Oral Language</p> <p>Ages: Birth through 6 years.</p> <p>Domains: Auditory Comprehension Expressive Communication</p>	<p>Individual administration</p> <p>Time: 20 to 45 minutes</p> <p>Examiner=s Qualifications: Master=s degree in psychology or education or training in assessment</p>	<p>Standard scores Percentile ranks Age-equivalent scores</p>	<p>Use: Assess child=s language skills</p> <p>Spanish Version: Yes</p> <p>Limitations: Should not be used alone to obtain a thorough language evaluation.</p>

Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>Receptive One-Word Picture Vocabulary Test (ROWPVT)</p> <p>Academic Therapy-Publications -1985-2000</p> <p>Test purchase restricted to qualified professionals. Publisher=s website: http://www.academictherapy.com</p>	<p>Measure receptive single-word vocabulary on what a child has learned from home or formal education. Spanish-Bilingual version assesses a child=s total acquired vocabulary.</p>	<p>Test of English Hearing Vocabulary</p> <p>Ages: 2 to 18 years</p> <p>Domains: Receptive</p>	<p>Individual administration</p> <p>Time: 10 to 15minutes</p> <p>Examiner Qualifications: Master=s degree in psychology or education or training in assessment</p>	<p>Standard scores Percentile ranks Age-equivalent scores Normal-Curve Equivalent scores Stanines</p>	<p>Use: Assess a child=s skill level of English hearing vocabulary.</p> <p>Spanish Version: Yes</p> <p>Limitations: Should be used along with other measures (i.e., Expressive One-Word Picture Vocabulary Test) to assess a child=s language abilities.</p>
<p>Tests of Early Language Development-Third Edition (TELD-III)</p> <p>PRO-ED, Inc. - 1991</p> <p>No test purchase restrictions. Publisher=s website: http://www.proedinc.com/</p>	<p>Measure receptive and expressive aspects of language structures (i.e., syntax, morphology, and phonology) and meaningful language (i.e., semantics)</p>	<p>Screening Instrument</p> <p>Ages: 3 to 7-11 years</p> <p>Domains: Receptive Language Expressive Language</p>	<p>Individually administered</p> <p>Time: 30 minutes.</p> <p>Examiner=s Qualifications: Does not require extensive experience in test administration</p>	<p>Standard scores Percentile ranks Age-equivalent scores</p>	<p>Use: A screening instrument to assess a child=s developmental level in receptive, expressive, and overall language.</p> <p>Spanish version: No</p> <p>Limitations: Limited usefulness for documenting progress in an intervention program.</p>

SOCIAL-EMOTIONAL					
Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>Ages and Stages Questionnaire: Social-Emotional</p> <p>Brookes Publishing Co.</p> <p>No test purchase restrictions.</p> <p>Publisher Website: http://www.kaplanco.com</p>	<p>Screen children=s social-emotional development.</p>	<p>Parent questionnaire</p> <p>Ages: 6 - 60 months</p> <p>Personal-Social Domain: Self-regulation, Compliance, Communication, Adaptive Functioning, Autonomy Affect Interaction with people</p>	<p>Time: 10-15 minutes</p> <p>Examiner Qualifications: Parent completes items written at the 4th to 6th grade reading level.</p>	<p>Total Scores are compared to cutoff scores</p>	<p>Use: Screen to assess a child=s developmental progress.</p> <p>Spanish Version: Yes</p> <p>Limitations: Has limited usefulness as a pre/post test measure</p>
<p>Devereux Early Childhood Assessment (DECA)</p> <p>Kaplan Early Learning Company, 1999</p> <p>No test purchase restrictions</p> <p>See publisher Website: http://www.kaplanco.com</p>	<p>Screen for emotional and behavioral problems.</p>	<p>Screening Instrument Observation-Based Behavior Rating (Parent, family caregiver, or early childhood professional)</p> <p>Ages: 2 to 5 years</p> <p>Domains: Initiative Self-Control Attachment Behavioral Concerns</p>	<p>Time: 10 minutes</p> <p>Examiner Qualifications: Rating form should be completed by an individual who has observed the child a minimum of 2 hours-day, 2 days-week over the preceding 4 weeks before completing ratings.</p>	<p>T-Scores Percentiles Normal-curve equivalents</p>	<p>Use: Screen to assess developmental level</p> <p>Spanish Version: Yes</p> <p>Limitations: Has limited usefulness as a pre/post test measure.</p>

Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>Social Skills Rating System</p> <p>American Guidance Service, 1990.</p> <p>Test purchase restricted to qualified professionals. Publisher Website: http://www.agsnet.com/</p>	<p>Assess positive social skills and measure child behaviors that can interfere with the development of interpersonal skills in educational and family settings.</p>	<p>Parent Rating Form (Also has teacher and student rating form for elementary level)</p> <p>Ages: 3 to 18 years</p> <p>Domains: Social Skills Cooperation Empathy Assertion Responsibility Self-Control Problem behaviors Externalizing Internalizing Hyperactivity</p>	<p>Time: 10-25 minutes for each questionnaire</p> <p>Examiner Qualifications:</p> <p>Items are written in behavioral terms so the parent does not have to use a high level of inference in order to respond.</p>	<p>Standard scores Percentile ranks Behavior levels (fewer, average, and more)</p>	<p>Use: Pre/Post measure</p> <p>Spanish Version: No</p> <p>Limitations: No significant limitations noted in test reviews.</p>
<p>Vineland Adaptive Behavior Scales</p> <p>American Guidance Service, 1935 - 1985</p> <p>Test purchase restricted to qualified professionals. Publisher website: http://www.agsnet.com/</p>	<p>Assess personal and social skills used for everyday living.</p>	<p>Structured Parent Interview</p> <p>Ages: Birth through 5-11 years</p> <p>Domains: Communication Daily Living Skills, Socialization Motor Skills.</p>	<p>Time: 20-60 minutes</p> <p>Examiner Qualifications: Interviewers should be thoroughly trained and have experience with open-ended interviewing techniques.</p>	<p>Standard scores Percentile ranks Age-equivalent scores Adaptive levels</p>	<p>Use: Screening instrument or to assess child=s developmental level of social competence</p> <p>Spanish Version: Yes</p> <p>Limitations: Not useful in comparing children across ages or performing longitudinal comparisons for the same child.</p>

Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>Vineland Social-Emotional Early Childhood Scales</p> <p>American Guidance Service, 1998</p> <p>Test purchase restricted to qualified professionals. Publisher website: http://www.agsnet.com/</p>	<p>Assess a child=s social skills that could contribute to independent functioning as well as monitor a child=s development.</p>	<p>Semi-Structured Parent Interview</p> <p>Ages: 3 to 18 years</p> <p>Domains: Interpersonal Relationships Play and Leisure Time Coping Skills.</p>	<p>Time: 15 to 25minutes.</p> <p>Examiner Qualifications: Interviewee should be the person most knowledgeable about the child=s social and emotional functioning.</p>	<p>Standard scores Percentile ranks with confidence intervals Age-equivalent scores</p>	<p>Use: Pre/Post measure</p> <p>Spanish Version: Yes</p> <p>Limitations: Norms were established from the Vineland Adaptive Behavior Scale in 1980 and need to be updated.</p>

DEVELOPMENTAL SCREENING INSTRUMENTS

Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>Ages and Stages Questionnaire</p> <p>Paul H. Brookes Publishing Co. - 1995</p> <p>No test purchase restrictions.</p> <p>Publisher=s website: http://brookespublishing.com</p>	<p>Screen to identify children to determine need for further assessment</p>	<p>Parent Questionnaire</p> <p>Ages: 4 months through 60 months (5 years)</p> <p>Domains: Communication Gross Motor Fine Motor Problem Solving Personal-Social</p>	<p>Parent Report</p> <p>Time: 10-15 minutes</p> <p>Examiner Qualifications: Parent completes items written at 4th to 6th grade reading level.</p>	<p>Total scores are compared to cutoff points in the five domains to separate young children who require referral and assessment from those who do not.</p>	<p>Use: Screening instrument to assess a child=s developmental progress.</p> <p>Spanish version: Yes</p> <p>Limitation: Should not be used as a pre/post measure.</p>
<p>AGS Early Screening Profiles</p> <p>American Guidance Service Publishers 1990</p> <p>Test purchase restricted to qualified professionals.</p> <p>Publisher website: http://www.agsnet.com</p>	<p>Screen major areas of functioning of children and their families.</p>	<p>Screening instrument</p> <p>Ages: 3 through 6 years 11 mos.</p> <p>Domains (child): Cognitive-language Motor Articulation.</p> <p>Parent Questionnaires: Home survey Health survey Self-Help Social Profile</p>	<p>Individual administration for child measures; Parent completes questionnaires</p> <p>Time: 15 to 30 minutes</p> <p>Examiner=s Qualifications: Can be administered by paraprofessionals.</p>	<p>Normal-curve equivalents Age-equivalent scores Stanines</p>	<p>Use: Screening instrument that provides some information on children=s expressive and receptive language skills.</p> <p>Spanish Version: No</p> <p>Limitations: Scores have relatively low agreement with those of other screening instruments such as the DIAL-R and Battelle.</p>

Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>Battelle Developmental Inventory Screening Test</p> <p>Riverside Publishing, 1984</p> <p>Test purchase restricted to qualified professionals. Publisher website: http://riverpub.com</p>	<p>A general screening measure for preschool and kindergarten children, preliminary assessment, and/or initial identification of developmental strengths and weaknesses.</p>	<p>Screening Instrument</p> <p>Ages: Birth to 8 years.</p> <p>Domains : Personal-Social Adaptive Motor (Gross/Fine Motor) Communication (Receptive, Expressive) Cognitive</p>	<p>Individual Administration, direct observation, and interview</p> <p>Time: Approximately 30 minutes.</p> <p>Examiner Qualifications: Requires extensive training to obtain valid scores.</p>	<p>Total scores are compared to cutoff scores Age-equivalent scores</p>	<p>Use: Can compare a child=s developmental level in the five domains with other children of the same age.</p> <p>Spanish Version: No</p> <p>Limitations: Not appropriate for Asian or Native American children because they were excluded from norming sample.</p>

Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>Brigance Screens</p> <p>Brigance Early Preschool Screen (ages three and four)</p> <p>Brigance K & I (Kindergarten and first grade)</p> <p>Curriculum Associates, Inc. 1978-1991</p> <p>Test purchase restricted to qualified professionals. Publisher website http://www.curricassoc.com</p>	<p>Screen a child=s basic skills and behavior in order to identify the child who should be referred for more comprehensive evaluation.</p>	<p>Screening Instrument</p> <p>Ages: 2 to 6 years</p> <p>Domains on Brigance Preschool Screen: Personal Data Response Identifies Body Parts Gross Motor Skills Identifies Objects Repeats Sentences Visual Motor Skills Number Concepts Builds Tower with Blocks Matches Colors Picture Vocabulary Plurals</p>	<p>Parent ratings, child=s performance on screening tasks, and observational data.</p> <p>Time: 15 minutes</p> <p>Examiner=s Qualifications: Can be administered by paraprofessionals.</p>	<p>Total scores are compared to a recommended cutoff score of 60.</p>	<p>Use: Quickly screen large numbers of children.</p> <p>Spanish version: No</p> <p>Limitations: Scores do not represent how well the child performed compared to other children of the same age.</p>
<p>Denver II</p> <p>Denver Developmental Materials, Inc. 1967-1990</p> <p>No test purchase restrictions. Publisher=s website: http://denverii.com</p>	<p>Screen for developmental delays</p>	<p>Screening Instrument</p> <p>Ages: Birth to 6 years</p> <p>Domains: Personal-Social Fine Motor-Adaptive Language Gross Motor</p>	<p>Individual administration</p> <p>Time: 20 minutes.</p> <p>Examiner Qualifications: Manual recommends that test administrators receive training in the administration of the instrument before using the test.</p>	<p>Each item is scored as pass, fail, no opportunity or refusal, and the performance is interpreted as normal, suspect, or untestable. Each of the 125 items is represented on the form by a bar than spans the ages at which 25%, 50%, 75%, and 90% of the standardization sample passed.</p>	<p>Use: Screen for developmental delays.</p> <p>Spanish Version: Yes</p> <p>Limitations: Few noted by test reviewers.</p>

Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>Developmental Indicators for the Assessment of Learning - Third Edition - DIAL-4</p> <p>American Guidance Service, Inc. 1983-1998</p> <p>Test purchase restricted to qualified professionals. Publisher's website: http://www.agsnet.com</p>	<p>Identify young children in need of further diagnostic assessment.</p>	<p>Screening Instrument</p> <p>Ages: 3 through 6-11 years.</p> <p>Domains: Motor Concepts Language Self-Help Social Skills.</p> <p>The short form of the test, Speed DIAL, measures Motor Skills, Concepts, and Language. Parent Questionnaire measures child's self-help and social skills.</p>	<p>Individual administration and parent questionnaire</p> <p>Time: 20 to 30 minutes Speed DIAL is 15 to 20 minutes.</p> <p>Examiner Qualifications: Coordinator of assessment team should have an early childhood or special education background.</p>	<p>Standard deviation and percentile cutoff points by chronological age at two-month intervals for total and area scores.</p>	<p>Use: Screen for developmental delays.</p> <p>Spanish Version: Yes</p> <p>Limitations: Norms for English-speaking children may be more valid than those indicated for Spanish-speaking children.</p>
<p>Early Screening Inventory-Revised (ESI-R)</p> <p>Pearson Early Learning 1976-1997</p> <p>No test purchase restrictions. Publisher's website: http://www.pearsonearlylearning.com</p>	<p>Identify children who may need additional support or services</p>	<p>Screening Instrument</p> <p>Ages: 3 to 6 years.</p> <p>Domains: Speech Language Cognition Perception Gross and Fine Motor Coordination</p>	<p>Individual administration accompanied by a parent questionnaire</p> <p>Time: 15 to 20 minutes</p> <p>Examiner Qualifications: It is recommended that test administrators view the two training videos before commencing testing children.</p>	<p>Total scores are reported as AOK@, ARescreen@ (1 to 2 standard deviations below mean) or ARefer@ (greater than 2 standard deviations below the mean).</p>	<p>Use: A measure of developmental skills.</p> <p>Spanish Version: Yes</p> <p>Limitations: The instrument does not measure behaviors/skills in the social-emotional domain, one of the five domains for school readiness.</p>

PARENTING KNOWLEDGE AND SKILLS					
Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>The Adult-Adolescent Parenting Inventory (AAPI-2)</p> <p>Family Nurturing Centers, Int. No test purchase restrictions. Publisher=s website: www://Nurturingparenting.com</p>	<p>To assess the parenting attitudes and child rearing practices of both adult parent and pre-parent populations as well as adolescent parent and pre-parent individuals.</p>	<p>Parent Report</p> <p>Ages: Parent and adolescent parent and adolescent non-parent populations.</p> <p>Sub-scales:</p> <ul style="list-style-type: none"> * Inappropriate Parental Expectations * Parental Lack of Empathy Towards Children=s Needs * Strong Belief in the Use of Corporal Punishment as a Means of Discipline * Reversing Parent-Child Role Responsibilities * Oppressing Children=s Power and Responsibilities 	<p>Individual or group administration</p> <p>Time: 20 minutes</p> <p>Items are written at the 5th grade level</p>	<p>Standard scores</p>	<p>Use: Pre/Post test measure</p> <p>Spanish Version: Yes</p> <p>Limitations: The AAPI-2 was not reviewed in the Buros Mental Measurements Yearbook.</p>

Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>Parent-Child Relationship Inventory</p> <p>Western Psychological Services - 1994</p> <p>Test purchase restricted to qualified professionals. Publisher website: http://www.wpspublish.com</p>	<p>Assess parents= attitudes toward parenting and their children.</p>	<p>Parent Questionnaire</p> <p>Ages: Parents aged 18 to 54 years.</p> <p>Content Scales: Parent Support Satisfaction with Parenting Involvement Communication Limit Setting Autonomy Role Orientation scale.</p>	<p>Parent Report</p> <p>Individual or group administration</p> <p>Time: 15 minutes</p> <p>Items are written at the fourth grade level</p>	<p>T-scores Percentiles Separate norms are provided for mothers and fathers.</p>	<p>Use: The measure can be used in both research and clinical settings to identify specific areas of difficulty between parents and their children.</p> <p>Spanish version: No</p> <p>Limitations: Parents with multiple children need to complete a test for each child. Demographic variables should be considered when interpreting scores.</p>

Instrument	Purpose	Description	Administration	Scores	Use and Limitations
<p>Parenting Stress Index - Third Edition</p> <p>Psychological Assessment Resources 1983-1995</p> <p>Test purchase restricted to qualified professionals. Publisher's website: http://www.parinc.com</p>	<p>For the early identification of parenting and family characteristics that fail to promote normal development and functioning in children, children with behavioral and emotional problems, and parents who are at risk for dysfunctional parenting.</p>	<p>Screening Instrument</p> <p>Ages: Parents of children 12 years old or younger.</p> <p>Child Domains: Distractibility/hyperactivity Adaptability Reinforces Parent Demandingness Mood Acceptability</p> <p>Parent Domains: Competence Isolation Attachment Health Role Restriction Depression Spouse</p>	<p>Parent Report</p> <p>Can be administered individually or as a group activity.</p> <p>Time: 20 minutes</p> <p>Items are written at a 5th grade reading level.</p>	<p>Percentile ranks Percentiles Profile analysis can be plotted on the scoring form. Total Stress score(the sum of the Parent and Child domain scores)</p>	<p>Use: Could be used as a screening tool to identify children who may be at risk for emotional or behavioral problems, as well as parents who may be in need of parent education.</p> <p>Spanish version: Yes</p> <p>Limitations: Poor support for subtest organization and profile analysis indicates that global scores should be used for screening purposes only.</p>

HOME ENVIRONMENT

<p>Home Observation for Measurement of the Environment (HOME)</p> <p>University of Arkansas at Little Rock Published: 1978 - 1984</p> <p>There are no test purchase restrictions. Email: Dr. Betty Caldwell [bmcaldwell@ualr.edu]</p>	<p>Measures the quantity and quality of the types of stimulation and support available to a child in the home environment.</p>	<p>Observation-Parent interview</p> <p>Ages: Birth to age 3, early childhood, middle childhood.</p> <p>The HOME consists of two inventories: one for Families of Infants and Toddlers (birth to age 3) with six subscales consisting of 45 items and a second for Families of Preschool Children (aged three to six) that has eight subscales and 55 items.</p>	<p>Two-thirds of the instrument involves observer assessment of parent-child interaction and one-third involves parent (or primary caregiver) report.</p> <p>Time: one hour</p> <p>Examiner Qualifications: Paraprofessionals can administer the instrument provided they follow the instructions in the manual and review the training tapes.</p>	<p>Items are scored as Ayes@ or Ano@ so that judgment is not required as to quality. A summary of items correct by subscale allows the user to assess the percentile band in which the HOME score falls.</p>	<p>Use: Pre/Post measure.</p> <p>Spanish version: No</p> <p>Limitations: Manual has limited information on the extent of training required to development reasonable proficiency in administering the instrument.</p>
--	--	---	--	--	--

Boehm Test of Basic Concepts - III -Preschool Version

Criterion-referenced Assessment

Author: Anne E. Boehm

Publisher: The Psychological Corporation, 1984-86

2003 price data: \$134 per complete kit including picture book, 35 individual record forms, manual and class record.

Administration Time

20-30 minutes.

Scores

Percentiles by 6-month age bands.

Purpose

Designed to measure preschool children=s (ages 3.0 to 5-11 years) knowledge of twenty-six basic relational concepts considered necessary for achievement in the beginning years of school.

Description of Instrument

The BTBC-PV tests knowledge of: (1) concepts about space, such as Ato@ and Anext to@; (2) concepts about quantity, such as Asome@ and Afew@; (3) concepts about time, such as Aafter@ and Anever@; and (4) miscellaneous concepts, such as Adifferent@ and Amatches@. The Preschool version is an individually administered test where the child responds to the items by pointing to pictures instead of having to provide a written response. There are two parallel forms and in each, all concepts are tested twice to determine the child=s understanding of it across contexts. Test materials are attractive and interesting enough to hold the attention of a child.

Strengths:

- § Administration is easy and quick, facilitated by the easel presentation and child=s pointing responses.
- § Manual is well organized, clear, and easy to read.
- § Scoring is straightforward, requiring minimal manipulation of test data in order to obtain percentiles and T-scores.
- § Children respond favorably to the colorful stimulus materials.
- § Each concept is tested twice to determine the child=s understanding of it across contexts.

Limitations

- § The test provides only some of the information necessary in screening for school readiness. For example, the expressive use of concepts is not measured.
- § Test directions-scoring require subjective judgment on the part of examiner. The directions indicate that the examiner can Ahelp@ the child with the warm-up items, if necessary, but it does not specify what Ahelp@ entails (repeating the item, eliminating one of the distractors, teaching a strategy, giving the answer, etc.). Because the manual directs the examiner to discontinue the test if two of the last four warm-up items are missed even with additional help, the degree of help provided on these items would affect the decision about whether or not to discontinue the test.

Bracken Basic Concept Scale (BBCS-R)

Full Battery Assessment

Author: Bruce Bracken (1998)

Publisher: Psychological Corporation

2003 price data: \$255 complete kit including examiner=s manual, stimulus manual, 25 diagnostic scale record forms and one screening test.

Administration Time

30 minutes.

Scores

Standard scores, age equivalency scores, and percentile ranks.

Purpose

Designed to measure basic concept acquisition and receptive language of children aged 2 2 to 8 years in the domains of Color, Letters, Numbers, Comparisons, Shapes, Direction, Social/Emotional connotations, Size, Texture, Quantity, and Time.

Description of Instrument

There are 300 concepts in eleven conceptual domains: Colors, Letters, Numbers-Counting, Sizes, Comparisons, Shapes, Direction-Position, Self-Social Awareness, Texture-material, Quantity, and Time-Sequence. The School Readiness Composite is based on the first six subtests: Colors, Letters, Numbers-Counting, Sizes, Comparisons and Shapes. The other five domains were meant to stand on their own. Test stimuli are full color pictures and are balanced in terms of race-ethnicity.

Strengths:

- § Can be used as a school readiness assessment.
- § Test correlates strongly with the Wechsler Preschool and Primary Scale of Intelligence-Revised.
- § Instrument is easy to administer and score. Children indicate their responses through pointing, vocalizing, eye gaze, or eye blink rather than being required to give a verbal response.
- § Testing materials are attractive to children and very Auser-friendly@. There is an 8.5 by 11-inch stimulus manual featuring colored artwork with tabbed dividers separating subtests, an expanded examiner=s manual, and a carrying case.
- § There is a Spanish version that is based on item translation and expert review.

Limitations

- § Examiner required to establish basal and ceiling levels during testing. However, there is a simplified procedure outlined in the manual for establishing a child=s basal level.
- § Additional assessment measures and tools should be employed when classification or placement is a goal of testing.
- § The Spanish version is only to be used as a criterion-referenced or curriculum-based measure due to sample size (n=193).

The Brigance Diagnostic Inventory of Early Development - Revised

Cognitive Screener or Skill Assessment

Author: Albert Brigance**Publisher:** Curriculum Associates, Inc. , 1978-1991**2003 price data:** \$124 per manual with tests; \$19.90 per 10 record books**Administration Time**

Not provided by publisher.

Scores

Each basic skill is regarded as distinct, and there is no cumulative score associated with a skill area or subarea.

Purpose

To assess and track the developmental skills of children aged birth to seven years in eleven broad skill areas: Preambulatory Motor Skills and Behaviors, Gross-Motor Skills and Behaviors, Fine-Motor Skills and Behaviors, Self-Help Skills, Speech and Language Skills, General Knowledge and Comprehension, Social and Emotional Development, Readiness, and Basic Reading Skills Manuscript Writing.

Description of Instrument

Methods used to assess skills include a parent interview, observation of the child, having the child perform tasks, engaging the child in conversation, and a teacher interview. Skills are sequenced within a subarea according to developmental age. This item arrangement is based on the link between the basic skills being evaluated and the approximate developmental age at which mastery is normally achieved. Given the chronological age of the child, examiners can use developmental age as an index for choosing an appropriate starting point within a skill subarea. When administering the Inventory, the examiner and child sit opposite each other with the Inventory placed between them. While the examiner reads the instructions, the child observes the visual material and responds to questions. Each broad skill area is broken down into overall goals and objectives, methods of assessment, assessment directions, required test materials, and the references used to validate the sequencing of skills. Questions are presented along with helpful hints for determining successful mastery. Each item is coded (a) not assessed, (b) assessed and set as an objective, (c) introduced but not achieved, and (d) skill has been achieved.

Strengths:

- § Broad skill areas evaluated by the test constitute a broad cross-section of the behaviors and skills associated with early childhood development.
- § Little specialized training is required to administer the test.
- § The Developmental Record Book, a bound booklet which consists of an ordered listing of all basic skills and behaviors, simplifies the process of recording child=s responses.
- § There is built-in examiner flexibility because the examiner decides which of the 11 broad skill areas should be evaluated and which subareas should be identified as goals and reassessed.
- § Used for program evaluation by several HIPPIY programs including sites in Arkansas, Louisiana, Maryland, and Texas.

Limitations:

- § Not useful for pre/post testing because there is no cumulative score associated with a skill area or subarea.
- § Requires discretion on the part of the examiner to determine when a sufficient number of questions have been answered correctly to justify area or subarea mastery.

Detroit Test of Learning Aptitude - Primary Second Edition (DTLA-P-2)

Cognitive Screening Instrument

Authors: Donald Hammill and Bryan Bryant

Publisher: PRO-ED, Inc., 1986-1991

2003 Price Data: \$164 per complete kit including picture book, 25 response forms, 25 profile-examiner record forms, and manual.

Administration Time

15 to 45 minutes.

Scores

Standard scores, percentile ranks, and age equivalency scores.

Purpose

Designed to identify the strengths and weaknesses of children aged 3-0 to 9-11 years in three developmental domains: Linguistic, Attention, and Motoric.

Description of Instrument

There are 100 dichotomously scored items (0=fail, 1=pass) that are classified into 16 different tasks measuring the ability to (a) articulate speech sounds, (b) match semantic concepts; (c) reproduce designs, (d) repeat digits, (e) draw a figure of a person, (f) sequence letters, (g) follow directions that involve manual dexterity, (h) sequence pictured objects, (i) follow oral directions, (j) identify fragmented pictures, (k) repeat sentences, (l) solve visual abstract reasoning problems, (m) call upon visual discrimination skills, (n) produce antonyms, (o) repeat a series of unrelated words, and (p) identify pictured objects. Each child begins the test at an entry point corresponding to his or her chronological age. Testing continues until eight consecutive items are missed (ceiling.). The number of correct responses is determined by established basal and ceiling points. It is recommended test examiners have formal training in assessment.

Strengths:

- § Has a clearly articulated, well-written administration manual. Discussion of the development, scoring, use, and interpretation of the instrument is written in a style and language readable by educators.
- § Scores correlate well with scores from other cognitive tests such as the Woodcock-Johnson Psycho-Educational Battery and the Peabody Picture Inventory Test-Revised.
- § Software scoring and report system available from publisher.

Limitations:

- § There are many different types of tasks, and the test authors recommend test examiners should be thoroughly familiar with the administration directions in the manual.
- § During testing the examiner fills in a formatted score sheet and is encouraged to write any anecdotal records that may aid in score interpretation.
- § It may be difficult for examiner to establish either a basal or ceiling point precisely for children with short attention spans.

Kaufman Survey of Early Academic and Language Skills (K-SEALS)

Full Battery Assessment

Authors: Alan Kaufman and Nadeen Kaufman

Publisher: American Guidance Service, 1993

2003 price data: \$223 per complete kit including manual, presentation easel, and 25 record booklets.

Administration Time

15-25 minutes.

Scores

Standard scores, age equivalency scores, and percentile ranks.

Purpose

Designed to measure the language skills (expressive and receptive vocabulary), pre-academic skills, and articulation skills of preschool and kindergarten children aged 3-0 to 6-11 years.

Description of Instrument

The K-SEALS is an expanded version of the AGS Early Screening Profiles (Cognitive-Language Profile) published in 1990. The test is composed of three subtests: Vocabulary, Numbers-Letters and Words, and Articulation Survey. The first two tests are split to make up two Language Scales (Expressive Skills and Receptive Skills) and two Early Academic Scales are to be used with children aged 5-0 to 6-11.

Strengths:

- § K-SEALS correlates highly with the Kaufman Assessment Battery for Children and the Stanford-Binet. The K-SEALS also correlates well with teacher ratings of child=s ability.
- § Test reviewers state the manual presents clear and complete instructions for administration and well as preparing raw score summaries.
- § Can be easily administered by a variety of professionals working with young children. It is helpful if examiner is knowledgeable about strategies to engage children with more challenging test behaviors.

Limitations:

- § Computation of standard scores and confidence intervals is somewhat complex because it involves combining information from several tables in the manual=s appendix.
- § Test reviewer felt it was unclear whether the primary use of the K-SEALS is that of a screening instrument or a diagnostic measure.
- § Test reviewer expressed some concern in interpreting test scores of children from culturally-linguistically diverse backgrounds and recommended using only the overall composite score.

McCarthy Scales of Children=s Abilities (MSCA)

Full Battery Assessment

Author: Dorothea McCarthy

Publisher: The Psychological Corporation, 1970-78

2003 price data: \$595 per complete kit including manual and all necessary equipment, 25 drawing booklets, and 25 record booklets.

Administration Time

45 minutes for children under five; 1 hour for older children.

Scores

Standard scores, Mental Age scores, and an overall General Cognitive Score which indicates mental competence but is not synonymous with an IQ score.

Purpose

Designed to measure the cognitive and motor development of children aged 2-4 to 8-7 years.

Description of Instrument

The McCarthy Sales of Children=s Abilities is a well standardized and psychometrically sound measure consisting of the following six scales: Verbal Scale (the ability to understand and process verbal stimuli and express thoughts); Perceptual-Performance Scale (visual-motor coordination and nonverbal reasoning through manipulation of concrete materials), Quantitative Scale (facility in dealing with numbers and understanding of quantitative concepts), Memory Scale (short-term memory across a wide range of visual and auditory stimuli), Motor Scale (gross and fine motor coordination)and General Cognitive Scale (reasoning, concept formation, and memory when solving verbal and numerical problems and when manipulating concrete materials). Test materials were designed to appear as games to children such as block building and puzzle solving. For example, in the block building task, the examiner will say, ASee these blocks we have to play with. Let=s see if you can make a tower just like this one (examiner points to picture).@

Strengths:

- § Administration manual is exemplary and contains elaborate information about the test=s psychometric soundness, the standardization process, norms tables, and guidelines for administration and interpretation.
- § Correlates very well with Sanford-Binet IQ test, the Metropolitan Achievement Tests and the Peabody Individual Achievement Test.
- § Has been used in the National Head Start Impact Study and Family and Child Experiences Study.
- § The game-like and nonthreatening nature of the test materials makes them very attractive to preschool children. For example, The sequence in which the subtests are ordered serves to engage young children.

Limitations:

- § Test administration requires approximately 45 to 60 minutes.
- § Proper administration of the 18 different subtests requires keen attention to detail acquired through considerable practice and careful readings of the manual.
- § The MSCA does not measure some areas of children=s functioning such as social comprehension, maturity, and judgment in younger children (social-emotional domain).

Woodcock-Johnson III

Full Battery Assessment

Authors: Richard Woodcock, Kevin McGrew, Nancy Mather, and Fredrick Schrank.

Publishers: Riverside Publishing Company, 1977-2001

2002 price data: \$700 per Tests of Cognitive Abilities Battery including Cognitive Standard and Extended test books, examiner=s manual, technical manual, audio cassette, 25 test records, 25 response booklets, Compuscore and profiles software program, and scoring guides for hand scoring.

Administration Time

Approximately five minutes per test; 35-115 minutes for each full battery.

Scores

Grade-equivalency scores, age equivalency scores and percentile ranks.

Purpose

Measures general and specific cognitive abilities, scholastic aptitude, oral language, and academic achievement for individuals aged 2 years to 90 years old.

Description of Instrument

The Woodcock-Johnson test consists of two separate batteries. The Woodcock-Johnson III Test of Cognitive Abilities measures general and specific cognitive functions, and the Woodcock-Johnson III Test of Achievement determines and describes the present status of an individual=s academic strengths and weaknesses. The batteries are further classified as standard or extended and both versions have parallel Forms A and B. Test materials include folding, easel-type binders that permit stimulus materials to be viewed easily by the child, while also affording a barrier so that the child cannot observe the recording and scoring of responses. There are separate binders for the Tests of Cognitive Abilities and Tests of Achievement as well as separate binders for the standard and extended versions of the tests. Although the Woodcock-Johnson III is appropriate for ages two to 90+, not all of the tests are administered at every age. At the preschool level, only six of the twelve cognitive ability tests and five of the 10 achievement tests are applicable. None of the five interests tests (reading interest, math interest, written language interest, physical interest and social interest) were designed for children below three years of age.

Strengths:

- § Test examiner can administer only selected subtests which reduces unnecessary testing.
- § Computerized scoring (Compuscore) is available. In addition to standard reports, the software allows the user to make modifications to printed output such as printing summary narratives of a child=s performance in English or Spanish.
- § Audio tapes are included for subtests requiring standardized presentation of oral material.
- § Test has been used extensively in longitudinal research and national program evaluations such as the Early Childhood Longitudinal Study-Kindergarten and the National Head Start Impact Study.

Limitations:

- § Test administrators need specific training in administration and scoring. Some parts of the test are strictly timed because they measure speed and fluency. Also, examiner is required to establish a child=s basal and ceiling level, the point at which the test begins and ends.

COMMUNICATION

Comprehensive Receptive and Expressive Vocabulary Test - Second Edition

Test of Oral Language Skills

Authors: Gerald Wallace and Donald Hammil

Publisher: PRO-ED, Inc. 1994-2002

2002 Price: \$226 per complete kit.

Administration Time

20 to 30 minutes.

Scores

Standard scores, percentile ranks and age equivalency scores.

Purpose

Identify deficiencies in oral vocabulary, identify discrepancies between receptive and expressive vocabulary; document progress in instructional programs and serve as a research tool that can be used to measure receptive skills of individuals aged 4 - 89 years and expressive skills of individuals aged 5 - 89 years.

Description of Instrument

The Comprehensive Receptive and Expressive Vocabulary Test-Second Edition is an untimed, individually administered test composed of one receptive and one expressive subtest. The Receptive Vocabulary Subtest consists of 76 items presented in 60 color photographs of people or objects. The child looks at a page containing six photographs and points to the one that best depicts the word the examiner says. There are 10 pages, each showing a different category: animals, transportation, occupations, clothing, food, personal grooming, tools, household appliances, recreation, and clerical materials. The Expressive Vocabulary Subtest consists of 29 single words spoken aloud by the examiner and the child is to provide a definition for each word presented.

Strengths:

- § Two equivalent forms of the test are available making it useful for pre and post testing.
- § Test materials are appealing (clear, realistic color photographs).
- § Administration and scoring are quick and simple. Children of all ages start with the first item.

Limitations

- § Test does not adequately distinguish individuals with language abilities that are above or below the average range.
- § The test employs normative data from previous versions that are up to 10 years old.
- § Reliability and validity data range from weak (criterion validity) to excellent (test-retest reliability).
- § Expressive test is appropriate only for children aged 5 years or older.

COMMUNICATION

Expressive One-Word Picture Vocabulary Test - 2000 Edition

Test of Expressive Vocabulary

Author: Morrison Gardner

Publisher: Academic Therapy Publications, 1990

2002 price data: \$140 per kit including manual.

Administration Time

10 - 15 minutes.

Scores

Standard scores, percentile ranks, and age equivalency scores.

Purpose

Assess the English speaking vocabulary of children aged two to eighteen years of age. The Spanish/Bilingual edition assesses the total acquired vocabulary of individuals who are bilingual in Spanish and English.

Description of Instrument

The examiner presents the child with a series of illustrations representing objects, concepts, or actions, and the child is asked to name each illustration. An approximation is made by the examiner to determine a starting point for the presentations. Testing is concluded when the child is unable to correctly name a number of consecutive illustrations. As the testing proceeds, the difficulty of the test items increases.

Strengths:

- § There are national norms and the instrument was co-normed with the Receptive One-Word Picture Vocabulary Test which permits the comparative analysis of both the child's ability to use language in speaking and in writing.
- § Scores have been correlated with vocabulary subtests of intelligence tests such as the Weschler Intelligence Scale for Children and the Stanford-Binet Intelligence Scales.
- § Speed of administration (Administration time is 10 to 15 minutes).
- § There is a Spanish/Bilingual edition.
- § Test has been used as an outcome measure in dialogic reading research and national program evaluations such as the National Longitudinal Study Children and Youth, the Early Childhood Longitudinal Study-Kindergarten and the National Head Start Impact Study.

Limitations

- § Should be used with the **Receptive One-Word Picture Vocabulary** test to measure a child's ability to use language in speaking and in writing and the ability to recognize and understand the spoken or written words of others.

COMMUNICATION

Illinois Test of Psycholinguistic Abilities - Third Edition

Full Battery Assessment

Authors: Donald Hammill, Nancy Mather, and Rhia Roberts

Publisher: PRO-ED, Inc. , 1961-2001

2003 price data: \$164 per kit including 25 profile-examiner record booklets.

Administration Time

45 to 60 minutes.

Scores

Standard scores, percentile scores, age equivalency scores and grade equivalency scores.

Purpose

Assess the ability of children aged 5 through 12 years to understand, organize, and express information in both an auditory-vocal and a visual-motor channel.

Description of Instrument

The Illinois Test of Psycholinguistic Abilities consists of 12 subtests designed to function across three levels of organization (reflexive, form, and content), three psycholinguistic processes (reception, expression, and association), two channels of input (auditory and visual), and two channels of output (vocal and manual). Six of the subtests are verbal and require the child to respond to a verbal presentation of an item, and the remaining six subtests are written requiring the child to respond to a written presentation of an item. The written subtests are appropriate for children aged 6 years 6 months or older.

Strengths:

- § Test scoring is simple with tables provided that convert raw scores to percentile scores, standard scores and age and grade equivalency scores.
- § Has been used as an outcome measure in dialogic reading and emergent literacy research.

Limitations:

- § Administration time is rather lengthy; however, the test can be administered in multiple sessions.
- § Examiner should have some specific knowledge regarding the evaluation of cognitive and linguistic abilities as well as general test administration experience (experience in establishing basal and ceiling limits).
- § As the age range covered by the test begins with age 5, the instrument would not be appropriate for program evaluation purposes. However, it could be utilized in a longitudinal study of language growth of HIPPIY children in public school.

COMMUNICATION

Oral and Written Language Scales Listening Comprehension and Oral Expression (OWLS)

Test of Oral and Written Language Skills

Author: Elizabeth Carrow-Woolfolk

Publisher: American Guidance Service, Inc., 1995

2003 price data: \$289.99 per complete kit including manual, Listening Comprehension Easel, Oral Expression Easel, and 25 record forms.

Administration Time

Depending on the age of the child, administration time for the Listening Comprehension Scale is 5 to 15 minutes and for the Oral Expression Scale, 10 to 25 minutes.

Scores

Standard scores, percentile ranks, stanines, and age equivalency scores

Purpose

The OWLS provides a broad measure of a child's competence in listening comprehension and oral expression in three domains: Written Expression, Oral Expression, and Listening Comprehension.

Description of Instrument

The OWLS is administered individually to children and young adults, aged 3 to 21. Listening Comprehension is measured by asking the child to select one of four pictures that best depicts a statement made by the examiner (e.g., "In which picture is she not walking to school"). Oral expression is assessed by asking the child to look at one or more line drawings and respond verbally to a statement made by the examiner (e.g., "Tell me what is happening here and how the mother feels."). An interesting feature of the Oral Expression subtest is that the examiner can conduct a descriptive analysis of correct and incorrect responses. For all but 30 of the 96 items on this subtest, correct responses can be categorized as preferred or acceptable responses, providing additional information on how well the child understood the oral expression task. Incorrect responses can further be classified as a miscue involving grammar or a miscue involving semantic and-or pragmatic aspects of language.

Strengths:

- § Both the Listening Comprehension and Oral Expression subtests were constructed on the basis of a strong theoretical foundation.
- § Both subtests are easy to administer, requiring only about 15 to 40 minutes for both tests depending on the age of the child.
- § OWLS scores are moderately to highly correlated with some measures of achievement and cognitive development (e.g., Kaufman Tests of Educational Achievement and the Peabody Individual Achievement Test).
- § Provides an opportunity to observe a child's ability to understand and produce connected language which is necessary to function in school or social settings

Limitations:

- § Some tasks on the tests are not typical of those found in the classroom.
- § The Listening Comprehension Scale should be used only as a screening device for children aged 6 to 9 because the reliability indices were below .80 for this particular scale.
- § Examiner must have specific education or training in children's assessment and experience in individual test administration.

COMMUNICATION

Peabody Picture Vocabulary Test-III

Test of Hearing Vocabulary

Authors: Lloyd M. Dunn and Leota M. Dunn

Publisher: American Guidance Service, 1959 - 1997

2003 price data: \$270 per complete kit including all testing materials including a carrying case that converts into an easel for convenient presentation of the picture plates.

Administration Time

10 to 12 minutes.

Scores

Standard scores, age equivalency scores, percentile ranks, normal curve equivalency scores and stanines.

Purpose

Originally designed as a brief intelligence test, the third edition is a measure of receptive vocabulary for Standard English of preschool children and adults (individuals aged 2 years 6 months to 90 years) and a screening test of verbal ability .

Description of Instrument

Test stimuli consist of large, clear, black-and-white drawings, with four pictures per plate. There are two alternate forms and each form has 204 items that are grouped into 17 sets of 12 items. Items are administered by complete sets and once testing begins, every item is administered in the set. A child is shown four pictures and is read a word. The child must either point to the picture most closely depicting the correct word or state the number of the correct picture.

Strengths:

- § The test folder converts into an easel for ease of administration.
- § Has two parallel forms for reliable testing and retesting.
- § Does not require specialized or highly technical knowledge to administer and score.
- § Quick administration (10 to 12 minutes) and scoring. Basal and ceiling points are established by sets, not by consecutive items and manual has step-by-step instructions for raw score conversions.
- § Scores correlate highly with Wechsler Intelligence Scale for Children-Revised
- § Computer scoring available (Computer ASSIST).
- § Has been used extensively as an outcome measure in dialogic reading and emergent literacy research and in national program evaluations such as the Comprehensive Child Development Program, Parents as Teachers, Early Childhood Longitudinal Study-Kindergarten, Head Start Impact Study 2001, and Family and Child Experiences Study (FACES).

Limitations:

- § Measures only the listening vocabulary which is a more restricted aspect of overall linguistic and cognitive functioning.

COMMUNICATION

Preschool Language Scale-Third Edition (PLS-4)

Test of Oral Language

Authors: Irla Zimmerman, Violette Steiner, and Roberta Pond

Publisher: Psychological Corporation, 1992

2003 Price data: \$235 per complete kit including 12 record forms, picture book, and manual.

Administration Time

20 to 45 minutes.

Scores

Standard scores, percentile ranks, age equivalency scores and language age equivalency scores.

Purpose

Designed to assess expressive language and auditory comprehension for children from birth to age 6.

Description of Instrument

This is a standardized, norm-referenced test that focuses on four language aspects: language precursors (attention, vocal development, social communication); semantics (vocabulary, and concepts including quality, quantity, spatial, and time-sequence); structure (morphology, syntax); and integrative thinking skills.

There is an easel-style picture manual containing uncluttered, full-color drawings as stimuli for specific items. Each of two 48-item subscales, Auditory Comprehension and Expressive Communication, includes four items or tasks at each of 12 age levels. Many tasks may be scored *pass* if spontaneously performed by the child at any time during testing, even if failed initially.

Strengths:

- § Examiner's manual contains especially clear, concise, and thorough administration and scoring directions, supplemented with abundant examples.
- § A Spanish version is available with normative data provided item by item, but not as a total score.
- § Standard scores, percentile ranks, and age equivalent scores for raw scores are provided in six-month age intervals for children aged one through 4 years and in 12-month intervals for five and six year olds.
- § Has been used by HIPPY programs in Nevada and New York HIPPY programs which can be contacted for further information regarding the suitability of the instrument for measuring HIPPY child outcomes.

Limitations:

- § Some test materials needed must be purchased separately for \$59.00: a ball, soft cloth, 5 blocks, 2 cars, 3 cups, 2 rattles, a box with lid, 3 spoons, 2 bowls, a squeaky toy, 2 wind-up toys, and a teddy bear.
- § Best used as a quick language assessment tool for children aged 3 through 5 years, but should not be used alone to obtain a thorough language evaluation.

COMMUNICATION

Receptive One-Word Picture Vocabulary Test (ROWPVT)

Test of English Hearing Vocabulary

Author: Morrison Gardner

Publisher: Academic-Therapy-Publications, 1985-2000..

2002 price data: \$140 per kit including manual, 25 record forms, and 170 full-color test plates.

Administration Time

10 to 15 minutes.

Scores

Standard scores, percentile ranks, age equivalency scores, normal curve equivalency scores and stanines .

Purpose

Designed to assess the English hearing vocabulary of individuals aged 2-0 through 18-11 years. The Spanish-Bilingual version is an assessment of the total acquired vocabulary of individuals aged 4-0 through 12-11 years.

Description of Instrument

The 170 test plates are presented in a spiral booklet with a flip-out easel and are ordered according to their difficulty level. Administration consists of presenting children a series of plates with four pictures per plate.

The examiner presents a word verbally to the child along with four full-color illustrations from the test plates. The child selects the illustration that best depicts the meaning of the word and responds by touching the picture or stating the number of the picture. Only a subset of the items, or critical range, is administered.

The critical range must be determined for each individual. The range begins with a series of items that are easy for the child and ends at a point where the responses are consistently incorrect. To establish the basal level, the examiner determines the chronological age and then starts either from that point or from a suggested starting point from the table in the manual.

Strengths:

- § The ROWPVT is sensitive to vocabulary growth associated with increased age in that the scores from the standardization sample systematically increased with age, and the correlation between age and raw score was .85.
- § Provides a practical, objective, and efficient starting point for the process of a comprehensive evaluation of language skills.
- § Manual presents interpretation of differences between scores from The Receptive One-Word Picture Vocabulary Test and scores from the Expressive One-Word Picture Vocabulary Test and gives the score differences required for significance at several levels of confidence. Appendix presents the frequency of occurrence of different discrepancy values with the norms group to determine if the difference is actually a clinical difference.
- § Speed and ease of test administration is quick (about 10 to 15 minutes).

Limitations:

- § The test is most often administered by speech-language pathologists, psychologists, counselors, learning specialists, and other personnel who are under the supervision of a professional familiar with assessment and interpretation.
- § The ROWPVT similar to other one-word vocabulary tests, is a very restrictive means of assessing vocabulary. By assessing vocabulary at the single word level, the role of context, language structure, and paralinguistic cues in word comprehension and use is minimized.

Tests of Early Language Development-Third Edition (TELD-III)

Screening Instrument

Authors: Wayne Hresko, Kim Reid, and Donald Hammill

Publisher: PRO-ED, Inc. , 1991

Price data 2003: \$272 per complete kit.

Administration Time

30 minutes.

Scores

Standard scores, percentile ranks, and age equivalency scores.

Purpose

Measure receptive and expressive aspects of language structures (i.e., syntax, morphology, and phonology) and meaningful language (i.e., semantics) of children ages 3 to 7-11 years.

Description of Instrument

The TELD was developed to meet the need for a quick, easily administered, oral language screening test. Test kit provides two forms, A and B. Each form has two subtests: Receptive language with 37 items (24 semantic and 13 syntax) and Expressive language with 39 items (22 semantic and 17 syntax). The picture cards are simple line drawings in black and white. Administration takes about 30 minutes. There is no time limit for a child's responses, however, it is recommended that all items be administered orally during a single session. The examiner shows the child one of the picture cards and says, "Show me the ----A." The child must point to the correct picture. Scoring criteria are specifically stated for each item.

Strengths:

- § The test has strong appeal for educators and clinicians desiring a useful and effective way of screening young children for potential language problems.
- § Test stimuli are colorful and attractive to children.
- § Test has high reliability and excellent validity (Content, Criterion-related, and Construct).
- § Extensive experience in test administration is not required for administering this measure.
- § Alternate forms are available.
- § Has been used as an outcome measure in the National Head Start Impact Study and the Family and Child Experiences (FACES) study.

Limitations:

- § Test is most effective with 4, 5, and 6 year old children but does have value at the 3 and 7 year age levels for identification of children who have difficulty in spoken language.
- § Some of the illustrations depict minority children, but most represent white, middle-class people and customs. This may be of concern when administering the test to children from low socioeconomic or minority backgrounds.

SOCIAL-EMOTIONAL

Ages & Stages Questionnaire: Social-Emotional

Screening instrument

Authors: Squires, J., Bricker, D., and Twombly, E.

Publisher: Brookes Publishing Co., 2003

2003 price data: \$125 per basic kit in English or Spanish including manual, 6 keys and 25 record forms.

Administration Time

10-15 minutes.

Scores

Total Scores are compared to cutoff scores .

Purpose

Screen children aged 6 to 60 months to assess social-emotional development in the Personal-Social domain: Self-regulation, Compliance, Communication, Adaptive Functioning, Autonomy, Affect, and Interaction with people.

Description of Instrument

Though the parent instrument, Ages & Stages Questionnaire, screens behavior in the Personal-social domain, questions in that instrument are limited to assessing solitary social play and play with toys and other children. Created in response to feedback for the Ages and Stages Questionnaire, this instrument focuses exclusively on the social and emotional behavior of children . There are eight color-coded questionnaires for use at 6, 12, 18, 24, 30, 36, 48, and 60 months. The questionnaire is to be completed by parent or most knowledgeable adult and takes about 10-15 minutes to complete. Every questionnaire has an easy-to-use form that allows professionals to score the test in 1-3 minutes and determine the child=s progress in that social-emotional area.

The reproducible scoring sheets all include referral considerations that help determine if the child needs further evaluation. Examples of questions on the 30-month questionnaire are:

Does your child look at you when you talk to him?

Does your child like to be hugged or cuddled?

Can your child settle himself down after periods of exciting activity?

Is your child interested in things around her, such as people, toys, and foods?

When you point at something, does your child look in the direction you are pointing?

Can your child move from one activity to the next with little difficulty, such as from playtime to mealtime?

Strengths:

§ Quick and easy administration .

§ Easy to score and interpret. Individual items are scored either 0 'Rarely or never'; 5 'Sometimes' or 'A concern'; or 10 'Most of the time'. Total scores are compared to cutoff scores to determine need for referral.

§ Parents and caregivers understand the items because they are written using common words that do not exceed a sixth-grade level. Additionally, quantitative descriptors (e.g., 15 minutes, within a 24-hour period) and concrete examples (e.g., kicks, bites other children) are included to aid in the interpretation of the items.

Limitations

§ A review of the psychometric properties of this instrument has not been published in *Buros Mental Measurement Yearbook* or *Test Critiques*, two reference sources for published tests. However, there is a Technical Report on ASQ-SE which can be downloaded for review from the publisher=s website.

§ As is true of all screening instruments, this measure would not be useful for programs to use as a pre/post test measure.

SOCIAL-EMOTIONAL

Devereux Early Childhood Assessment (DECA)

Screening Instrument

Observation-Based Behavior Rating

Authors: Paul A. LeBuffe and Jack A. Naglieri

Publisher: Kaplan Early Learning Company, 1999

2002 Price Data: \$199.95 per kit including 40 record forms, user=s guide, technical manual, classroom strategies guide, 20 parent guides, and classroom observation journal in a carrying case.

Administration Time

10 minutes

Scores

Standard scores, percentiles, and normal curve equivalents.

Purpose

Designed to screen for emotional and behavioral problems of children aged 2 to 5 years of age in the domains of Initiative, Self-Control, Attachment, and Behavioral Concerns.

Description of Instrument

The DECA is a four-subscale, 37-item observation-based behavior rating instrument intended to assess within-child protective factors in preschool children. This paper-and-pencil measure rates children=s behavior using parents, family caregivers, and/or early childhood professionals as raters. Test materials include a list of Head Start Performance Standards that are evaluated by the DECA. Behaviors are rated on a 5-point frequency scale by observers. Examples of test items include:

ADuring the past 4 weeks, how often did the child...

Act in a way that made adults smile or show interest in her/him?

Do things for himself/herself?

Choose to do a task that was challenging for her/him?

Strengths:

§ One of the few measures of child protective factors (resilience).

§ Requires minimal training and time (about 10 minutes) to administer and score.

§ Interpretation guidelines in test manual are thorough and provide an effective way to give constructive feedback to parents and practitioners.

§ Questions are designed to measure directly observable behaviors.

Limitations:

§ Test-retest correlations indicate that teachers were more consistent raters than parents. Scores obtained from parent raters require some caution in interpretation and use.

§ Many of the behaviors measured are more likely to occur in the preschool context.

§ Although the test authors state the items can be read to raters with limited English proficiency, they do not address the historic difficulty of obtaining reliable oral ratings, in general, or from those who speak minimal English, in particular

Social Skills Rating System

Parent Rating Form (Also has teacher and student rating form for elementary level)

Authors: Frank M. Gresham and Stephen N. Elliott

Publisher: American Guidance Service, 1990

2003 price data: \$218.99 per preschool/elementary level starter set including 10 copies of each form and level questionnaires, 10 assessment-intervention records, and manual.

Administration Time

10-25 minutes for each questionnaire.

Scores

Standard scores, percentile ranks, and behavior levels (fewer, average, and more).

Purpose

Designed to assess social skill behaviors (Cooperation, Empathy, Assertion, Responsibility, and Self-Control) and problem behaviors (Externalizing, Internalizing, and Hyperactivity) of children aged 3 to 18 years. Preschool level is appropriate for ages 3-0 to 4-11. Five-year old HIPPY children will take Kindergarten version.

Description of Instrument

The SSRS offers three methods of evaluating student social behavior: Parent Form, Teacher Form, and Student Form. The parent and teacher versions of the SSRS are divided into three developmental levels: preschool, kindergarten through sixth grade and seventh through twelfth grades. The student version is written at the third grade reading level. The number of items on the Parent Form range from 49 (Preschool Level) to 55 (Elementary Level). Each item asks parents to rate the frequency of a specified behavior: 0 Never; 1, Sometimes; or 2, Very often). Sample Items from the Parent Form, Preschool level include:

Cooperation Subscale (behaviors such as helping others, sharing materials, and complying with rules and directions.)

Puts away toys

Keeps room clean

Helps with tasks

Assertion Subscale (initiating behaviors, such as asking others for information, introducing oneself, and responding to the actions of others)

Makes friends

Shows interest in things

Responsibility Subscale (behaviors that demonstrate ability to communicate with adults and regard for property or work)

Questions unfair rules

Invites others home

Self-Control Subscale (Behaviors that emerge in conflict situations, such as responding appropriately to teasing, and in nonconflict situations that require taking turns and compromising)

Controls temper with you

Ends disagreements calmly

Attends to instructions

Strengths:

- § One of the few social skills rating scales that include a comprehensive parent report version.
- § The instrument has strong psychometric properties.
- § Items are written in behavioral terms so rater does not have to use a high level of inference in order to respond.
- § Quick administration (takes about 20 minutes) and easy to score.
- § Scores on the SSRS correlate highly with other somewhat similar measures (e.g. Social Behavior Assessment, Piers-Harris Children's Self-Concept Scale, and the Child Behavior Checklist).
- § Use of a multirater system provides information on how significant others in a child's life perceive

Limitations:

- § Each user should closely examine the normative sample described in test manual to decide if test is valid for their use. Normative sample underrepresented rural communities, the Northeast and West regions of the county, and Hispanic children. Also, parent sample contained proportionately too few parents with less than a high school education.

SOCIAL-EMOTIONAL

Vineland Adaptive Behavior Scales

Structured Parent Interview

Authors: Sparrow, S.S., Balla, D.A. and Cicchetti, D.V.

Publisher: American Guidance Service, 1984

2003 price data: \$189.99 for the complete kit.

Administration Time

20-60 minutes.

Scores

Standard scores, percentile ranks, age equivalency scores, and adaptive behavior composite (combination of four domains: Communication, Daily Living Skills, Socialization, and Motor Skills).

Purpose

Assesses personal and social adaptability of individuals from birth to adulthood.

Description of Instrument

The Vineland Adaptive Behavior Scales is a structured interview, conducted by a trained interviewer, with a respondent who is familiar with the subject's everyday behavior, i.e. daily activities required for personal and social sufficiency by interviewing the primary caregiver. There are three versions: Interview Edition, Survey Form with 297 items; Interview Edition, Expanded Form with 577 items; and the Classroom Edition, 244 items. Four Domains are measured: Communication (receptive and expressive language); Daily Living Skills (self-care activities of eating, dressing, washing, etc.); Socialization (interpersonal relations, play, and leisure); and Motor Skills (gross and fine coordination).

Strengths:

- § Administration and scoring are relatively simple. Examiner begins testing based upon child's chronological age. Manual sets forth criteria for scoring the items. Items are scored on a three-point scale (0 to 3) that indicates the extent to which the child performs the activity.
- § A Spanish version is available.
- § Computer scoring software assistance is available.
- § Training video is available.
- § Has been used in research studies and national program evaluations such as the National Longitudinal Study Children and Youth, Canada.

Limitations:

- § Interviewer qualifications include a graduate degree and specific training in early childhood assessment and interpretation. Interviewers should be thoroughly trained and have experience with open-ended interviewing techniques.
- § Despite having specified criteria for scoring items, there is an element of subjectivity in scoring items. However, an audio cassette tape providing sample interviews is available.

SOCIAL-EMOTIONAL

Vineland Social-Emotional Early Childhood Scales (Vineland SEEC Scales)

Semi-Structured Parent Interview

Authors: Sparrow, S.S., Balla, D.A., & Cicchetti, D.V.

Publisher: American Guidance Service, 1998

2003 price data: \$61.99 per complete kit including manual and 25 record forms.

Administration Time

15 to 25minutes.

Scores

Standard scores, percentile ranks with confidence intervals, and age equivalency scores.

Purpose

The Vineland SEEC Scales is designed to identify strengths and weaknesses in specific areas of social and emotional functioning of children from birth through 5 years 11 months of age. Assessment focuses on Interpersonal Relationships, Play and Leisure Time, and Coping Skills.

Description of Instrument

The Vineland SEEC Scales are a subset of items taken from the Socialization domain of the Vineland Adaptive Behavior Scale. Administration is through a semi-structured interview with a respondent who is familiar with the child=s behavior. The Interpersonal Relationships scale consists of 44 items that describe the child=s ability to communicate in social contexts, establish and maintain friendships, and appropriately recognize and express emotions. The Play and Leisure Time scale also consists of 44 items and is designed to assess how the child plays with toys, constructs make-believe activities and plays and shares with other children. The Coping Skills scale has 34 items and assesses the child=s ability to use manners, follow rules, engage in impulse control, and manage feelings. Ratings on all 122 items yield a Social-Emotional Composite score considered by the test authors to be the most statistically reliable score.

Strengths:

- § Instrument is well suited for both clinical and educational settings to assist in the screening and early identification of developmental delays and to chart developmental progress (pre- post testing).
- § Test administration is relatively quick (about 15 to 25 minutes)and scoring simple.
- § There are moderate correlations between the Vineland SEEC Scales and other developmental scales such as the Battelle Developmental Inventory.
- § English and Spanish language versions of the Parent Report are available.
- § Could be used in longitudinal studies that involve assessing the development of a child=s social or emotional skills.
- § Computerized scoring and interpretation are available.

Limitations:

- § Interviewer qualifications include a graduate degree and specific training in early childhood assessment and interpretation. Test administrator does not read test items, or allow the respondent to read test items but asks open-ended questions designed to elicit responses which requires the ability to monitor appropriately the course of the interview.

§ Norms from the Vineland SEEC are based upon modifications of the norms from the Vineland Adaptive Behavior Scale standardized in 1980, which is almost two decades old.

DEVELOPMENTAL SCREENING

Ages and Stages Questionnaire

Screening Instrument

Authors: Jane Squires, LaWanda Potter, and Diane Bricker

Publisher: Paul H. Brookes Publishing Co., 1995

2003 price data: \$190 per complete system including 11 reproducible questionnaires, 11 reproducible, age-appropriate scoring sheets, and user=s guide.

Administration Time

10-15 minutes.

Scores

Total scores are compared to cutoff points in the five domains to separate young children who require referral and assessment from those who do not.

Purpose

Screen to identify children aged 4 months to 60 months to determine need for further assessment.

Description of Instrument

The ASQ screening system is composed of 19 questionnaires with varying intervals of ages. Each question is divided into five different areas: Communication, Gross Motor, Fine Motor, Problem Solving, and Personal-Social. Each questionnaire contains 30 developmental items that are written in simple, straightforward language at grade levels ranging from 4th to 6th grade. Illustrations are provided to assist parents and caregivers in understanding the items. For the 30 developmental items, parents check *yes* to indicate that their child performs the behavior specified in the item, *sometimes* to indicate an occasional or emerging response from their child, or *not yet* to indicate that their child does not perform the behavior. Responses are converted to a point value which are totaled and compared to established screening cutoff points. Sample items from the 60 month 5 Year Questionnaire include:

Communication Domain:

Does your child use four- and five-word sentences? For example, does your child say, "I want the car"?

Gross Motor Domain:

Does your child catch a large ball with both hands?

Fine Motor Domain:

Ask your child to trace on the line below with a pencil. Does your child trace on the line without going off the line more than two times? Mark "sometimes" if your child goes off the line three times.

Problem Solving Domain::

Does your child count up to 15 without making mistakes? If so, mark "yes". If your child counts to 12 without making mistakes, mark "sometimes".

Personal-Social Domain:

Does your child wash her hands and face with soap and water and dry off with a towel without help?

Strengths:

- § The ASQ is an easily understood instrument (4th to 6th grade reading level) that is quick to administer and easy to score.
- § There is a Spanish version of the instrument.
- § Test administrators do not need to have specialized course work or training in test administration, however, the test manual emphasizes the importance of consistency when administering the instrument.
- § The ASQ has been used as a developmental screener in HIPPIY programs in the states of Nevada and Florida, Healthy Families America, and Early Head Start.

Limitations

- § Though the ASQ can provide information on a child=s developmental progress and can serve as a screener for the early detection of developmental problems, it has limited usefulness for measuring child progress over time (pre and post-test measure) or for comparing groups of children.
- § May be difficult for parent to complete on their own.

DEVELOPMENTAL SCREENING

AGS Early Screening Profiles

Screening instrument

Authors: Patti Harrison, Slan Kaufman, Nadeen Kaufman, Robert Bruininks, John Rynders, Steven Illmer, Sara Sparrow and Domenic Cicchetti

Publisher: American Guidance Service, 1990

2003 price data: \$350 per complete kit including test plates in easel, 25 test records, 25 self-help social profile questionnaires, sample home-health history survey, 25 score summaries, tape measure, beads and string.

Administration Time

15 to 30 minutes.

Scores

Normal Curve Equivalents, age equivalency scores and stanines. Profile scores are converted into Screening indexes, which correspond to standard deviation units on the normal curve to determine which children require further assessment.

Purpose

To screen cognitive, language, motor, self-help, and social development of children aged 2 through 6 years. This test also surveys the child=s articulation, home environment, health history, and test behavior.

Description of Instrument

The three basic components, called Profiles, are supplemented by four Surveys. All of the Profiles and Surveys can be administered or they may be used separately or in various combinations. The **Cognitive-Language Profile** is administered individually to the child and assesses reasoning skills, visual organization and discrimination, receptive and expressive vocabulary, and basic school skills. The profile can be separated into cognitive (nonverbal) and language (verbal) subscales. The **Motor Profile**, also individually administered, assesses both gross and fine motor skills, such as walking a straight line, imitating arm and leg movements, tracing mazes, stringing beads, and drawing. The **Self-Help/Social Profile** is a questionnaire completed by the child=s parent, caregiver, or a combination of them. It assesses the child=s typical performance in the areas of communication, daily living skills, socialization, and motor skills.

The **Articulation Survey** measures the child=s ability to pronounce 20 words selected to test articulation problems in the initial, medial, and final positions of words. The **Home Survey** is completed by the parent and asks nonintrusive questions about the child's home environment such as types of play materials, content and frequency of parent-child interactions (i.e. reading), and responsibilities given to the child. The **Health History Survey**, also completed by the child's parent, is a brief checklist of any health problems the child has had. The fourth Survey, the **Behavior Survey** is used by the examiner to rate the child's behavior during administration of the Cognitive/Language and Motor Profiles. The child is rated in categories such as attention span, frustration tolerance, and response style.

Strengths:

- § Administration is relatively quick. Profiles can be administered in less than 30 minutes and the Surveys require an additional 15-20 minutes.
- § Publisher indicates paraprofessionals can administer the Profiles and Surveys.
- § The Cognitive-Language Profile can be administered nonverbally which would be useful for screening children with limited English proficiency, language difficulties, or hearing problems.
- § There is an AGS Early Screening Profiles training video that can be purchased from the publisher.

Limitations:

- § Predictive validity for the Behavior, Motor and Home subscales is rather low.

- § Only four ratings are obtained on the Behavior Survey (activity level, attention span, cooperativeness, and independence), fewer than would be expected from a behavioral measure.
- § Scores have relatively low agreement with scores from other screening instruments such as the DIAL-R and Battelle.
- § It has limited usefulness for measuring child progress over time (pre and post-test measure) or for comparing groups of children.

DEVELOPMENTAL SCREENING

Battelle Developmental Inventory Screening Test

Screening Instrument

Authors: Jean Newborg, John Stock, Linda Wnek, John Guidubaldi, and John Svinicki

Publisher: Riverside Publishing, 1984

2003 price data: \$150 per screening test including 30 test booklets, test item book, and examiner=s manual

Administration Time

Approximately 30 minutes.

Scores

Total scores are compared to cutoff scores; raw scores are converted to age equivalency scores.

Purpose

Can be used as a general screening measure, preliminary assessment or initial identification of possible developmental strengths and weaknesses for preschool and kindergarten children from birth to eight years old.

Description of Instrument

The BDI Screening Test, part of the complete Battelle Developmental Inventory, is a behaviorally based measure that screens development in the following domains:

Personal-Social Domain: Adult Interaction, Expression of Feelings/Affect, Self-Concept, Peer Interaction, Coping, and Social Role.

Adaptive Domain: Attention, Eating, Dressing, Personal Responsibility, and Toileting.

Motor Domain: Muscle Control, Body Coordination, Locomotion, Fine Muscle, and Perceptual Motor

Communication Domain: Receptive Communication and Expressive Communication

Cognitive Domain: Perceptual Discrimination, Memory, Reasoning and Academic Skills, and Conceptual Development.

Testing procedures for the BDI Screening Test include using a combination of direct observation and interview assessment methods. The Screening Test for the BDI consists of 96 items: 20 items each from the Personal-Social, Adaptive, and Motor Domains, and 18 items each from the Communication and Cognitive Domains. The number of items actually administered to an individual child will rarely be more than six per domain. Examples of items in the Personal-Social Domain for a 60-71 month old child are:

Adult Interaction: The child uses adults other than parents as resources.

Expression of Feelings/Affect: The child shows a positive attitude toward school.

Self-Concept: The child performs for others.

Peer Interaction: The child participates in competitive play activities.

Coping: The child waits his or her turn for a teacher=s or adult=s attention.

Social Role: The child recognizes the feelings of others.

Strengths:

- § Age-Equivalency scores (the age at which a raw score is average) are provided in table form which is useful for making comparisons about a child=s relative performance compared to an age group.
- § There are multiple cut-off points that correspond to 1.00, 1.28, 1.50, 1.65, and 2.33 standard deviations below the mean for each age group. Multiple cutoff points are useful because the degree of handicap required to become eligible for special services varies among states and often within a state.
- § Has been used as one of the outcome measures in the Early Head Start Evaluation.

Limitations:

- § Effective administration requires a trained test examiner. Items must be administered in a specified order and basal and ceiling rules must be applied. It is recommended test examiner conduct several practice administrations with normally developing children before actually administering the test.
- § Clinical judgement in interviewing is necessary for acquiring reliable parent information.
- § Black and Hispanic children were included in normative sample, but not Asian or Native American children.
- § It has limited usefulness for measuring child progress over time (pre and post-test measure) or for comparing groups of children.

DEVELOPMENTAL SCREENING

The Brigance Screens

Screening Instrument

Author: Frances P. Glascoe

Publisher: Curriculum Associates, Inc., 1978-1991

2003 price data: \$124 per manual with tests; \$19.90 per 10 record books.

Administration Time

15 minutes.

Scores

Total scores are compared to a recommended cutoff score of 60.

Purpose

To identify children who need further assessment to determine their eligibility for special programs, stimulation, or extra assistance.

Description of Instrument

The Brigance Screens consist of four separate screening instruments: Brigance Infant and Toddler Screen, Brigance Early Preschool Screen (ages 2 to 2 2), Brigance Preschool Screen (ages 3 and 4), and the Brigance K & I Screen (ages 5 and 6). Each of these four screening instruments have items drawn from the 1982 edition of the Brigance Diagnostic Inventory of Early Development, a complete assessment battery assessing language, motor, cognitive, self-help, academic, and readiness skills.

The Preschool Screen is an individually administered instrument that assesses development in eleven broad skill areas: Personal Data Response, Identifies Body Parts, Gross Motor Skills, Gross Motor Skills, Identifies Objects, Repeats Sentences, Visual Motor Skills, Number Concepts, Builds Tower with Blocks, Matches Colors, Picture Vocabulary, and Plurals.

Each basic skill is regarded as distinct, and there is no cumulative score associated with a skill area or subarea. A score does not represent how well the child performed compared to other children of the same age. The score is also not a simple percentage, but is a summation of weighted raw scores. The PBS materials suggest that scores can be rank ordered and the lowest scoring children referred for evaluation. Also, cutoff scores can be determined. Test author recommends that any child scoring below 60 be referred for further assessment.

Strengths:

- § Test is quick (10 to 20 minutes) and easy to administer by paraprofessionals with limited background and training in test administration.
- § There is a Spanish version of the instrument.
- § The Brigance Screens include many of the same items that are commonly found in more comprehensive psychoeducational assessment batteries used with preschool children.
- § A 20-minute training video covering administration and scoring practice examples is available.
- § Test materials include a Developmental Record Book which facilitates record keeping for purposes of multiple testing because it is an ordered listing of all basic skills and behaviors.
- § A scoring software CD-ROM can be purchased from the publisher that creates student lists with chronological age, cut-off scores, growth indicators, percentiles, and age equivalents.

Limitations:

- § The Brigance Preschool Screen is appropriate for three and four year old children. A second Brigance screening instrument, Brigance K & I, will need to be purchased for use with children aged five.
- § This instrument is of limited usefulness for measuring child progress over time or comparing groups of children.

DEVELOPMENTAL SCREENING**Denver II**

Screening Instrument

Authors: Frankenburg, W., Dodds, J., Archer, P., Bresnick, B., Maschka, P., Edelman, N. and Shapiro, H.

Publisher: Denver Developmental Materials, Inc., 1967-1990

2003 price data: \$88 per complete package (English) and \$118 per complete package (Spanish) including test forms, screening manual, technical manual, and the materials required for screening.

Administration Time

25 minutes.

Scores

Each item is scored as pass, fail, no opportunity or refusal, and the performance is interpreted as normal, suspect, or untestable. Each of the 125 items is represented on the form by a bar that spans the ages at which 25%, 50%, 75%, and 90% of the standardization sample passed.

Purpose

The Denver II is designed to screen for developmental delays in children from birth to six years old.

Description of Instrument

The Denver II is a revision of the Denver Developmental Screening Test, a commonly used developmental screening test. Like its predecessor, the Denver II is an individually administered developmental screening test. There are 125 items organized into four domains of development: Personal-Social, Fine Motor-Adaptive, Language, and Gross Motor. Items are arranged in approximate chronological order according to the ages at which most children can be expected to perform them. A special feature of the test is the provision of age norms for each of the 125 items that are represented on the form by a bar that spans the ages at which 25%, 50%, and 90% of the standardization sample passed that item. This provides a comparison of a child's performance to other children of the same age on each of the test items. There is a scale for rating the child's behavioral characteristics during the test such as compliance, alertness, fearfulness, attention span, and typical behavior.

An Age Scale is used to determine the starting point of the test. The child's age is calculated and shown as a line on the test form, and each item that is intersected by the age line determines the starting point of the test.

Strengths:

- § Ease of administration (about 25 minutes) and scoring (Child's responses are scored as Pass, Fair, No Opportunity, and Refusal)
- § Manual recommends that test administrators receive training in the administration of the instrument before using the test.
- § The capability of comparing a child's performance to same age peers for each of the 125 items is an added feature of this test.

Limitations:

- § Test is a reliable screening instrument provided it is administered by trained examiners who follow the test protocols provided in the technical manual, a limitation which would apply to all individually administered instruments.

- § The interpretation of a child's development (normal, suspect, or untestable) is based on the number of Delay and Caution scores rather than from established norms.

DEVELOPMENTAL SCREENING

Developmental Indicators for the Assessment of Learning - Third Edition - DIAL-3

Screening Instrument

Authors: Mardell-Czudnowski, C and Goldenberg, D.

Publisher: American Guidance Service Inc., 1983-1998

2003 price data: \$410 per complete kit (English or Spanish) including manual, 50 record forms (English), 1 record form (Spanish), 50 cutting cards, 50 parent questionnaires (English), manipulatives, dials, operator's handbooks in English and Spanish for motor, concepts and language areas, Speed DIAL, and training packet.

Administration Time

20 to 30 minutes; **Speed DIAL** is 15 to 20 minutes.

Scores

Standard deviation and percentile cutoff points by chronological age at two-month intervals for total and area scores.

Purpose

Used to identify young children aged three through six years who may be in need of further diagnostic assessment.

Description of Instrument

The DIAL-3 is an individually administered instrument that assesses five developmental areas: Motor, Concepts, Language, Self-Help, and Social Skills while the short form of the test, Speed DIAL, measures Motor Skills, Concepts, and Language. The DIAL-3 Parent Questionnaire provides normed scores for the child's self-help and social skills as well as information on personal-medical information, and overall development. Test materials include large cardstock wheels that are rotated to reveal picture prompts (one at a time). The test name derives from these wheels or Adials®.

Though the DIAL-3 is designed as an individual screener, it is administered in a setting in which three children at a time are observed in a specially designed screening area staffed by several members of an administration team. This team consists of a coordinator (with an early childhood or special education background) and three observers that staff three stations in the screening area. Administration requires children move through the three stations and, at each station, to respond to questions or perform tasks related to one of the areas assessed. As the child progresses from station to station, trained observers are recording behaviors that may indicate the need for further assessment in social, affective, or speech-language areas. Scores are reported as either Potential Delay or OK. At the same time the child is being assessed, the parent completes the Parent Questionnaire on Self-Help Development and Social Development.

Strengths:

- § Has abbreviated form, Speed DIAL, which takes about 15 minutes to administer.
- § Screening of program participants could be accomplished in a relatively short time compared to individually testing each child.
- § There is a simplified scoring procedure. Raw score conversions are printed directly on the scoring forms where the child's performance is recorded.
- § Videotaped instructional materials are available for training screening program personnel and self tests are included for assessing the competence of those who will be administering the screening. Questions are also suggested that could be used for a program to evaluate the screening program at its site.
- § Has been used in HIPPIY programs in the states of Colorado and Louisiana where program staff can be consulted regarding their experiences with the instrument.

Limitations:

- § The administration manual does not provide guidance regarding how a screening coordinator would select the cutoff points for Refer or Do not refer decisions.

§ Caution is suggested when interpreting scores using the norms for Spanish-speaking children.

DEVELOPMENTAL SCREENING

Early Screening Inventory-Revised (ESI-R)

Screening Instrument

Authors: Meisels, S., Marsden, D., Wiske, M., and Henderson, L.

Publisher: Pearson Early Learning, 1976-1997

2003 price data: \$115 for ESI-P kit and \$115 for ESI-K including examiner=s manual, Parent Questionnaires, Score sheets, and screening materials.

Administration Time

15 to 20 minutes.

Scores

Total scores are reported as AOK@, ARescreen@ (1 to 2 standard deviations below mean) or ARefer@ (greater than 2 standard deviations below the mean).

Purpose

The appropriate use of the ESI-R is to identify the possibility of a learning or handicapping condition that might affect a child=s overall potential for success in school.

Description of Instrument

The Early Screening Inventory-Revised was developed from items selected from the Eliot-Pearson Screening Inventory (EPSI). The instrument is available for two age groups: The Early Screening Inventory -Preschool (ESI-P) is for children aged 3 to 4 2 years and the Early Screening Inventory-Kindergarten (ESI-K) covers children aged 4 2 to 6 years. The ESI is divided into four sections: 1) Initial Screening Items (a Draw-a-Person Task and Letter Writing) , 2) Visual-Motor-Adaptive domain, 3) Language and Cognition domain, and 4) Gross Motor/Body Awareness domain. The **Visual-Motor Adaptive** assesses fine motor skills, eye-hand coordination and the ability to remember visual sequences, to draw two-dimensional visual forms, and to reproduce three-dimensional visual structures. **Language and Cognition** items focus on language comprehension and verbal expression, the ability to reason and count, and the ability to remember auditory sequences. **Gross Motor/Body Awareness** items examine balance, large motor coordination used in hopping and skipping, and the ability to imitate body positions from visual cues. A separate informal Parent Questionnaire accompanies the ESI. Its purpose is to collect relevant case history information about the child=s family background, school attendance, prenatal and postnatal medical problems, current health, and developmental skills. Rather than a formal score, the Parent Questionnaire is intended to give perspective and possibly some insights into the etiology of a child=s presenting problem or alert examiners to conditions placing a child at risk for normal development.

Strengths:

- § Test administration is quick (about 15 to 20 minutes) and scoring interpretation is clearly described in the manual.
- § A Spanish version of the instrument is available.
- § There are two training videos available from the publisher. The first video provides an overview of what developmental screening is, what it is not, and issues related to the presence of parents during screening. The second video provides the trainee an opportunity to view how an examiner administers the test to four different children who are between the ages of three and six years.
- § Has been used by HIPPIY programs in Colorado.

Limitations:

- § Two instruments will be needed to assess children from age three to five years. The Early Screening Inventory - Preschool (ESI-P) is appropriate for children ages 3 to 4 2 years and the Early Screening Inventory-Kindergarten (ESI-K) covers children ages 4 2 to 6 years.
- § The instrument does not include a social-emotional domain, however, there may be some items on the Parent Questionnaire that cover this area.

§ The instrument should be administered only by examiners knowledgeable in early childhood behavior and development who have received supervised practice. Some items such as Draw-A-Man test require scoring judgments.

PARENTING KNOWLEDGE AND SKILLS

The Adult-Adolescent Parenting Inventory (AAPI-2)

Questionnaire

Author: Stephen J. Bavolek

Publisher: Family Nurturing Centers, Int.

2003 price data: \$122 for AAPI-2 kit in English or Spanish including Pre and Post Test forms, stencils, handbook and worksheets

Administration Time

Parents can complete instrument in approximately 20 minutes.

Scores

Standard scores.

Purpose:

Designed to assess the parenting attitudes and child rearing practices of adult parent and adolescent parent individuals. It has been used to identify high-risk child rearing and parenting practices that could lead to physical or emotional abuse, or neglect of children and to assess changes in parenting and child rearing practices after treatment.

Description of Instrument:

The AAPI-2 is used to determine the degree to which respondents agree or disagree with parenting behaviors and attitudes known to contribute to child abuse and neglect. There are 40 test items that divided into five separate subscales:

Inappropriate Parental Expectations

Parental Lack of Empathy Towards Children=s Needs

Strong Belief in the Use of Corporal Punishment as a Means of Discipline

Reversing Parent-Child Role Responsibilities

Oppressing Children=s Power and Responsibilities

Responses are recorded by the parent on a 5 point Likert scale from Strongly Agree to Strongly Disagree.

The five subscales total raw scores are calculated by adding the numerical values. Raw score totals for each subscale are converted into standard scores forming the norm tables of the AAPI. Normative data is provided by age (adult/adolescent), and gender (male/female). The standard scores are plotted on the AAPI Parenting Profile providing an index of risk for abusive and/or neglecting behaviors.

Sample test items:

Children should keep their feelings to themselves.

A certain amount of fear is necessary for children to respect their parents.

Parents who nurture themselves make better parents.

It=s OK to spank as a last resort.

Strengths:

- § Takes approximately 20 minutes to administer and can be administered in a small group setting or individually.
- § Items are written at the fifth grade reading level and can be administered orally to non-readers.
- § There is a Spanish version of the instrument.
- § There are two forms of the AAPI-2 - Form A and Form B.
- § Computer scoring is available on a CD-Rom which automatically scores and provides printed results.
- § Technical report of test reported that research studies that employed the AAPI-2 as a pre/post test measure reported parents who completed their parenting programs had more positive parenting attitudes and behaviors than those who dropped out of the parenting program.
- § Test has been used for several years by HIPPO programs in Nevada.

Limitations:

- § The AAPI-2 was not reviewed in the Buros Mental Measurements Yearbook. However, research summaries of several evaluation studies conducted by the author indicate the instrument has adequate psychometric properties for use as a Pre/Post test measure.
- § Some of the test items require parents to disclose information that they may consider sensitive.

PARENTING KNOWLEDGE AND SKILLS

Parent-Child Relationship Inventory

Questionnaire

Author: Anthony Gerard

Publisher: Western Psychological Services, 1994

2003 price data: \$106 per complete kit including manual, 25 answer sheets, and 2 prepaid mail-in answer sheets for computer scoring and interpretation.

Administration Time

Parent can complete instrument in approximately 15 minutes.

Scores

Standard scores (T-scores) and percentiles. Separate norms are provided for mothers and fathers.

Purpose

To assess parents= attitudes toward the task of parenting and how they feel about their children.

Description of Instrument

This is a 78-item self-report measure designed for use with mothers or fathers of 3- to 15-year old children. It is to be completed by the parents for only one child in the family at a time in which parents (mothers or fathers) respond to each item using a 4-point Likert scale. The items are grouped into seven content scales: Parent Support, Satisfaction with Parenting, Involvement, Communication, Limit Setting, Autonomy, and the Role Orientation scale. There are five Social Desirability items and 10 pairs of highly correlated items to assess response inconsistency. The measure can be used in both research and clinical settings to identify specific areas of difficulty between parents and their children.

Strengths:

- § Items are written at the fourth grade level and takes approximately 15 minutes to complete through either individual or group administration.
- § There is a Spanish version available.
- § Raw scores are converted to scaled scores (T-scores and percentiles) and separate norms are provided for mothers and fathers.
- § Publisher offers an option to purchase a microcomputer scoring and report program. The microcomputer scoring program is easy to install and use and is compatible on any system running Microsoft Windows 3.1 or above. The computerized report produced from the microcomputer scoring program can be a useful aid in interpretation of scores. Computer scoring (raw data is mailed in and interpreted by publisher) is also available through Western Psychological Services as well as FAX service scoring.

Limitations

- § Families with multiple children need to complete a test for each child.
- § There is a maximum number of missing responses that can be present without jeopardizing the test=s validity.
- § Norming sample was better educated and less diverse than the U.S. population as a whole. Young parents, Blacks, Hispanics, and less educated parents were under represented..

PARENTING KNOWLEDGE AND SKILLS

Parenting Stress Index - Third Edition

Screening Instrument

Authors: Richard Abidin, PAR staff, Noriel Ona

Publisher: Psychological Assessment Resources, 1983-1995

2003 price data: \$139 per complete kit 10 reusable item booklets, 25 hand-scorable answer sheet-profile forms, and manual.

Administration Time

Parent can complete the instrument in approximately 20 minutes.

Scores

Percentiles, percentile ranks, and a Total Stress score (the sum of the Parent and Child domain scores). The Profile analysis can be plotted on the scoring form.

Purpose

Though titled Parenting Stress Index, this instrument is a measure of overall child and/or parent maladjustment rather than stress per se. Results are useful in guiding decisions regarding whether or not professional intervention might be necessary. It has been used as an outcome measure to show the effects of various stress-reduction interventions.

Description of Instrument

This is a parent self-report instrument that measures three source domains of stress: those sources associated with the child=s characteristics, those associated with parent characteristics, and life event stressors. The child domain of the instrument was designed to measure a parent=s perceptions of four child temperament characteristics: Adaptability, Mood, Demandingness, and Distractibility-Hyperactivity. The parent domain of the instrument measures three [Personality-Pathology](#) variables: Sense of Competence in parenting role, Depression, and Parental Attachment. The other four parent stress variables are situational in nature: Relationship with Spouse, Social Isolation, Parental Health, and Role Restriction. The Life Stress scale measures stress due to unforeseen major life events: death of a family member, loss of job, divorce, etc.

Subscale scores in the child domain are added together to form a Child Domain score, and subscale scores in the parent domain add together to form a Parent Domain score. A Total Stress score is the sum of the Parent and Child domain scores. The Life Stress scale does not contribute to the Total Stress score. There are percentile norms for Total Stress and the domain scores. The norms are broken down by age of target child. No standard scores are reported, but raw scores are converted to percentile ranks; percentiles and a profile analysis can be plotted on the scoring form.

Strengths:

- § Because so few instruments exist that measure family and child adjustment, this instrument has been utilized extensively as a screening instrument during the initial interview process to determine the need for individual or family counseling.
- § A computer software scoring program is available.
- § There is a Spanish version which was normed on an Hispanic sample.

Limitations

- § Problems include poor support for subtest organization and profile analyses. It is recommended that global scores be used for screening purposes only.

HOME ENVIRONMENT

Home Observation for Measurement of the Environment (HOME)

Observation measure accompanied by a parent interview

Authors: Bettye M. Caldwell and Robert Bradley

Publishers: Bettye M. Caldwell and Robert H. Bradley, University of Arkansas for Medical Sciences and the University of Arkansas at Little Rock 1978 - 1984

2003 price data: \$30 per administration manual; \$9.00 for 50 scoring forms.

Purpose

The HOME is designed as a screening instrument to measure the quantity and quality of the types of stimulation and support available in the child=s home.

Description of Instrument

The HOME is intended to be an environmental Aprocess@ measure that reflects ongoing patterns of nurturance and stimulation provided to children in their homes. It was intended by its authors to be a more useful alternative to Astatus@ indices such as demographic attributes of parents (education, occupation, socioeconomic status, etc.) Two-thirds of the instrument involves observer assessment of parent-child interaction and one-third involves a parent (or primary caregiver) interview. There are two inventories: one for Families of Infants and Toddlers (birth to age 3) with six subscales consisting of 45 items: Emotional and Verbal Responsibility of Parent, Acceptance of Child=s Behavior, Organization of Physical and Temporal Environment, Provision of Appropriate Play Materials, Parent Involvement with Child, and Opportunities for Variety in Daily Stimulation

The second form for Families of Preschool Children, aged three to six that has eight subscales and 55 items: Learning Stimulation, Language Stimulation, Physical Environment, Pride, Affection and Warmth, Academic Stimulation, Modeling and Encouragement of Social Maturity, Variety in Experience, and Acceptance/Physical Punishment.

As a process measure, HOME involves direct observation of children=s home environments and, to a lesser degree, parent (or primary caregiver) report. Items are scored Ayes@ or Ano@ reflecting the occurrence or nonoccurrence of particular maternal behaviors, presence or absence of types of play materials, or types of in-home or out-of-home activities in which the child has or has not participated. The instrument needs to be administered during a home visit while the child is awake in order that the observer-interviewer can see the interaction between the child and the primary caregiver. The time required for the home visit is about one hour. Detailed suggestions are provided for conducting the interview.

Strengths:

- § Validation studies show the HOME correlates with SES measures to a moderate degree as well as early measures of cognitive development.
- § HOME has been used in several national program evaluations: Hawaii Healthy Start Program Evaluation, Nurse Home Visitation Program, Multisite Evaluation of the Parents as Teachers Home Visiting Program conducted by SRI.
- § Some of the HOME subscales might be useful in measuring the extent to which HIPPY has contributed to ACreating an educational milieu in the home that encourages literacy@, which is one of the HIPPY program goals.

Limitations:

§ Test examiners need practice with the instrument to develop the level of observational skills necessary to assure reliable scoring.

GLOSSARY MEASUREMENT AND ASSESSMENT TERMS

Ability

A characteristic indicative of an individual's competence in a particular field. The word Ability is frequently used interchangeably with aptitude.

Age-equivalent scores

A score that is derived by determining the average score obtained on a test by children of various ages. Age equivalency scores provide information that is easily understood, place the obtained score in a developmental context, and provide a comparison of the relative performance of two groups of children.

Age norms

The distribution of test scores by age of test takers. For example, a norms table may be provided for children aged five. This age-norms table would present such information as the percentage of five year olds who score below each raw score on the test.

Aptitude test

A measure of abilities that are assumed to be relevant to future performance in a specific type of skill or area of achievement.

Assessment

An ongoing process using both formal and informal procedures by appropriately qualified personnel to determine the child's unique strengths and needs and the services appropriate to meet those needs. Formal procedures for ongoing assessment may include the use of published developmental profiles or checklists; health and medical tests and procedures, and/or structured observations. Informal procedures include conversations with parents/caregivers or informal observations of the children in their daily routines.

Checklist

A specific list of skills the child is judged as having accomplished or not. A checklist may be completed by the parent or teacher through observation of the skills being performed or based on their knowledge of the child.

Confidence Intervals (also called Confidence Range or Confidence Bands)

A band or range of scores that has a high probability of containing a child's true score. A statistical term representing an upper and lower value that are likely to contain the actual or true score. The standard error of measurement provides the basis for forming the confidence interval. The band may be large or small depending on the degree of certainty (confidence) desired. For example, a 95% confidence interval can be thought of as the range in which a child's true score would be found 95 percent of the time.

Correlation

The degree of relationship between two measured variables. A correlation of 0.00 denotes a complete absence of relationship. A correlation of plus or minus 1.00 indicates a perfect (positive or negative) relationship. Correlation coefficients are used in estimating test reliability and validity.

Criterion-referenced instrument

An instrument that is used to determine if a child has achieved mastery in a particular domain. The child's behavior is measured in relation to a specific behavior, rather than to a normative group. The focus is on what the child knows or can do, not on how they compare to others.

Cutoff score

A number that identifies the point for pass/fail for each domain for each chronological age. Scores falling above the cutoff point mean the child is progressing as expected for his/her developmental age. Scores below the cutoff point mean a child may need further diagnostic assessment.

Derived score

A transformation of a raw score (e.g., age equivalents) to reveal the individual's performance relative to a norming group.

Developmental Screening Instrument

An instrument that is used to determine if developmental skills are progressing as expected, or if there is a cause for concern and further evaluation is necessary. Screening for sensory, behavioral, or developmental concerns is a first step and results from a screening does not lead to a decision about whether a child has a developmental problem.

Diagnostic Test

A test used to diagnose or analyze or locate an individual's specific areas of weakness or strength, to determine the nature of his/her weaknesses or deficiencies, and, if possible, to suggest their cause. Diagnostic achievement tests are commonly used for the skill subjects.

Domain-referenced instrument

An instrument that includes a random sample of items drawn from an item pool that is representative of all possible test items for a well-defined content area.

Equivalent forms

Any of two or more forms of a test that are closely parallel with respect to content and the number and difficulty of the items included. Equivalent forms should also yield very similar average scores and measures of variability for a given group. Also called parallel or alternate forms.

Grade equivalency scores

A norm-referenced score; the grade and month of the school year for which a given score is the actual or estimated average. A grade equivalent is based on a 10-month school year. For example, if a student scores at the average of all fifth graders tested in the first month of the school year, he/she would obtain a G.E. of 5.1.

Grade norms

The distribution of test scores by the grade of the test takers.

Individual-referenced measurement

Involves comparing an individual's performance on a test at one point in time to that individual's performance on the same test at another point in time.

Mastery level

The cutoff score on a criterion-referenced or mastery test. People who score at or above the cutoff score are considered to have mastered the material. A mastery is an arbitrary judgment because cutoff scores can be determined by several different methods, and each method results in a different cutoff score.

Mean

The arithmetic average of a set of scores.

Median

The middle score in a distribution or set of ranked scores; the point (score) that divides a group into two equal parts; the 50th percentile.

Mode

The score or value that occurs most frequently in a distribution.

Normal curve equivalents

Normalized standard scores with a mean of 50 and a standard deviation of 21.06. The standard deviation of 21.06 was chosen so that NCE=s of 1 and 99 are equivalent to percentiles of 1 and 99. There are approximately 11 NCE=s to each stanine.

Norms

The distribution of test scores of some specified group called the norm group. Norms are not standards. Norms are indicators of what students with similar characteristics did when answering the same test items as those taken by students in the norm group. Norm allow you to compare children of the same age with children in a national sample. Test manuals have test norms for each raw score.

Norm-referenced measurement

Any test in which the score acquires additional meaning by comparing it to the scores of people in an identified norm group. A test can be both norm- and criterion-referenced. Most standardized achievement tests are referred to as norm-referenced.

Normal curve equivalency score

A type of standard score with a distribution that has a mean of 50 and a standard deviation of 21.06

Observation

Collection information by observing the child and not directly requesting that the child perform specific tasks.

Percentile

A point on the norms distribution below which a certain percentage of the scores fall. For example, if 70% of the scores fall below a raw score of 56, then the score of 56 is at the 70th percentile.

Percentile ranks

The percentage of scores falling below a certain point on a score distribution. (Percentile and Percentile Rank are sometimes used interchangeably.)

Pretest

A measure that is administered prior to an intervention.

Posttest

A measure that is administered following an intervention or participation in a program in order to determine the effects of the intervention or program.

Profile

A graphic presentation of several scores expressed in comparable units of measurement for an individual or a group. This method of presentation presents easy identification of relative strengths or weaknesses across different tests or subtests.

Psychometric properties

The statistical properties of an instrument such as the reliability and validity indices.

Questionnaire

A measure that presents a set of written questions to which all individuals in a sample respond.

Range

A measure of the amount of dispersion in a distribution of scores. It is expressed as the lowest and highest scores in the distribution.

Raw score

A person's observed score on a test, i.e. the number correct. Raw scores should not be used to make comparisons between performance on different tests, unless other information about the characteristics of the tests is known. For example, if a child answered 24 items correctly on a reading test, and 40 items on a mathematics test, it should not be assumed the child is better in math than reading.

Reliability

The extent to which test scores are consistent; the degree to which the test scores are dependable or relatively free from random errors of measurement. Stability is the key reliability issue for preschool children because of the appreciable fluctuation in the test behaviors of young children. Reliability is determined by statistical analysis. No test is 100% reliable due to measurement error. The reliability of a test is improved when the testing conditions remain uniform - the same environment, testing conditions, how instructions are presented, the materials used, etc.

Response bias

A predisposition to give the same type of answer to some or all the items in a test rather than an answer to each item based on careful consideration of that item's content.

Scaled score

A mathematical transformation of a raw score. Scaled scores are useful when comparing test results over time. Most standardized achievement test batteries provide scaled scores for such purposes. Several different methods of scaling exist, but each is intended to provide a continuous score scale across the different forms and levels of a test series.

Standard age scores

Normalized standard scores provided for specified age groups on each battery of a test. Typically, standard age scores have a mean of 100 and a standard deviation of 15.

Standard deviation

A measure of the variability, or dispersion, of a distribution of scores. The more the scores cluster around the mean, the smaller the standard deviation. In a normal distribution of scores, 68.3% of the scores are within the range of one S.D. below the mean and one S.D. above the mean.

Standard score

A general term referring to scores that have been transformed for reasons of convenience, comparability, ease of interpretation, etc. and have a given mean and standard deviation. They express how far an examinee's score lies from the mean of the distribution in terms of the standard deviation.

Standardization

Standardization refers to the uniformity of procedure in administering and scoring the test. This is the process the test developer uses to choose the test items or questions and the conditions under which the test

should be administered (i.e., verbal instructions to the child, if and how the test administrator can demonstrate a task, how many times the individual can attempt the task, etc.)

Stanines

A normalized standard score that has a mean of 5 and a standard deviation of 2. Only the integers 1 to 9 occur. Sometimes, the first three stanines are interpreted as being *below average*, the next three as *average* and the top three stanines as *above average*.

T-score

A standard score that has a mean of 50 and a standard deviation of 10.

Validity

The validity of a test is a statistical measure that refers to how well it measures what it is designed to measure. It cannot be determined in general terms, such as high or low, but only in reference to the particular use for which the test was designed. There are different types of validity:

Content validity: The extent to which the content of the test represents a balance and adequate sampling of the outcomes (domain) about which inferences are made.

Criterion-related validity: The extent to which scores on the test are in agreement with (concurrent validity) or predict (predictive validity) some criterion measure, e.g., scores on an academic aptitude test administered in high school to grade-point averages over four years of college.

Construct validity: The extent to which a test measures some relatively abstract psychological trait or construct. Construct validity is applicable in evaluating tests such as personality, verbal ability, mechanical aptitude, critical thinking, and creativity.

Variability

The spread of dispersion of test scores, most often expressed as a standard deviation.

REFERENCES

- Armbruster, B.B., Lehr, F., & Osborn, J. (2003). *A child becomes a reader: Birth to preschool: proven ideas from research for parents*. National Parent Information Network. Washington, DC: U.S. Department of Education.
- Atkins, D.H., Kelly, K.T., & Morrison, G.S. (2001). Development of the child evaluation measure: an assessment of children's learning across disciplines and in multiple contexts. *Educational and Psychological Measurement*, *61*(3), 505-511.
- Baker, A. J.L., Piotrkowski, C.S., & Brooks-Gunn, J. (1999). The Home Instruction Program for Preschool Youngsters (HIPPY). *The Future of Children: Home Visiting:Recent Program Evaluations*, *9* (1), p.116-134.
- Blankemeyer, M., Culp, R.E., Hubbs-Tait, L. & McDonald-Culp, A. (2002). A tool for identifying preschoolers' deficits in social competence: The preschool taxonomy of problem situations. *Education and Treatment of Children*, *25*(2) 208-224.
- Bradley, R.H. & Gilkey, B. (2002). The Impact of the Home Instructional Program for Preschool Youngsters (HIPPY) on School Performance in 3rd and 6th grades. *Early Education and Development*, *13*(3), 301-311.
- Buros Mental Measurement Yearbook, Vol. II, III, IV, V, VI, VII, VIII, IX, X, XI, XII, XIII, XIV, XV
- Center for Human Investment Policy, University of Colorado, Denver: The Colorado Parent and Child Foundation: Home Instruction for Parents of Preschool Youngsters (2000).
Http:www.coloradoparentandchildfoundation.org/hippy_research.htm
- Child Trends (2001). *School Readiness: Helping Communities Get Children Ready for School and Schools Ready for Children*. Research Brief. Washington, DC. www.childtrends.org
- Cooahan, K., McWayne, C., Fantuzzo, J. & Grim, S. (2002). Validation of a multidimensional assessment of parenting styles for low-income African-American families with preschool children. *Early Childhood Research Quarterly*, *17*(3), 356-373.
- Crnicek, K. & Lamberty, G. (1994). Reconsidering school readiness: Conceptual and applied perspectives. *Early Education and Development*, *5*(2).
- Denton, K., West, J. & Walston, J. (2003). *Reading-Young Children's Achievement and Classroom Experiences*. NCES 2003-070. Washington, DC: U.S. Department of Education, Institute of Education Sciences.
- Denton, K. & West, J. (2000). *Children's Reading and Mathematics Achievement in Kindergarten and First Grade*. U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- Dickinson, D.K. & DeTemple, J. (1998). Putting parents in the picture: Maternal reports of preschoolers' literacy as a predictor of early reading. *Early Childhood Research Quarterly*, *13*(2), 241-261.
- Fantuzzo, J., Manz, P.H., & McDermott, P. (1998). Preschool Version of the Social Skills Rating System: An empirical analysis of its use with low-income children. *Journal of School Psychology*, *36*(2). 199-214.
- Fantuzzo, J., Hightower, D., Grim, S. & Montes, G. (2002). Generalization of the Child Observation Record: A validity study for diverse samples of urban, low-income preschool children. *Early Childhood Research Quarterly*, *17*(1), 106-125.

Hatch, J.A. & Grieshaber, S. (2002). Child observation and accountability in early childhood education: Perspectives from Australia and the United States. *Early Childhood Education Journal*, 29(4). 227-231.

Head Start National Reporting System (2003). *Description of the Direct Child Assessment Battery*. [Http://www.rivhsa.org/OverviewoftheHSNRS_omb.pdf](http://www.rivhsa.org/OverviewoftheHSNRS_omb.pdf)

Head Start Bureau (2003). *The National Reporting System*. Administration on Children, Youth and Families, U.S. Department of Health and Human Services. [Http: www.headstartinfo.org](http://www.headstartinfo.org).

Head Start Bureau (2003). *Overview of the Head Start National Reporting System Full Implementation*. Administration on Children, Youth and Families, U.S. Department of Health and Human Services. [Http: www.headstartinfo.org](http://www.headstartinfo.org).

Irvine, B., Biglan, A., Smolkowski, K. & Ary, D.V. (1999). The value of the *Parenting Scale* for measuring the discipline practices of parents of middle school children. *Behaviour Research and Therapy*, 37(2), 127-142.

Jacobson, A.L. & Ramiesetty-Mikler, S. (2000). *The HIPPY Corps Initiative: Getting Things Done. 1998-1999 Annual Program Evaluation Report*. Center for Parent Education, Denton, TX: University of North Texas.

Kagan, S.L., Moore, E., & Bredekamp, S. (Eds.). 1995. *Reconsidering children's early development and learning: Toward common views and vocabulary*. Report of the National Education Goals Panel, Goal 1 Technical Planning Group. Washington DC: Government Printing Office. [ED 391576](#).

Lombard, A.D. (1994). **The Effect on Children**. In A. D. Lombard, *Success Begins At Home* (pp.35-56). Sluice Dock, Guilford, CT: The Dushkin Publishing Group, Inc.

Lonigan, C.J. & Whitehurst, G.J. (1998). Relative efficacy of parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly*, 13(2). 263-290.

Love, J.M. (2001). *Instrumentation for State Readiness Assessment: Issues in Measuring Children's Early Development and Learning*. Mathematica Policy Research. Princeton, NJ. Paper prepared for discussion as the Assessing the State of State Assessments Symposium, Atlanta, GA.

Marcon, R.A. (1999). Positive relationships between parent school involvement and public school inner-city preschoolers' development and academic performance. *School Psychology Review*, 28(3), 395-412.

Mann, T., & Powers, S. (1997). *Screening and assessment for Head Start Programs serving infants and toddlers*. Paper presented at the 24th National Head Start Association Training Conference. Boston, MA.

Manset-Williamson, G., St. John, E., Hu, S., & Gordon, D. (2002). Early Literacy Practices as Predictors of Reading Related Outcomes: Test Scores, Test Passing Rates, Retention, and Special Education Referral. *Exceptionality*, 10(1), 11-28.

McConnel, S.R. (2000). Assessment in early intervention and early childhood special education: Building on the past to project into our future. *Topics in Early Childhood Special Education*, 20, p. 43.

McConnell, S.R., Priest, J.S., Davis, S.D., & McEvoy, M.A. (2000). Accountability Systems for Children Between Birth and Age Eight. Early Childhood Research Institute on Measuring Growth and Development, *Center on Early Education and Development*. Minneapolis, MN: University of Minnesota.

McMahon, R.J. & Kotler, J.C. (2001). Differentiating anxious, aggressive, and socially competent preschool children: Validation of the Social Competence and Behavior Evaluation-30 (parent version). *Behaviour Research and Therapy*, 40, 947-959.

Meisels, S.J. (1998). *Assessing Readiness*. Center for the Improvement of Early Reading Achievement (CIERRA Report No. 3-002). Ann Arbor, MI: University of Michigan School of Education.

Merrell, K.W. (2001). Assessment of children's social skills: Recent developments, best practices and new directions. *Exceptionality*, 9(1&2), 3-18.

Morrongiello, B.A. (1997). *Tapping School Readiness in the NLSCY: Measurement Issues and Solutions* (T-98-1E). Income Security and Social Development Studies, Applied Research Branch, Strategic Policy, Human Resources Development, Canada.

National Association for the Education of Young Children. (1988). NAEYC position statement on standardized testing of young children 3 through 8 years of age. *Young Children* 43(3), 42-47, [EJ 367920](#).

National Parent Information Network (1998). *Assessing readiness*.
[Http://www.npin.org/library/2002/n00751/n00751.html](http://www.npin.org/library/2002/n00751/n00751.html)

National Parent Information Network (2001). *Appropriate assessment of readiness*.
[Http://www.npin.org/library/2002/n00759/n00759.html](http://www.npin.org/library/2002/n00759/n00759.html)

Paris, A.H. & Paris, S.G. (2003). Assessing narrative comprehension in young children. *Reading Research Quarterly*, 38, p.36.

Pfannenstiel, J. (1999). *School Entry Assessment Project: Report of Findings*. Overland Park, KS: Missouri Department of Elementary and Secondary Education.

Pickstone, C., Hannon P., & Fox L. (2002). Surveying and screening preschool language development in community-focused interventions programmes: a review of instruments. *Child: Care, Health & Development*, 28, 251-264.

Piotrkowski, C.S., & Botsko, M., & Matthews, E. (2000). Parents' and teachers' beliefs about children's school readiness in a high-need community. *Early Childhood Research Quarterly*, 15(4) 537-558.

Prakash, N., West, J. & Denton, K. (2003). Schools' Use of Assessments for Kindergarten Entrance and Placement: 1998-99. NCES 2003-004. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Pretti-Frontczak, K., Kowalski, K., Brown, R.D. (2002). Preschool teachers' use of assessments and curricula: A statewide examination. *Exceptional Children*, 69, p. 109.

Puma, M., Bell, S., Shapiro, G., Broene, P., Cook, R., Friendman, J. & Heid, C. (2001). *Building Futures: The Head Start Impact Study: Research Design Plan*.
[Http://www.acf.dhhs.gov/programs/care/ongoing_research/hs/impactreports.html](http://www.acf.dhhs.gov/programs/care/ongoing_research/hs/impactreports.html)

Sharif, I., Ozuah, P.O., Dinkevich, E.I., & Mulvihill, M. (2003). Impact of a brief literacy intervention on urban preschoolers. *Early Childhood Education Journal*, 30(3), 177-180.

Shepard, L.A., Taylor, G.A., and Kagan, S.L. (1996). *Trends in Early Childhood Assessment Policies and Practices*. Boulder, CO: University of Colorado, School of Education.

U.S. Department of Education, National Center for Education Statistics (2000). *Children's Direct Cognitive Assessments*. Washington, DC: National Center for Education Statistics.

U.S. Department of Education, National Center for Education Statistics. (2000). *Home Literacy Activities and Signs of Children's Emerging Literacy, 1993 and 1999*. NCES 2000-026, pp.1-16. Washington, DC: U.S. Department of Education.

U.S. Department of Health and Human Services. *Head Start FACES 2000: A Whole-Child Perspective on Program Performance. Chapter 3: Matrix of Measures and Indicators*. Administration on Children, Youth and Families.

Wagner, M., Lida, E., & Spiker, D. *The Multisite Evaluation of the Parents as Teachers Home Visiting Program: Three-Year Findings from One Community*. (SRI International Project PO7283). Menlo Park, CA: SRI International.

Wagner, M. & Spiker, D. (2001). *Multisite Parents as Teachers Evaluation: Experiences and Outcomes for Children and Families*. (SRI International Project PO7283). Menlo Park, CA: SRI International.

West, J., Denton, K., & Germino-Hausken, E. (2000). *America's Kindergartners: Findings from the Early Childhood Longitudinal Study - Kindergarten Class of 1998-99, Fall 1998*. National NCES 2000-070. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

West, J., Kenton, K., & Reaney, L.M. (2000). *The Kindergarten Year: Findings from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99*. Washington, DC: National Center for Education Statistics.

Whitehurst, G.J., Epstein, J.N., Angell, A.L., Payne, A.C., Crone, D.A., & Fischel, J.E. (1994). Outcomes of an Emergent Literacy Intervention in Head Start. *Journal of Educational Psychology*, 86(4), 542-555.

Ysseldyke, J.E., Thurlow, M.L., & Erickson, R.N. (1994). *Possible sources of data for early childhood (age 3) indicators*. U.S. Department of Education, National Center on Educational Outcomes. Minneapolis, MN: University of Minnesota.