DOCUMENT RESUME

ED 329 559   TM 015 817

TITLE   The Bilingual Test Information System.
INSTITUTION   Evaluation Assistance Center--East, Arlington, VA.;
               Georgetown Univ., Washington, D.C.; RMC Research
               Corp., Hampton, N.H.
SPONS AGENCY   Office of Bilingual Education and Minority Languages
               Affairs (ED), Washington, DC.
REPORT NO   TAC-B-55
PUB DATE   Jul 87
NOTE   137p.
PUB TYPE   Guides - Non-Classroom Use (055) -- Book/Product
               Reviews (072) -- Reference Materials - General (130'

EDRS PRICE   MF01/PC06 Plus Postage.
DESCRIPTORS   *Achievement Tests; Affective Measures; *Bilingual
               Education; Criterion Referenced Tests; *Diagnostic
               Tests; Elementary Secondary Education;
               Identification, Information Systems; Language
               Proficiency; *Language Tests; Norm Referenced Tests;
               Oral Language; Program Evaluation; *Standardized
               Tests; *Student Placement; Test Reviews
IDENTIFIERS   *Placement Tests

ABSTRACT
               A resource is presented for those who work in
bilingual education and who are involved in test selection. Initial
information is provided regarding tests that might be used for
student identification, placement and diagnosis, progress monitoring,
exiting decisions, and program evaluation purposes. No endorsement or
recommendation is implied and no test should be selected based on
these reviews. An overview of testing and test selection discusses:
the importance of testing in program evaluation; functions of
norm-referenced and criterion-referenced tests; and four categories
of test summaries. Test reviews are grouped in four categories:
standardized achievement; standardized placement or diagnostic; oral
language proficiency; and affective. The individual test summaries,
when completed, will provide the following information as appropriate
and available: (1) descriptive information about the category; (2)
test description; (3) test structure; (4) language skills chart; (5)
test administration; (6) norming information; (7) scoring
information; (8) out-of-level testing; (9) equivalent tests; and (10)
sources for test documentation information. This document contains a
complete overview, and information on 12 standardized achievement
tests, one standardized placement/diagnostic test, and 6 oral
language proficiency tests. (SLD)

****
*   Reproductions supplied by EDRS are the best that can be made   *
*   from the original document.   *
****

TAC-12-55

ED329559

THE BILINGUAL

TEST

INFORMATION

SYSTEM

EVALUATION ASSISTANCE CENTER (EAST)

Georgetown University
1916 Wilson Boulevard, Suite 302
Arlington, Virginia 22201
(202) 673-3540
(800) 626-5443

RMC Research Corporation
400 Lafayette Road
Hampton, NH 03842
(800) 258-0802
In NH only (800) 582-7175
(603) 926-8888

BEST COPY AVAILABLE

TM015817

THE BILINGUAL TEST INFORMATION SYSTEM
July, 1987

The document which you have received is the first installment of the Bilingual Test Information System -- a resource notebook for those who work in the field of bilingual education and who are, in some way, involved in the process of test selection. This document is meant as a resource, to provide initial information regarding tests which might be used for student identification, placement/diagnosis, progress monitoring, exiting decisions and program evaluation purposes. The information provided within this document can help the user decide what category of test is suitable for a particular evaluation need and to narrow down a list of tests for more in-depth review and analysis.

This document is not meant as an endorsement or recommendation for any test described within. Nor should any test be selected based on these reviews. In all cases, one should identify a group of potentially appropriate tests, then obtain copies for review prior to making a decision regarding which test to use for a particular purpose.

The Bilingual Test Information System consists of five parts -- an overview of testing and test selection, followed by four sections containing test summaries. The overview of testing and test selection discusses the following points:

- the importance of testing in program evaluation;

- functions of norm-referenced and criterion-referenced tests; and,

- four categories of test summaries, namely, _hievement, placement/diagnostic, oral language proficiency and affective.

The individual test summaries, where applicable and available, provide information on the following, as appropriate: descriptive information regarding the test category, general test descriptions, test structure, language skills chart, test administration information, norming information, scoring information, out-of level testing, and sources for test documentation information (e.g., reliability, validity, objectives tested). Specific entries for each test do vary, depending upon the test category and available information.

Please note that this document is the first series of the Bilingual Test Information System and is by no means complete. When referring to the document you will notice that, at this time, the status of each section is as follows:

- The overview of testing and test selection is complete.

- The section on standardized achievement tests includes an introduction and reviews of 12 tests. Additional reviews and materials of use for test selection purposes will be added at a later date.

- The section on placement/diagnostic tests includes only a description of one test at this time. Introductory materials, additional test reviews and other materials will be added at a later date.

- The section on oral language proficiency tests includes descriptions of six tests at this time. The introductory materials, additional reviews and other materials will Le added at a later date.

- The section on affective measures has not been included at this time. The entire section will be added at a later date.

As these additional materials become available, updated versions of the Bilingual Test Information System will be distributed.

The Bilingual Test Information System is produced by the Evaluation Assistance Center - East, one of two national centers funded by the Office of Bilingual Education and Minority Language Affairs to provide technical assistance and training in the areas of program evaluation and testing issues. If you would like further information regarding this document, can recommend particular tests for possible inclusion, or would simply like addition information about the EAC-East, please contact either J. Michael O'Malley at Georgetown University (800-626-5443) or Susan L. Reichman at RMC Research Corporation (800-258-0802 or 800-582-7175 in New Hampshire).

OVERVIEW

THE BILINGUAL

TEST

INFORMATION

SYSTEM

5

OVERVIEW OF THE TEST INFORMATION SYSTEM

INTRODUCTION

The purpose of this system is to provide a resource to those in the field of bilingual and ESL (English as a second language) education who require information about tests for use in program evaluation. The Test Information System includes sections which describe: the importance of testing as it relates to various program evaluation activities; the four categories of tests which are of relevance for program evaluation activities, along with characteristics which are of importance when selecting that particular type of test; and, summaries of commonly used tests in each category.

The Test Information System is intended as a resource, to be referenced in order to: (1) make a decision regarding what category of test would be most suitable for meeting a particular evaluation need, and (2) provide initial screening information on possible tests. It should be noted that the test summaries provided in this document should not be used as a basis for actually selecting a test, but only for narrowing down the potential options. It is always necessary for those who are selecting the test to order and review actual copies of tests and accompanying manuals before making final decisions.

THE IMPORTANCE OF TESTING IN PROGRAM EVALUATIONS

Test information is a critical part of the program evaluation process, whether that evaluation is being conducted to determine how well the program is working or to establish where the program could be improved. In order for this test information to provide relevant and useful data for evaluation purposes, the test must be appropriate for its intended purpose and have certain key qualities in terms of technical standards. To look at it another way, an inappropriate or "poor" test can result in questionable data, which could potentially mislead those conducting the evaluation activities.

There are actually seven areas in which test data contributes to program evaluation activities, including:

- conducting needs assessments in order to determine whether programs are required and, if so, the general types of programs which would meet the existing needs;

- designing entry procedures for selecting students for participation in the program;

- designing exit procedures for determining when students can successfully move out of the program;

- planning specific programs and the instructional components within that program which will be most responsive to student needs;

- establishing whether a program is operating as intended;

1

6

- determining whether the program has an effect on those involved and, if so, the type and extent of that effect; and,

- finding ways to improve various program components.

In considering the wide range of applications which test data may have and the many ways in which the data can be used, the importance of testing as part of the evaluation process becomes clear. A brief overview of how test data contributes to each of these seven types of evaluations follows.

<u>Test Data and Needs Assessments</u>   Before a program is actually designed and implemented, it is necessary to conduct a needs assessment to determine whether programs are required and, if so, the general types which would meet the existing needs. As part of a needs assessment, information should be collected on the status of students who are potential participants in the program. This part of the needs assessment should provide data on the status of these students, including their present level of functioning in comparison to other students who probably do not require any additional services, specific instructional needs, strengths and limitations. While part of this information can be descriptive in nature, test results should provide a large portion of the necessary data.

<u>Test Data and Entry Procedures</u>   Once a decision has been made to actually implement a program, then a method should be established for selecting students from the population of all potentially eligible students to participate in the program. A quantifiable method for selecting those students who would benefit from participation -- entry procedures -- should be established. Entry procedures are often composed of more than one piece of information, such as student records, teacher recommendations, test results, interviews and so forth. Student screening and selection can be done using a composite score approach where a pre-established set of information is collected on each potential student, then this information is used to assign the student a score which would represent need for the program. Test data, both that existing from previous administrations and administered specifically to determine the need for program participation, can contribute relevant data for making entry decisions.

<u>Test Data and Exit Procedures</u>   Part of any program is also a procedure for determining when students are ready to move into the regular program on a full-time basis. As with entry procedures, these exit procedures should be established prior to actual program implementation and should specify exactly what evidence will indicate that the student has mastered enough of the project skills to move out of the program. Exit procedures also generally consist of more than one indicator which is used to make up a composite score. For example, samples of student work, interviews, teacher recommendations and test results may be combined to result in a composite score for determining whether a student is ready to exit the program.

<u>Test Data and Program Planning</u>   Planning specific programs and the instructional components within the program which will be most responsive to student needs entails collecting enough data to develop a complete picture of the specific needs which exist, then using this information to plan specific instructional events which will address the necessary areas. For

a program to be effective, the instructional events must be targeted at a level which matches the entry capabilities of students, provide instruction that will maximize student learning and growth, and at a later date, facilitate successful transition into the mainstream program. Test data on both individual and group needs, strengths, weaknesses and present level of functioning is a critical element in terms of program planning and development.

Test Data and Program Operations   Once a program has been established, it is helpful to take a close look at all its aspects to determine whether the project is actually operating as intended. If there are discrepancies between how the program was intended to operate and what is actually occurring, then it is very difficult to attribute any evaluation results to the program. While much of the data collected for this type of discrepancy evaluation is descriptive in nature, test data also makes a contribution to the process. For example, if the program was originally intended to serve students at a certain entry level, test data could be used to verify whether this, in fac*, is occurring. Similarly, if students were supposed to exit the program only after they had gained certain skills, then test data might be used to help determine whether the exit procedures planned as part of the program operations were functioning as originally intended.

Test Data and Program Effects   Another area which an evaluation might address is that of the effects that the program had on those involved, including students, staff, parents and others directly or indirectly involved with the project. Program effects may range from the more academic areas, such as an improvement in English language proficiency or increased skills in mathematics, to the affective areas, such as changing a students' attitudes toward school. Program effects could also be viewed in terms of staff attitudes and skills, parent attitudes, changes in other areas of student behaviors, such as attendance, and so forth. Clearly, test data on student achievement in subject matter areas, language skills, and affective measures can all contribute relevant data in order to answer questions about program effects.

Test Data and Program Improvement   An ongoing process within any project should be that of conducting evaluations to gather program improvement information, implementing changes, then conducting another evaluation to determine both the effect of the previous change and where additional improvements should be made. Information useful for program improvement purposes could include attitudinal surveys, existing records, observations and test data, to name a few. In regard specifically to test data, for example, it may be helpful to review group performance on achievement test items. These results can point directly to areas which students have yet to master. If that area has supposedly been taught in the program, then a close look needs to be taken at the instructional process and materials being used to determine why students are still having difficulty. If that area is not presently being taught in the program, a potential improvement could be made by adding instruction to help students learn the material. However, just as with other types of evaluations, the test data must be collected using an appropriate and psychometrically sound test. ·

3

TYPES OF TESTING APPROACHES

On a broad level, testing approaches fall into two general categories: the norm-referenced approach and the criterion-referenced approach. These distinctions are important because the type of scores reported for each approach varies. When conducting an evaluation, it is necessary to ensure that the resulting scores provide the type of data required to respond to the intended evaluation purposes.

A **norm-referenced test** is one that is used to determine an individual's status with respect to the performance of other individuals on that test. This definition implies that the purpose of a norm-referenced test is to compare the performance of an individual with that of others. When standardized norm-referenced tests are used in a school setting, the comparison is usually made with those of a similar age or grade level. Performance comparisons are typically made by the use of norms tables, which have been derived through previous administrations of the test to a group of students selected to represent the entire population.

In terms of program evaluations, norm-referenced testing would be used to address issues focusing on comparisons of individual to group performances on the test. For example:

- Norm-referenced tests could be used to determine how performance of students in the program compares to performance of students not in the program. Students not in the program could be peers who met the entry level requirements, but did not participate, or students who did not meet the entry requirements.

- Norm-referenced tests could be used to determine the overall effect of a particular program. In this case, pre- and posttest results would be collected to establish the size of the gains shown by students who did participate in the program. This type of information could contribute to an evaluation aimed at establishing the success of the program.

- Norm-referenced tests could be used to obtain an overall picture of a student's level of functioning at entry into the program and at predetermined points during program participation (e.g., after the first year, second year and third year). While this longitudinal information would not be tied to specific skills, it would provide information about growth over time in the area covered by the test.

- Norm-referenced tests could be used to obtain an overall picture of the areas in which instruction should be provided. Group scores on a norm-referenced test could point directly to areas (e.g., basic math, reading comprehension, study skills) in which there is the greatest need for instruction.

4

Norm-referenced tests, then, are best used when the evaluation activities are investigating comparisons between groups of students who took a particular test.

A **criterion-referenced test** is one that is used to determine an individual's status with respect to specified objectives or skills. For each objective or skill assessed in the test, a set of items is developed to determine whether the student has, in fact, mastered the objective. Prior to the administration of the test, a criterion is established which is then used as the standard to determine whether the resulting scores indicate mastery or non-mastery of the objective(s) tested. A criterion-referenced test may cover one or more instructional objectives, depending upon the purpose of the test.

For program evaluation purposes, criterion-referenced testing would be used to address questions focusing on the mastery or non-mastery of specific objectives or skills. For example:

- Criterion-referenced tests could be used to assess the strengths and weaknesses of individuals as they enter the educational program. Information regarding the degree to which various instructional objectives have already been mastered will help in developing an instructional program most responsive to student needs. This same information can also be used to plan and evaluate the instructional programs.

- Criterion-referenced tests could be used to assess the status of individual students with respect to specific instructional sequences within the program. This information could then be used upon entry into the program to determine where to place the student in regard to instruction; or, after completion of an instructional sequence to determine whether additional instruction is required in that area or whether the student is ready to move on to a new area of instruction. This same information could also be used to evaluate the effectiveness of the instructional program and the curriculum materials being used to teach specific instructional sequences.

- Criterion-referenced tests could be used to assess the strengths and weaknesses of a group in order to determine where to place emphasis within the overall instructional program. This type of information could be used for program development, grouping, staff assignments, and in the evaluation of the later success of the instructional components.

- Criterion-referenced tests could be used to determine the overall effect of a program. The information obtained could be used to describe program effects over time by looking at the length of time the typical student takes to master certain instructional sequences. Areas in which the students continue to have difficulty over time would indicate a potential need for program improvements.

Criterion-referenced test results, then, are best used when the evaluation activities are investigating comparisons regarding how students performed in view of specified objectives or skills.

A few additional differences in regard to norm-referenced and criterion-referenced tests are worth keeping in mind. First, because norm-referenced testing focuses on individual-to-group comparisons, test content is often general and may not be matched to a particular, well-defined area of instructional content. Second, in order to facilitate comparisons in regard to student performance, items are selected to deliberately spread out the distributions of scores.

In comparison, with criterion-referenced testing the focus is on individual comparisons to prespecified standards, so the content is typically much more specific and thus easier to relate to a particular instructional content. Second, items are selected primarily to determine whether a student has mastered a specified objective or skill and not to spread out the range of student scores.

One other point should be kept in mind about both norm-referenced and criterion-referenced tests -- commercial test publishers sometimes offer to score a criterion-referenced test in a norm-referenced manner as well. In this case, student performance would be reported in terms of mastery and non-mastery of specific objectives or skills and also in terms of how the student performed in comparison to the norming group. Norm-referenced tests, on the other hand, generally cannot be scored in a criterion-referenced manner because, as mentioned, the items have been developed to result in a ranking of students along a continuum, rather than to determine whether students have mastered specific objectives.

TYPES OF TESTS

There are four types of tests which are most typically used for program evaluation purposes: achievement tests, placement/diagnostic tests, oral language proficiency devices and affective measures. Each of these particular tests will provide specific types of information for one or more of the seven previously discussed areas of program evaluation. As with the distinction between norm-referenced and criterion-referenced testing, it is necessary to keep in mind the type of information required for the particular evaluation activity in order to select the most appropriate test.

A standardized achievement test is one which is intended to sample a student's present level of learning across a range of general skills areas. The content tested is typically related to formal learning experiences at school and because the intent is to sort and rank students, covers a wide range of topic areas. Standardized achievement tests generally have sound psychometric characteristics and norms which indicate how large groups of students at various grade or age levels performed on the test. Using the test norms, performance of students in the program can be compared to that of a "typical" group of students -- the norming group -- in order to determine general progress and areas of need.

A standardized placement/diagnostic test is one which is intended to identify what the student's strengths and weaknesses are in specific content

areas and how to best help the student overcome any particular limitations. The purpose of a diagnostic test (from this point on, the placement/ diagnostic test will simply be referred to as a diagnostic test) is to identify what the student needs to learn and how best to teach those particular areas. Because diagnostic tests are intended to pinpoint spec'fic problems, the items cover a narrow range of content area, but in depth. As with standardized achievement tests, a standardized diagnostic test generally provides norms which indicate how students at a similar grade or age level performed on the test. Using these test norms, student performance can be compared to that of the "typical" performance for a student of the same grade or age level in order to determine the specific skills which should be taught.

An oral language proficiency device is a method used to determine how well a student is functioning in regard to a specifir spoken language. Oral language proficiency devices use a variety of ways to measure a student's level of performance, ranging from structured interviews to the presentation of pictures with discussions. Oral language proficiency devices can measure a wide range of language acquisition skills, ranging from those necessary for conducting basic interpersonal communications to the those necessary for conducting more difficult activities, such as school learning. Depending upon the particular language acquisition model followed by the test developers, oral language proficiency devices may measure comprehension, fluency, vocabulary, pronunciation, grammar, or functional uses of language. The results of an oral language proficiency device can help determine whether a student is progressing in oral language proficiency and what weaknesses should be addressed in the future.

An affective measure is a device used to measure student attitudes or interests. The purpose of an affective device is generally to establish how a student thinks about types of people, specific situations, types of objects, experiences or other defined areas. While affective measures are not directly useful in terms of planning or providing activities to address specific instructional needs, the results can be useful in terms of overall program planning and evaluation. For example, the results of an affective measure can help establish whether desired changes in student attitudes occurred over the course of a certain time period.

ORGANIZATION OF THE REMAINDER OF THIS DOCUMENT

Following these introductory materials are fou‧ sections, each of which addresses one of the four types of tests (achievement, diagnostic, oral language proficiency and affective). Each section provides a more detailed discussion on that particular type of test, terminology of special relevance, key issues which should be kept mind when reviewing and selecting tests, and summaries describing important features of commonly used tests. Again, keep in mind when using the test descriptions in each section that they are not intended as a way to make a final test selection, but rather to provide preliminary information in order to determine the type of test which will be most appropriate for meeting a particular evaluation need and to help narrow down the choices by providing preliminary information. In all cases, actual copies of the potential tests should be reviewed before a final selection is made.

SOME IMPORTANT TERMINOLOGY

There is some important terminology which should be kept in mind when selecting and reviewing tests for possible program applications. Because these terms are of relevance to more than one type of test, they are highlighted here. Additional terminology, specific to a single type of test, is presented within that section of this document.

Test Features

Standardized (Test) - A test which has items which have been systematically structured to measure a specified content area; and, includes specific instructions for administration, scoring and interpretation of the results. Most standardized tests also typically have certain psychometric characteristics, such as a high reliability and validity indices, as well as test norms for use in interpreting the results.

Test Norms - A set of standards which indicate how the typical student at various grade or age levels performed on the test.

Empirical Norms - These are test norms established by the publisher through the administration of each test level to students of the appropriate age on certain dates in a given year. These dates are actual testing dates and are referred to as the empirical norm, or norming, dates.

Interpolated Norms - Norms which are mathematically derived from empirical norm values. Interpolated norms are **not** actual norms, but rather derived in order to expand the dates on which one can administer a test and then compare the results to some type of test norms.

Administration Considerations

In-Level Testing - The administration of the test level designated by the publisher as being appropriate for a particular age or grade to students of that age or grade.

Functional Level Testing (Out-Of-Level Testing) - The administration of a test one or more levels above or below the level of the test designated by the publisher as appropriate for the student's age or grade. This is done in order to provide a functionally appropriate level of the test and to avoid "floor" or "ceiling" effects.

Ceiling Effects - In a situation where the student has been able to answer all or nearly all of the items on a publisher's recommended level of the test, that student has not had the opportunity to demonstrate his/her true performance level and has encountered the "ceiling." If students, on the average, are getting three-fourths or more of the total items correct, there is probably a ceiling effect and a higher level of the test should be administered.

Floor Effects - In a situation where the student knows answers to only a few items on the publisher's recommended level of a test, that student's score may reflect chance or guessing as much as actual knowledge and the

student is said to have encountered the "floor." If students, on the average, are getting one-third or fewer of the items correct on a test there is probably a floor effect and a lower level of the test should be administered.

Test Scores

Raw Score - The number of items answered correctly.

Mean - The average score, determined by dividing the sum of a group of scores by the number of scores in the group.

Median - The middle score in a set of scores; the point at which 50% of the scores are above and 50% below.

Percentile - A number, running from 1 to 99, indicating the percentage of students scoring at or lower than the test score in question. The percentile scale is not an equal-interval scale -- one percentile point change near the ends of the scale (e.g., between 1 and 2 or 98 and 99) represents a much larger difference than does one point near the middle of the scale (e.g., between 50 and 51). For this reason, percentile scores are not useful for evaluation and comparison purposes.

Percentile Rank - The percent of scores equal to or below a given score.

Standard Score - Raw scores which have been recalculated to have a predetermined mean and standard deviation. A standard score indicates how far a raw score is above or below a mean in terms of standard deviation units. Unlike percentiles, standard scores are equal interval scales and one standard score point always represents the same number of raw score points regardless of the position in the distribution.

Normal Curve Equivalent (NCE) - An equal-interval, normalized standard score with a mean of 50 and a standard deviation of 21.06. NCEs have a range from 1-99 and match the percentile rank at values 1, 50 and 99. Because this score has equal intervals, it can be used for evaluation and comparison purposes.

Expanded Standard Score - This is a special type of standard score used to link different levels of a test. The scale is constructed to span all levels so that the scores from one test level can be converted to their equivalent values on another level. Having expanded standard scores is especially important in regard to using pre/posttest scores for evaluation purposes.

Ps_nometric Characteristics

Confidence Band (Confidence Interval, Confidence Limits) - This band or interval provides an estimate of the range of scores that might be obtained should the student be retested using the same, or parallel, forms of the test.

9

14

Correlation - The relationship between two or more variables, such as two sets of test scores; correlation coefficients range from +1.00 (the maximum possible relationship) to -1.00 (the maximum negative relationship).

Item-Objective Congruence - An index which addresses how well test items measure intended tasks and meet the intended purposes of the test.

Reliability - The degree to which a test consistently measures what it claims to measure. The consistency with which a student's score will remain fairly stable over repeated testings, across retestings using a parallel form of the test, or through calculations based on split halves of the same test.

Standard Deviation - A statistic which indicates the spread or variability of a set of scores. The more the scores fall close to the mean, the smaller the standard deviation; the greater the variation of the test scores, the larger the standard deviation.

Standard Error of Measurement (SEM) - A statistic which estimates the amount of possible error in an obtained score. The SEM estimates how much the obtained score may vary from the "true" score.

Variance - A measure of the spread or variability of a set of test scores. The greater the spread of the scores, the larger the variance.

Variability - The spread of test scores, most typically reported in terms of the standard deviation.

Validity - The degree to which the test measures what it claims to measure. Content validity refers to whether the content of the test is appropriate and in enough depth to actually measure what the test claims to measure. Construct validity refers to how well the test measures the theoretical construct (e.g., reading comprehension, problem solving) it claims to measure. Predictive validity, sometimes also called criterion validity, refers to the relationship between test scores and some other type of criterion (e.g., an achievement test score and course grades).

STANDARDIZED

ACHIEVEMENT

TESTS

THE BILINGUAL

TEST

INFORMATION

SYSTEM

16

STANDARDIZED ACHIEVEMENT TESTS

INTRODUCTION

A standardized achievement test is one which is intended to sample a student's present level of learning across a range of general skills areas. The content tested is typically related to formal learning experiences at school and because the intent is to sort and rank students, covers a wide range of topic areas. Standardized achievement tests generally have sound psychometric characteristics and norms which indicate how large groups of students at various grade or age levels performed on the test. Using the test norms, performance of students in the project can be compared to that of a "typical" group of students -- the norming group -- in order to determine general progress and areas of need.

Standardized achievement tests are typically developed using extensive trials and revisions in order to result in a test which has good psychometric characteristics; that is to say, the test is field-tested and revised until it satisfactorily measures what it purports to measure in a consistent manner. Standardized achievement tests, as mentioned, include test norms. which indicate how large groups of students at different age or grade levels performed on the test as a whole and on any subtests.

Standardized achievement tests provide a way for a project to obtain a measure of how **groups of students** are progressing over time. The results of standardized achievement tests also provide an indication of strengths and weaknesses of a group, thus pointing to possible instructional needs which should be addressed. Achievement test results do **not** indicate what a student might accomplish in the future, nor should the results be used to assess or "grade" individuals.

Four points of special interest should be kept in mind in regard to the use of standardized achievement tests with limited English proficient students.

- First, because the test does measure school learning, the contents may be somewhat culturally bound. Cultural background and experience could therefore affect how a certain student scores on the test.

- Second, because the group on which the test was normed may be different from those students in the project, it is necessary to keep these differences in mind when comparing group test results to the test norms.

- Third, because standardized achievement tests sample a wide range of skills, the test may not be sensitive enough to pick up growth for students who have been in a program for a limited time period.

- Fourth, because limited English proficient students may be given instruction over skills not provided to mainstreamed students, standardized achievement tests may measure very different skills from those being taught in the project.

SELECTING ACHIEVEMENT TESTS

Regardless of the type of program being evaluated, there are some general issues which should be addressed when selecting an achievement test for program evaluation purposes. When reviewing tests for possible use, be sure to keep the following six areas in mind.

1. **The match between the content measured on the test and the program objectives and related instruction.** Clearly a major consideration when selecting a test for program evaluation purposes is how well the content measured on the test matches the program objectives and related instruction. A preliminary determination for initial screening purposes can be made based on the test structure information presented by the publisher. This test structure information generally presents a list of constructs measured on the test, along with general skill areas for each construct (e.g., the construct of reading may include vocabulary, reading comprehension, oral comprehension and word analysis). After a preliminary set of tests have been selected, it is necessary to review copies of the tests to determine: the number of program objectives covered by the test; how well the items match specific program objectives; and, whether the objectives are tested in enough depth to provide information necessary to meet the intended applications.

2. **The psychometric characteristics of the test.** One advantage of standardized achievement tests is that they generally meet high technical standards, including those of validity and reliability. These high technical standards help ensure that the test measures what it purports to measure and does so fairly consistently, with as little error as possible. However, the psychometric characteristics will vary, depending upon the process used to develop and revise the test items, the number of students used to field test the items, and the constructs being measured on the test (some constructs, such as basic math, are easier to measure than others such as problem solving, which is less well defined). All other factors being equal, the test with the best psychometric characteristics would be the test of choice for program evaluation purposes.

3. **Scoring options available.** The scoring options available should also be taken into consideration when selecting a test. One choice involves whether the test can be hand scored, machine scored or either option taken. Clearly, if large numbers of students will be taking the test and complex scoring is required, then machine scoring is desirable. On the other hand, if small numbers of students are involved and the test results are wanted very quickly, then hand scoring would probably be preferable. Additionally, it is also important to find out what types of scores and analyses (e.g., standard scores, NCEs, percentiles, total test scores, subtest scores, objective-based mastery scores, individual and group test results, pre/posttest gains) are available. The types of scores and analyses available must match the information needs of the project and also be suitable for answering the evaluation questions at hand.

4. **Administration requirements.** Of consideration here are issues such as the level of training required for administering the test, the setting (individual, small group, large group) in which the test must be administered and the length of time required for students to take the test. Obviously these administrative requirements affect the suitability of a particular test. For example, a test which requires extensive training for administration and must be administered on an individual basis would not be a good choice for a project which intended to test a large number of students and where the staff did not have time to be trained in the administrative procedures.

5. **Availability and suitability of the test norms.** Another consideration when selecting a test is whether there are test norms which can be used for comparison purposes and, if so, whether the norms were developed in a manner such that the comparisons would be appropriate. There are three frequently occurring limitations to test norms. First, the numbers of students used in the norming group may be too small to make comparisons meaningful. For example, a norming group of 100 students, over eight grade levels would not really be a good norming sample on which to base generalizations.

Second, the students included in the norming group may not be enough like students in the project group to legitimately make certain kinds of comparisons. For example, comparing growth in reading comprehension skills over the course of a year of students who are limited English proficient to their English proficient peers would not be a legitimate comparison. However, one might compare growth to determine whether or to what degree the limited English proficient students are catching up to their English proficient peers.

Third, it is also necessary to take into account the dates on which the test must be administered in order for the empirical norms to be used. (Remember that the empirical norming dates are those established by the publisher through administration of each test level to students in the norming group.) If the test cannot be administered at a time period within these dates, then the test norms should not be used for program evaluation purposes.

6. **Examinee Appropriateness.** Finally, when selecting a test, it is necessary to consider whether that test is appropriate for the students who will be taking it. Appropriateness involves a range of issues, including the difficulty level of the test, content suitability, cultural factors and language versions available. Each one of these factors can result in test results which are not an accurate reflection of the student's true score.

The difficulty level of the test addresses whether a test is too easy or too difficult for a particular student or groups of students. Test scores resulting from items which are much too difficult or to which the student has had little or no exposure, will reflect primarily guessing, rather than the student's knowledge. Test scores resulting from items which are too easy will not reflect the upper limits of the student's knowledge. If a particular test does not have a range of test levels available, it may be difficult to obtain an accurate reflection of the student's knowledge.

1-3

Keep in mind also that students must be able to read the test in a manner sufficient to understand and respond to the test items. The ability of the student to read must be considered for both tests administered in English and in the student's native language. Finally, remember that the student's ability to conduct interpersonal communications orally may be at a higher level than his/her ability to use language for school-based activities.

The issue of content suitability is, in many ways, related to the difficulty levels of the test items. Most test publishers make recommendations regarding how far above or below the recommended level a test may be administered and the recommendation is generally one level. This is especially important in regard to using lower levels of a test with older students -- while the knowledge tested may be suitable, often the materials are inappropriate for an older student (e.g., the drawings are more suitable for a very young child and are inappropriate for the older, more mature, student). Inappropriate materials can cause students to not score as accurately as they might had the materials been more suitable for their age level.

Somewhat related to this issue of context are cultural factors, which may vary by cultural group. The background of the student may have an affect on how certain items or situations are interpreted or understood and thus, how the item is answered. Similarly, students with certain types of backgrounds and experiences may be unfamiliar with some standardized testing approaches and item formats. For example, the student may have had experience with essay tests, but not multiple choice items or the use of answer sheets. This unfamiliarity can affect student's test scores, distorting the test results.

Language versions available will, of course, be especially important when the interest is in evaluating achievement in various content areas, such as in science and mathematics. The test must be available in the language in which the student is proficient. Otherwise, the issue of language interferes with obtaining a score which accurately reflects student's achievement levels.

ORGANIZATION OF THE ACHIEVEMENT TEST SUMMARY REVIEWS

For each achievement test summary review, the following information is provided:

Descriptive Information - the test title; forms available; range of grade/age levels for which the test is intended; and, the publisher's address and telephone number

General Test Description - brief statements summarizing the purpose or function of the test; the content areas tested; any special test characteristics (e.g., whether practice tests are available, whether the test has to be administered orally or individually); and, language versions available

Test Structure - information regarding the various test and subtest scores available for each test level; the availability of in-level norms; and, the number of items and corresponding administration times for the entire test and subtests

Norming Information - the publisher's recommended testing periods (norming dates) and any special notes regarding norming (e.g., whether norms are empirical or extrapolated, whether limited English proficient students were represented in the norming group)

Scoring Information - the types of scoring services available (hand, machine); the types of scores available (e.g., standard scores, objective-based mastery scores, subtest analyses); how to obtain standard scores and NCEs; and, any special scoring cautions which should be noted (e.g., the lack of availability of certain types of scores)

Out-of-Level Testing - whether out-of-level testing can be done and, if so, the range of out-of-level testing which is appropriate; the name of the expanded score; and, whether locator tests are available to determine the most appropriate level

Equivalent Tests - whether the test is linked to any other tests; whether there are corresponding subtests across tests; and, if appropriate, tables which depict any linkages across tests and subtests

Test Documentation - identification of the publisher's source documents for additional information on item objectives or content classification, norming sample descriptions, score conversion tables, reliability and validity

STANDARDIZED ACHIEVEMENT TESTS:

CALIFORNIA ACHIEVEMENT TEST, 1985

```
        FORMS:     E and F

        GRADES:    K-12

PUBLISHER:  CTB/McGraw Hill
            Del Monte Research Park
            Monterey, CA 93940

        (800) 538-9547
```

1. DESCRIPTION

- A series of nationally normed tests designed to measure achievement in reading, mathematics, language arts, spelling, and study skills. Optional tests are available in science and social studies.

- Levels 10-12 are available in Form E only. Levels 13-20 are available in both Forms E and F.

- An additional level (Level 20) has been added to the CAT E/F to provide better content coverage in high school.

2. TEST STRUCTURE

TEST LEVELS AND IN-LEVEL NORMS

Subtest and Composite Scores Available	LEVEL 10 Kg. F,S Gr. 1-F		LEVEL 11 Kg.-S Gr. 1-F,S		LEVEL 12 Gr. 1-S Gr. 2-F,S		LEVEL 13 Gr 2-S Gr. 3-F,S		LEVEL 14 Gr. 3-S Gr. 4-F,S	
READING:	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
Vocabulary	30	25	30	30	30	30	35	28	55	30
Reading Comprehension	--	--	30	30	30	35	35	43	55	50
Oral Comprehension	22	25	--	--	--	--	--	--	--	--
Total Prereading	52	50	--	--	--	--	--	--	--	--
Total Reading	--	--	60	60	60	65	70	71	110	80
Visual Recognition	24	20	--	--	--	--	--	--	--	--
Sound Recognition	20	20	--	--	--	--	--	--	--	--
Word Analysis	--	--	40	20	45	33	40	28	24	10
LANGUAGE ARTS:										
Language Expression	20	20	25	20	30	35	30	30	55	50
Language Mechanics	--	--	--	--	30	23	35	28	35	17
Total Language Arts	--	--	--	--	60	58	65	58	90	67
Spelling	--	--	--	--	30	20	30	20	35	15
Study Skills	--	--	--	--	--	--	--	--	25	-
MATHEMATICS:										
Mathematics	20	20	--	--	--	--	--	--	--	--
Computation	--	--	24	20	24	24	44	40	50	45
Concepts and Applications	--	--	30	20	36	35	48	40	55	50
Total Mathematics	--	--	54	40	60	59	92	80	110	95

Subtest and Composite Scores Available	LEVEL 15 Gr. 4-S Gr. 5-F,S		LEVEL 16 Gr. 5-S Gr. 6-F,S		LEVEL 17 Gr. 6-S Gr. 7-F,S		LEVEL 18 Gr. 7-S Gr. 8-F,S		LEVEL 19 Gr. 8-S Gr. 9-F,S		LEVEL 20 Gr. 10-S Gr. 11-F,S Gr. 12-F,S	
READING:	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
Vocabulary	55	30	55	30	55	30	55	30	55	30	55	30
Reading Comprehension	55	50	55	50	55	50	55	50	55	50	55	50
Oral Comprehension	--	--	--	--	--	--	--	--	--	--	--	--
Total Prereading	--	--	--	--	--	--	--	--	--	--	--	--
Total Reading	110	80	110	80	110	80	110	80	110	80	110	80
Visual Recognition	--	--	--	--	--	--	--	--	--	--	--	--
Sound Recognition	--	--	--	--	--	--	--	--	--	--	--	--
Word Analysis	24	10	24	10	--	--	--	--	--	--	--	--
LANGUAGE ARTS:												
Language Expression	55	50	55	50	55	50	55	50	55	50	55	50
Language Mechanics	35	17	35	17	35	17	35	17	35	17	35	17
Total Language Arts	90	67	90	67	90	67	90	67	90	67	90	67
Spelling	35	15	35	15	35	15	35	15	35	15	35	15
Study Skills	25	25	25	25	25	25	25	25	25	25	25	25
MATHEMATICS:												
Mathematics	--	--	--	--	--	--	--	--	--	--	--	--
Computation	50	45	50	45	50	45	50	45	50	50	50	45
Concepts and Applications	55	50	55	50	55	50	55	50	55	50	55	50
Total Mathematics	110	95	110	95	110	95	110	95	110	95	110	95

TEST LEVELS AND IN-LEVEL NORMS (header spanning the table)

3. NORMING INFORMATION

Recommended Testing Periods

	Grades	Levels	Earliest Administration	Midpoint	Latest Administration
Fall	K-12	10-20	Oct. 9	Oct. 22	Nov. 5
Mid-Year			Not Applicable		
Spring	K-12	10-20	April 18	May 2	May 16

Special Notes On Norming

- CAT E/F has a single set of Norms Books and includes tables for converting raw scores to scale scores, and scale scores to national percentiles, grade equivalents, or NCEs.

- The publisher provides Quarter-Month norms tables with both empirical and interpolated norms within the compliance period (from 3-6 weeks on either side of the week of standardization). These tables may by obtained from the publisher's regional offices.

- Panel of educators of various ethnic groups reviewed test items for possible bias; empirical item bias test procedures were used. Percent of Hispanic students tested in schools was from 5.8% to 9.1%.

4. SCORING INFORMATION

Method

Hand and machine

Scoring Services

- School districts can choose between the traditional scoring method based upon the number of correct responses (raw scores) or a newer method based on Item Response Theory (three-parameter IRT Model). Scale scores and other derived scores can be obtained through either scoring method. IRT scoring is available only through the publisher's scoring service.

- Norm-referenced scores avail le include raw scores, scale scores, percentile ranks and stanines, grade equivalents, and NCEs. Criterion-referenced scores available include percent mastery and an average objectives performance index.

- The basic scoring service provides the Class Record Sheet, including up to six scores chosen by the district.

- A variety of additional, optional scoring services are available such as the Individual Test Record, the Class Grouping Report, the Objectives Performance Report, and Evaluation Summary among others.

- Districts with microcomputer capability can score and report the CAT E/F locally using CTB's Microcomputer Instructional Management System.

Obtaining Standard Scores and NCEs

RS to SS:	Raw Score to Scale Score
RS to NCE:	Raw Score to Scale Score Scale Score to NCE

Scoring Cautions

Be sure to note correct grade level when looking up norms.

5. OUT-OF-LEVEL-TESTING

Name of Expanded Standard Score

Scale Score

Range of Out-Of-Level Testing

The publisher has advised that districts may test two levels out-of-level.

Locator Test

Two locator tests are available: Test 1 for Grades 1-6; Test 2 for Grades 6-12.

6. EQUIVALENT TESTS

- The CAT/E has been equated with both norm-referenced and criterion-referenced tests.

- Equated norm-referenced tests include the CAT/C and the CTBS/U.

- Equated criterion-referenced tests include the DMI/MS and the PRI/RS.

7. TEST DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	List of Category Objectives; Test Coordinator's Handbook: pp. 23-64, 101-113
NORMING SAMPLE DESCRIPTION	Test Coordinator's Handbook: pp. 21-22; Classroom Management Guide; Technical Bulletin No. 1: pp. 24-28, 56-61
SCORE CONVERSION TABLES	Multi-Level Norms Booklets (one each for Fall and Spring)
RELIABILITY	Technical Bulletin No. 2: pp. 49-57, 127-141
VALIDITY	Technical Bulletin No. 1: pp. 10-23; and Technical Bulletin No. 2: pp. 1-3

CIRCUS, 1972-1979

FORMS: X and Y

GRADES: Pre-K-3

PUBLISHER: Addison Wesley Publishing Company
              South Street
              Reading, MA   01867

              (617) 944-3700

1. DESCRIPTION

- CIRCUS comprises the first four test levels of Sequential Test of
  Educational Progress (STEP), Series III. CIRCUS Levels C and D are
  related through content and a common standard scale score to CIRCUS
  Levels A and B and to some measures in Step III.

- The CIRCUS can be administered individually or in groups.

- Forms X and Y are available only for Levels C and D of the Basic
  Assessment Tests. Otherwise, only one form is available.

- CIRCUS is also available in a Spanish version called El Circo.

2. TEST STRUCTURE

	TEST LEVELS AND IN-LEVEL NORMS							
Total and Subtest Scores Available	LEVEL A Prekg.-M Kg.-F		LEVEL B Kg.-S Gr. 1-F		LEVEL C Gr. 1-S Gr. 2-F		LEVEL D Gr. 2-S Gr. 3-F	
	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
READING:								
Word Puzzles/Prereading	--	--	39	--	--	--	--	--
Sounds of Words**	--	--	--	--	6	--	--	--
Consonants*	--	--	--	--	8	--	9	--
Whole Words*	--	--	--	--	24	--	21	--
Phonetic Analysis	--	--	39	--	38	30	30	25
Vocabulary*	--	--	--	--	16	--	12	--
Comprehension*	--	--	--	--	29	--	30	--
Total Reading	--	--	39	40	45	40	42	40
LANGUAGE ARTS:								
Listen to the Story/Listening	25	30	38	35	40	40	40	40
Writing Skills:								
Spelling*	--	--	--	--	--	--	11	--
Word Structure*	--	--	--	--	--	--	21	--
Capitalization & Punctuation*	--	--	--	--	--	--	11	--
Writing Skills	--	--	--	--	--	--	43	40
Total Lang je Arts	--	--	--	--	--	--	--	--
MATHEMATICS:								
How Much and How Many	40	--	58	--	--	--	--	--
Computation*	--	--	--	--	10	--	12	--
Numerical & Mathematical Concept*	--	--	--	--	15	--	14	--
Relational Concepts*	--	--	--	--	15	--	14	--
Total Mathematics	40	40	58	40	40	40	40	40

* Domain scores are available; expanded standard scores are not available for domains.
**Sounds of Words is not a separate domain or norm-referenced score, but items are included in the Phonetic Analysis score.

3. NORMING INFORMATION

Recommended Testing Periods

	Grades	Levels	Earliest Administration	Midpoint	Latest Administration
Fall	K-3	A,B,C,D	Oct. 1	Oct. 15	Oct. 29
Mid-Year	Pre-K	A	Jan. 1	Jan. 15	Jan. 29
Spring	K-2	B,C,D	April 26	May 10	May 24

Special Notes on Norming

- Norms are available only for the Basic Assessment Tests.

- National norm group contained 3.9% of Hispanics and 2.0% "other."

4. SCORING INFORMATION

Method

- Hand and machine

- A variety of scores are available for CIRCUS. They include raw scores, standard scores, percentile ranks, percentile bands, stanines, grade level indicators, domain scores, and verbal reports.

- Scoring for Levels A and B emphasize verbal reports: pupil performance is expressed both in general categories and in sentence reports.

Obtaining Standard Scores and NCEs

```
RS to SS:   Raw Score to Standard Score
            (available only for Basic Assessment Tests)

RS to NCE:  Raw Score to Percentile
            Percentile to NCE
```

Scoring Cautions

None

5. **OUT-OF-LEVEL TESTING**

Name of Expanded Standard Score

Standard Score

Range of Out-of-Level Testing

Testing more than one level out-of-level is not recommended by the publisher.

Locator Test

None

6. EQUIVALENT TESTS

- Cooperative Primary Scaled Scores comparable to CIRCUS Standard Scores are available for the following subtests:

COOP	CIRCUS
Reading	Reading
Writing	Writing Skills
Mathematics	Mathematics
Word Analysis	Word Puzzles and Phonetic Analysis
Listening	Listening

- Continuous coordinated assessment with CIRCUS through test content and at the expanded standard score is available through the administration of tests listed below.

	CIRCUS			STEP III
A	B	C	D	E-J
	Word Puzzles (Prereading)	Reading	Reading	Reading
	Word Puzzles (Prereading)	Phonetic Analysis	Phonetic Analysis	
Listen to the Story	Listen to the Story	Listening	Listening	Listening
How Much & How Many	How Much & How Many	Mathematics	Mathematics	Math Basic Concepts & Math Computation
			Writing Skills	Writing Skills

7. TEST DOCUMENTATION

Test Information   **Source Document/Page**

ITEM OBJECTIVES OR
CONTENT CLASSIFICATION

STEP Test Development and Content Description; User's Guide (one for each test for both Levels A and B combined and one for both Levels C and D combined); CIRCUS Manual and Technical Report, 1979: Levels A and B, pp. 14-24; Level D, pp. 25-31

NORMING SAMPLE DESCRIPTION

CIRCUS Manual and Technical Report, 1979: pp. 68-101

SCORE CONVERSION TABLES

CIRCUS Levels A-D Individual Norms Tables; STEP Individual Norms Tables (CIRCUS C and D Basic Assessment, STEP III, SCAT III, End-of-Course Tests); CIRCUS C and D User's Guide, Basic Assessment, Preliminary Edition (provides domain scores)

RELIABILITY

CIRCUS Manual and Technical Report, 1979: pp. 106-124

VALIDITY

CIRCUS Manual and Technical Report, 1979: pp. 115-119

COMPREHENSIVE TESTS OF BASIC SKILLS, SPANISH EDITION
1978 (CTBS ESPANOL)

```
FORMS:     One only

GRADES:    K-12


PUBLISHER: CTB/McGraw-Hill
           Order Service Center
           2500 Garden Road
           Monterey, CA 93940

           (408) 649-8400
           (800) 538-9547
           (800) 682-9222 (in CA)
```

1. DESCRIPTION

- Test battery designed to measure achievement through Spanish in word recognition, reading vocabulary, reading comprehension, mathematics, computation and mathematics concepts and applications.

- Only Levels B through 3 of CTBS Form S were adapted for the Spanish Edition.

- Listed below are the test levels and corresponding grades for both the CTBS Form S and CTBS Espanol:

TEST LEVEL	CTBS, FORM S GRADES	CTBS ESPANOL GRADES
B	K.6 - 1.6	1
C	1.6 - 2.9	2
1	2.5 - 4.9	3 and 4
2	4.5 - 6.9	5 and 6
3	6.5 - 8.9	7 and 8[*]

*Can be used up to grade 12.

2. TEST STRUCTURE

Total and Subtest Scores Available	TEST LEVELS AND IN-LEVEL NORMS									
	LEVEL B Gr. 1		LEVEL C Gr. 2		LEVEL 1 Gr. 3&4		LEVEL 2 Gr. 5&6		LEVEL 3 Gr. 7&8	
	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
Word Recognition 1 (Reconocimiento de Palabras I)	19	un-timed or 16	--	--	--	--	--	--	--	--
Word Recognition 2 (Reconocimiento de Palabras II)	19	14	--	--	--	--	--	--	--	--
Reading Vocabulary (Vocabulario de Lectura)	--	--	33	15	40	15	40	14	40	12
Reading Comprehension (Comprension de Lectura)	24	15	41	41	45	35	45	35	45	35
TOTAL READING	62	45	74	56	85	50	85	49	85	47
Mathematics Computation (Computation de Matematicas)	32	32	28	34	48	40	48	40	48	40
Mathematics Concepts and Applications (Conceptos y Aplicaciones de Matematicas)	24	un-timed or 26	25	25	50	35	50	35	50	35
TOTAL MATH	56	58	53	59	98	75	98	75	98	75

3. NORMING INFORMATION

Recommended Testing Periods

	Grades	Levels	Earliest Administration	Midpoint	Latest Administration
Fall	Not Applicable				
Mid-Year	Not Applicable				
Spring*	1-8	B-3	April 24	May 8	April 22

*Spring norms are the only empirical norms available.

Special Notes on Norming

• New norms were not developed for CTBS Espanol. Percentiles in the norms tables are not norms; assuming basic competency in both English and Spanish, they are estimates of the percentile a student with a given raw score on CTBS Espanol would have received if CTBS Form S had been taken.

• The equating study used to develop the percentiles occurred in the spring. To avoid systematic bias in derived scores, districts should administer CTBS Espanol in the spring.

• Only spring norms (end of grade) are empirical. Beginning and middle of grade norms are extrapolated.

4. SCORING INFORMATION

Method

* Hand and machine

* For Levels B, C and 1, hand- and machine-scorable test books are available; for Levels 1, 2 and 3, CompuScan answer sheets can be obtained.

Scoring Services

The following prepared reports are available: classroom; pretest-posttest; school; district; individual item; item analysis; frequency distributions and evaluator's summaries; and, selected student reports.

Obtaining Standard Scores

> RS to SS: Raw Score to Scale Score

5. OUT-OF-LEVEL TESTING

Name of Expanded Standard Score

Scale Score

Range of Out-of-Level Testing

None indicated

Locator Test

The Spanish-English Reading and Vocabulary Screening (SERVS) is a pretest developed for use with CTBS Espanol to determine the student's dominant language and whether the student is to be tested with the English or Spanish version of the CTBS.

6. EQUIVALENT TESTS

CTBS Espanol is an adaptation of the CTBS, Form S. CTBS Espanol is similar to CTBS Form S in rationale and process/content classification scheme.

7. TEST DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	Technical Report: p. 2; Examiner's manual, Level B: pp. 3-6; Examiner's Manual, Level C: pp. 5-7; Examiner's Manual, Level 1: pp. 5-8; Examiner's Manual, Level 2: pp. 3-7; Examiner's Manual, Level 3: pp. 3-6
NORMING SAMPLE DESCRIPTIONS	Technical Bulletin No. 1: pp. 5-14 (for CTBS, Form S, 1973-1975) Technical Report: pp. 1, 8-11; Examiner's Manual, Level B: pp. 4-5; Examiner's Manual, Level C: pp. 6-7; Examiner's Manual, Level 1: p. 6; Examiner's Manual, Level 2: p. 4; Examiner's Manual, Level 3: p. 4

Test Information	Source Document/Page
SCORE CONVERSION TABLES	Examiner's Manual, Level B: pp. 31-40; Examiner's Manual, Level C: pp. 37-46; Examiner's Manual, Level 1: pp. 37-57; Examiner's Manual, Level 2: pp. 23-45; Examiner's Manual, Level 3: pp. 23-40.
RELIABILITY	Technical Report: p. 22
VALIDITY	Technical Report: pp. 1-2

1-25

41

COMPREHENSIVE TESTS OF BASIC SKILLS, 1981-82

```
             FORMS:    U and V

             GRADES:   K-12

 PUBLISHER:  CTB/McGraw Hill
             Delmonte Research Park
             Monterey, CA   93940

             (408) 649-8400
```

1. DESCRIPTION

- Achievement test battery measuring reading, language, spelling, mathematics, reference skills, science and social studies.

- Form V covers levels D through J only.

- Level K has a higher ceiling than Level J and is more appropriate for testing students in college preparatory programs.

2. TEST STRUCTURE

Total and Subtest Scores Available	TEST LEVELS AND IN-LEVEL NORMS							
	LEVEL A Kg.-F,S Gr. 1-F		LEVEL B Kg.-S Gr. 1-F		LEVEL C Gr. 1-F,S		LEVEL D Gr. 1-S Gr. 2-F.S	
READING	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
Reading Vocabulary	16	19	17	20	25	37	25	19
Reading Comprehension	--	--	--	--	25	35	25	28
Oral Comprehension	15	21	15	20	--	--	--	--
Total Reading	32	40	32	40	50	72	50	47
Visual Recognition	18	20	--	--	--	--	--	--
Sound Recognition	17	21	--	--	--	--	--	--
Word Attack	--	--	32	28	30	27	40	38
LANGUAGE ARTS:								
Mechanics	--	--	--	--	--	--	20	15
Expression	--	--	17	21	20	30	25	27
Total Language	--	--	--	--	--	--	45	42
Spelling	--	--	--	--	--	--	25	17
MATHEMATICS:								
Computation	--	--	--	--	20	20	20	18
Concepts & Applications	15	21	20	20	25	23	30	33
Total Mathematics	--	--	--	--	45	43	50	51

TEST LEVELS AND IN-LEVEL NORMS

Total and Subtest Scores Available	LEVEL E Gr. 2-S Gr. 3-F,S		LEVEL F Gr. 3-S Gr. 4-F,S		LEVEL G Gr. 4-S Gr. 5&6-F,S		LEVEL H Gr. 5-S Gr. 7&8-F,S		LEVEL J Gr. 8-S Gr. 10,11&12	
READING:	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
Reading Vocabulary	30	27	45	26	45	26	45	26	45	26
Reading Comprehension	30	35	45	43	45	43	45	43	45	43
Oral Comprehension	--	--	--	--	--	--	--	--	--	--
Total Reading	60	62	30	69	90	69	90	69	90	69
Visual Recognition	--	--	--	--	--	--	--	--	--	--
Sound Recognition	--	--	--	--	--	--	--	--	--	--
Word Attack	35	30	--	--	--	--	--	--	--	--
LANGUAGE ARTS:										
Mechanics	20	17	30	20	30	20	30	20	30	20
Expression	25	21	45	37	45	37	45	37	45	37
Total Language	45	38	75	57	75	57	75	57	75	57
Spelling	25	16	30	14	30	14	30	14	30	14
MATHEMATICS:										
Computation	36	34	40	33	40	33	40	33	40	33
Concepts & Applications	40	33	45	37	45	37	45	37	45	37
Total Mathematics	76	67	45	70	85	70	85	70	85	70

44

3. NORMING INFORMATION

Recommended Testing Periods

	Grades	Levels	Earliest Administration	Midpoint	Latest Administration
Fall	K-12	A-J	Sept. 30	Oct. 14	Oct. 28
Mid-Year		Not Applicable			
Spring	K-12	A-J	April 15	April 29	May 13

Special Notes on Norming

- The publisher had divided the ten months of the school year into 40 Quarter Months (QM).

- Districts should carefully select test administration dates to coincide with calendar dates that correspond to specific Quarter Month norms tables, as appropriate.

Calendar Dates:	Quarter Month (QM) Table to Use:

Sept.   01-07......................................QM 0
                 08-15......................................QM 1
                 16-22......................................QM 2
                 23-30......................................QM 5

Oct.   01-07......................................QM 5
                 08-15....................Week of standardization Use QM 5
                 16-23......................................QM 5
                 24-30......................................QM 5

Nov.   01-07......................................QM 8
                 08-15......................................QM 9
                 16-22......................................QM 10
                 23-30......................................QM 11

45

| | Quarter Month (QM) |
| Calendar Dates: | Table to Use: |

Mar.	08-15...................................QM 8
	16-23...................................QM 26
	24-31...................................QM 27

Apr.	01-07...................................QM 28
	08-15...................................QM 31
	16-22...................................QM 31
	23-30...................Week of standardization Use QM 31

May	01-07...................................QM 31
	08-15...................................QM 31
	16-23...................................QM 34
	23-32...................................QM 35

| June | 01-07...................................QM 36 |
| | 08-15...................................QM 37 |

● Review panel screened test items for appropriateness of language, subject matter and portrayal of ethnic groups.

● A chart listing all Quarter Months for which empirical or interpolated norms tables are available and which are within six weeks or norming period midpoint is presented as follows:

4. SCORING INFORMATION

Method

● Hand and machine

● If machine-scoring services are used, the scale scores may be obtained using one of two different methods: number-correct or item response theory (IRT).

● If hand-scoring is done, the scale scores may be obtained using only the number-correct method.

46

Scoring Services

● Special scoring services include: (1) Pre-Post Class: Record Sheet (Class, School, or District Report) showing each student's pretest and posttest scores and difference between the two and also, average pretest and posttest scores and their difference; (2) Pre-Post Evaluator's Summary (Class or School and District Report) showing pretest means, posttest means, and median scores; and (3) Individual Longitudinal Report (a graphic, norm-referenced report) designed to track a student's growth in achievement over several school years. NCE scores are included among the score options available.

Obtaining Standard Scores and NCEs

RS to SS:	Raw Score to Scale Score
RS to NCE:	Raw Score to Scale Score
	Scale Score to NCE

Scoring Cautions

● When ordering machine-scoring services or when planning for hand-scoring, identify the proper Quarter Month corresponding to the dates of testing.

● Tables for converting raw scores to scale scores are available in the regular Norm Books. There are three of these, one for each levels A-E, F-G, and H-J. Tables for converting scale scores to grade equivalents are also available in these norms books.

● The tables for converting scale scores to national percentiles, national stanines, and NCEs available in the regular Norms Books, are appropriate only for testing done in Quarter Months 3-7 and 29-33 (the week of standardization plus/minus two weeks). If testing is done in Quarter months 0-11 or 25-37 (2-6 weeks from the week of standardization), these conversions must be done using the Interpolated Norms Books. There are two of these, one for fall and one for spring, covering all levels of the test.

5. OUT-OF-LEVEL TESTING

Name of Expanded Scale Score

Sca'e Score

Range of Out-of-Level Testing

The publisher had advised that districts test no more than two levels cut-of-level for functional level testing.

Locator Test

There are two locator tests for reading and mathematics combined. Test 1 is for grades 1-6 (levels A-G); Test 2 is for grades 6-12 (levels H-J).

6. EQUIVALENT TESTS

- Scale scores for Form U were equated with those for Form S. The number-right method for Form U was used. Only levels C, 1, 2, 3, and 4 of Form S were equated with levels D-J of Form U. The only normed total scores that were equated were Total Reading, Total Mathematics, Total Language and Total Battery.

- Scale Scores for Form U in reading and mathematics can be predicted from category objective raw scores from the PRI/DMI, respectively. Therefore, normative data can be estimated for the PRI/RS and DMI using Form U norms. This procedure is available through the publisher's scoring service only.

7. TEST DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	Test Coordinators Handbook: pp. 19-39
NORMING SAMPLE DESCRIPTION	Technical Report: pp. 75-93
SCORE CONVERSION TABLES	Regular Norms Books (one for each of levels A-E, F-G, and H-J). Interpolated Norms Books (one for fall and one for spring - all levels)
RELIABILITY	Technical Report: pp. 115-144
VALIDITY	Technical Report: pp. 9-74

IOWA TESTS OF BASIC SKILLS, 1978-1982

FORMS: 7 and 8

GRADES: K-9

PUBLISHER: Riverside Publishing Company
8420 Byrn Mawr Road
Chicago, IL 60631

(800) 323-9540
(312) 693-0040 (IL, HI, AK)

1. DESCRIPTION

- This test edition replaces the ITBS 1976 and is designed to provide comprehensive assessment of student achievement in the areas of basic skills: reading comprehension, vocabulary, language skills, work study skills, and mathematics skills.

- Uniform directions and time limits allow simultaneous administration of Levels 9 through 14 within the same classroom. Score reporting services also accommodate multi-level testing.

- Norms are provided for both the 1978 and 1982 standardizations of the ITBS.

2. TEST STRUCTURE

	TEST LEVELS AND IN-LEVEL NORMS							
Total Subtest Scores Available	EARLY PRIMARY 5 Kg.-F,S Gr. 1-F		EARLY PRIMARY 6 Kg.-S Gr. 1-F,S		PRIMARY 7 Gr. 1-S Gr. 2-F		PRIMARY 8 Gr. 2-S Gr. 3-F	
READING:	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
LI: Listening	31	25	31	25	32	16	32	16
WA: Word Analysis	35	25	35	25	49	20	57	20
V: Vocabulary	29	20	29	20	30	14	30	14
R: Reading	--	--	70	45	66	34	67	34
Total Reading	--	--	--	--	--	--	--	--
LANGUAGE ARTS:								
L: Language	29	20	27	20	--	--	--	--
L-1: Spelling	--	--	--	--	27	13	29	13
L-2: Capitalization	--	--	--	--	66	12	75	12
L-3: Punctuation	--	--	--	--	50	13	69	13
L-4: Usage	--	--	--	--	24	9	23	9
Total Language Skills	--	--	--	--	167	47	195	47
W-1: Visual Materials	--	--	--	--	29	24	32	24
W-2: Reference Materials	--	--	--	--	33	25	38	25
Total Work Study Skills	--	--	--	--	62	49	70	49
MATHEMATICS:								
M: Mathematics	33	25	33	25	--	--	--	--
M-1: Concepts	--	--	--	--	33	15	36	15
M-2: Problem Solving	--	--	--	--	22	18	24	18
M-3: Computation	--	--	--	--	26	22	28	22
Total Mathematics	--	--	--	--	81	55	88	55

TEST LEVELS AND IN-LEVEL NORMS

Total Subtest Scores Available	LEVEL 9 Gr. 3-F.S		LEVEL 10 Gr. 4-F.S		LEVEL 11 Gr. 5-F.S		LEVEL 12 Gr. 6-F.S		LEVEL 13 Gr. 7-F.S		LEVEL 14 Gr. 8&9-F.S	
READING:	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
LI: Listening	--	--	--	--	--	--	--	--	--	--	--	--
WA: Word Analysis	--	--	--	--	--	--	--	--	--	--	--	--
V: Vocabulary	30	15	36	15	39	15	42	15	43	15	44	15
R: Reading	44	42	49	42	54	42	56	42	57	42	58	42
Total Reading	--	--	--	--	--	--	--	--	--	--	--	--
LANGUAGE ARTS:												
L: Language	--	--	--	--	--	--	--	--	--	--	--	--
L-1: Spelling	30	12	36	12	40	12	42	12	43	12	44	12
L-2: Capitalization	28	12	29	12	30	12	30	12	31	12	32	12
L-3: Punctuation	28	14	29	14	30	14	30	14	31	14	32	14
L-4: Usage	27	14	29	14	30	14	30	14	31	14	32	14
Total Language Skills	113	52	123	52	130	52	132	52	136	52	140	52
W-1: Visual Materials	36	40	40	40	46	40	49	40	52	40	54	40
W-2: Reference Materials	37	25	44	25	45	25	46	25	47	25	48	25
Total Work Study Skills	73	65	84	65	91	65	95	65	99	65	102	65
MATHEMATICS:												
M: Mathematics	--	--	--	--	--	--	--	--	--	--	--	--
M-1: Concepts	28	25	32	25	37	25	40	25	42	25	44	25
M-2: Problem Solving	23	25	25	25	27	25	29	25	30	25	32	25
M-3: Computation	39	20	42	20	45	20	45	20	45	20	45	20
Total Mathematics	90	70	99	70	109	70	114	70	117	70	121	70

52

3. NORMING INFORMATION

Recommended Testing Periods

	Grades	Levels	Earliest Administration	Midpoint	Latest Administration
Fall	K-3	5-8	Oct. 14	Oct. 28	Nov. 11
	3-9	9-14	Oct. 16	Oct. 30	Nov. 13
Mid-Year	Not Applicable				
Spring	K-3	5-8	April 18	May 2	May 16
	3-9	9-14	April 14	April 28	May 12

Special Notes on Norming

Standardization sample contained .4% of Native Americans, .6% of Asians and 3.5% of Hispanics.

4. SCORING INFORMATION

Method

Hand and machine

Scoring Services

● Standard service includes the List Report which provides, for each pupil, a main score (standard score or grade equivalent), national percentile rank, and stanine. Raw score, NCE, and other scores are optional. List Report provides class, building, and district averages.

● Optional Evaluation Service includes: (1) means and standard deviations for matched pretest-posttest scores, pretest only scores, and posttest only scores; (2) correlations, t-tests, and joint frequency distributions for pretest-posttest scores; and (3) raw scores, grade equivalent scores, and NCE scores for three subtests. Report will not include raw score averages if pupils in same group were tested with both in-level and out-of-level tests.

Obtaining Standard Scores and NCEs

RS to SS:	Raw Score to Standard Score (Standard scores are in a separate publication)
RS to NCE:	Raw Score to Grade Equivalent or Standard Score
	Grade Equivalent or Standard Score to NCE (Conversions are in a separate publication)

Scoring Cautions

None

5. OUT-OF-LEVEL-TESTING

Name of Expanded Standard Score

Standard Score

Range of Out-of-Level Testing

The publisher suggests that tests may be administered one or more levels out-of-level, as appropriate.

Locator Test

None

6. EQUIVALENT TESTS

- ITBS 1971-76 and ITBS 1978-82 editions are equated.

- The 1978 and 1982 version of the ITBS are the same edition; the ITBS 1978 version was renormed in 1982.

- The expanded standard scores on the ITBS 1978-82, are continuous with those of TAP, 1978, Form T, on the following tests:

ITBS (1978-82)	TAP
Reading Comprehension	Reading Comprehension
Mathematics Total	Mathematics
Language Skills Total	Written Expression
Total Work-Study Skills	Using Sources of Information

7. TEST DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	Teacher's Guide (one each for Early Primary, Primary and Multi-level Editions); ITBS 1982, Detailed Skills Objectives
NORMING SAMPLE DESCRIPTION	Preliminary Technical Summary Second Edition: pp. 6, 26-29; and Development of the 1982 Norms
SCORE CONVERSION TABLES	Teacher's Guide (one each for Early Primary, Primary and Multi-level Editions); NCE Norms for Primary and Multilevel Batteries
RELIABILITY	Teacher's Guide (one each for Early Primary, Primary, and Multi-Level Editions); Preliminary Technical Summary, Second Edition: pp. 10-12
VALIDITY	Teacher's Guide (one each for Early Primary, Primary and Multi-Level Editions); Preliminary Technical Summary, Second Edition: pp. 12-15

LA PRUEBA RIVERSIDE DE REALIZACION EN ESPANOL, 1984

```
          FORMS:     Form A

          GRADES:    K-9

PUBLISHER:  Riverside Publishing Company
            8420 Bryn Mawr Avenue
            Chicago, IL  60631

            (800) 323-9540
            (312) 693-0040
```

1. DESCRIPTION

- Spanish language achievement test which determines the degree to which students are literate in Spanish and assesses achievement in reading, language, mathematics, social studies, and science.

- Based in part on 3-R's Test for which it provides Spanish translations of reading and mathematics tests (Levels 6-14) and language tests (Levels 9-14).

2. TEST STRUCTURE

Total and Subtest Scores Available	TEST LEVELS AND IN-LEVEL NORMS									
	Level 6 Gr. K		Level 7 Gr. 1		Level 8 Gr. 2		Level 9 Gr. 3		Level 10 Gr. 4	
	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
Reading/Language	25	30	28	35	28	30	--	--	--	--
Reading	--	--	--	--	--	--	30	30	30	30
Language	--	--	--	--	--	--	25	30	25	30
Mathematics	25	30	25	30	25	30	30	25	30	25
Composite*	50	60	53	65	53	60	85	75	85	75
Social Studies	--	--	--	--	28	25	28	20	28	20
Science	--	--	--	--	28	25	28	20	28	20

Total and Subtest Scores Available	TEST LEVELS AND IN-LEVEL NORMS							
	Level 11 Gr. 5		Level 12 Gr. 6		Level 13 Gr. 7		Level 14 Gr. 8	
	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
Reading/Language	--	--	--	--	--	--	--	--
Reading	30	30	30	30	30	30	30	30
Language	25	30	25	30	25	30	25	30
Mathematics	30	25	30	25	30	25	30	25
Composite*	85	75	85	75	85	75	85	75
Social Studies	28	20	28	20	28	20	28	20
Science	28	20	28	20	28	20	28	20

* Composite raw scores are obtained by averaging Reading, Language, and Mathematics scores.

3. NORMING INFORMATION

Recommended Testing Periods

La Prueba was field tested in spring 1983. No other test periods are recommended.

Special Notes on Norming

- No national norms are provided. Manual provides spring norms for a Texas Reference Group only. Users (i.e.. states, regions, city) should contact the Company regarding the availability of other norms and the possibility of developing their own.

- Test norms are referenced to the indicated test level, not to the student's grade level.

- Reference group sample was comprised of 6500 Spanish-speaking students, grades K-8, from twenty-three school districts in Texas.

4. SCORING INFORMATION

Method

Hand and machine

Scoring Services

- Standard service includes List Report of scores for individual students and Raw Score Frequency Distribution. Optional services include Individual Student Report, Press-On Labels, and Rank List by Grade.

- List Report includes the following four scores for each subtest and the composite: raw score, local percentile rank, reference-group percentile rank, and reference-group stanine. Unless otherwise specified, Texas norms for the spring 1983 testing will be used in supplying the reference-group percentile ranks and stanines. Normal curve equivalents (NCEs) and/or local stanines can be requested.

Obtaining Scale Scores and NCEs

> RS to NCE: Raw Score to Normal Curve
> Equivalent

Scoring Cautions

None

5. OUT-OF-LEVEL TESTING

Name of Expanded Standard Score

None

Range of Out-of-Level Testing

● Divisions for out-of-level testing are based on students' proficiency in Spanish Literacy Skills and indicated in the following chart:

SPANISH LITERACY SKILLS

Grade	High Average or Above Average ("on-level")	Low Average or Below Average
K	Level 6	Level 6
1	Level 7	Level 7
2	Level 8	Level 8
3	Level 9	Level 8
4	Level 10	Level 9
5	Level 11	Level 10
6	Level 12	Level 11
7	Level 13	Level 12
8	Level 14	Level 13
9		Level 14

- If Spanish-language ability is very limited, even out-of-level testing may result in relatively small gains.

- Texas reference-group norms are available only for on-level testing.

- Using on-level norms for out-of-level testing could result in information that might be misleading.

Locator Test

None

6. EQUIVALENT TESTS

La Prueba is the Spanish edition of The 3-R's Test, Form A. Scores have not been equated.

7. TEST DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	Teacher's Guide and Technical Summary: pp. 16-40
NORMING SAMPLE DESCRIPTION	Teacher's Guide and Technical Summary: pp. 9-10
SCORE CONVERSION TABLES	Teacher's Guide and Technical Summary: pp. 57-67
RELIABILITY	Teacher's Guide and Technical Summary: pp. 11-15
VALIDITY	Teacher's Guide and Technical Summary: p. 11

METROPOLITAN ACHIEVEMENT TESTS SURVEY BATTERY, 1985

FORMS: L and M

GRADES: K-12

PUBLISHER: Psychological Corporation
555 Academic Court
San Antonio, Texas 78204-0952

(512) 299-1061

1. DESCRIPTION

● Nationally normed achievement tests in reading, language arts, mathematics, science, and social studies.

● The 1985 edition of the MAT is the sixth edition of the Metropolitan Achievement Tests series.

● The MAT6 Survey Battery and the MAT6 Diagnostic Battery, scheduled to be published in 1986, are coordinated in content and have compatible norm-referenced scores.

● The Reading Comprehension Test yields an Instructional Reading Level (IRL) in the form of a criterion-referenced score which can be used by teachers to plan reading instruction. This test is designed like an informal reading inventory with passages spanning grade levels above and below the on-grade level of the test. Students' independent reading levels and frustration levels are also available, in addition to the IRL.

● Similarly, the MAT6 also provides an Instructional Mathematics Level (IML) in the form of a criterion-referenced score derived from the Math Concepts Test and the Math Computation Test. The IML can be used by teachers to place students in the proper level of a mathematics instructional sequence by matching student performance to graded levels of math textbooks.

o Most of the tests through the Elementary Level can be dictated or partially dictated, thereby minimizing reliance on student reading skills as a requirement of test-taking.

o MAT6 also assesses study skills. This is not a separate test but a compilation of items across test areas yielding a "Research Skills" score.

o Achievement-Ability Comparisons can be obtained when the MAT6 is used in conjunction with the Otis-Lennon School Ability Test.

2. TEST STRUCTURE

	TEST LEVELS AND IN-LEVEL NORMS							
Total Subtest Scores Available	PREPRIMER Kg.-F,S		PRIMER Kg.-S Gr. 1-F,S		PRIMARY 1 Gr. 1-S Gr. 2-F,S		PRIMARY 2 Gr. 2-S Gr. 3-F,S	
READING:	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
Reading	54	48	--	--	--	--	--	--
Word Recognition Skills	--	--	36	27	28	20	28	20
Vocabulary	--	--	15	10	22	15	22	15
Reading Comprehension	--	--	38	44	53	43	55	40
Total Reading	--	--	89	81	103	78	105	75
LANGUAGE ARTS:								
Language	24	25	23	18	24	20	35	25
Spelling	--	--	--	--	21	10	21	10
Total Language Arts	--	--	--	--	45	30	56	35
MATHEMATICS:								
Mathematics	24	25	41	35	--	--	--	--
Concepts	--	--	--	--	32	27	30	21
Computation	--	--	--	--	22	18	30	22
Problem Solving	--	--	--	--	24	22	24	22
Total Mathematics	--	--	--	--	78	67	84	65

Total Subtest Scores Available	ELEMENTARY Gr. 3-S Gr. 4-F.S		INTERMEDIATE Gr. 5 & 6-F.S		ADVANCED 1 Gr. 7, 8 & 9-F.S		ADVANCED 2 Gr. 10, 11 & 12-F.S	
TEST LEVELS AND IN-LEVEL NORMS								
READING:	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
Reading	--	--	--	--	--	--	--	--
Word Recognition Skills	29	20	--	--	--	--	--	--
Vocabulary	22	15	24	15	24	15	24	15
Reading Comprehension	60	40	60	40	60	40	50	30
Total Reading	111	75	84	55	84	55	74	45
LANGUAGE ARTS:								
Language	42	30	50	30	50	30	48	35
Spelling	21	10	25	10	25	10	25	10
Total Language Arts	63	40	75	40	75	40	73	45
MATHEMATICS:								
Mathematics	--	--	--	--	--	--	55	40
Concepts	35	25	35	25	35	25	--	--
Computation	30	23	30	28	30	28	--	--
Problem Solving	30	27	30	30	30	30	--	--
Total Mathematics	95	75	95	83	95	83	--	--

3. NORMING INFORMATION

Recommended Testing Periods

	Grades	Levels	Earliest Administration	Midpoint	Latest Administration
Fall	K-12	Pre-Primer-Advanced 2	Oct. 1	Oct. 15	Oct. 30
Mid-Year	Not Applicable				
Spring	K-12	Primer-Advanced 2	April 11	April 25	May 9

Special Notes on Norming

● Interpolated mid-year norms are available through the publisher.

● Week-of-testing norms will be available from the publisher's scoring service.

● In addition to national norms, non-public and large city norms will be available.

● Standardization sample contained 8.4% of Spanish surnamed students, .5% of Native Americans, 1.6% of Asians and .7% "other." Panel of minority educators reviewed test items, standardization plans and user manuals.

4. SCORING INFORMATION

Scoring Method

Hand and machine

Scoring Services

- MAT6 has two basic services. Basic Service #1 provides two copies of an Individual Profile, Class Summary, Building Summary, and System Summary. Basic Service #2 provides two copies of a List Report and Summaries for the Class, Building, and System.

- A variety of optional services are available such as a Cluster Analysis Report, a Pupil Narrative Report, a Group Item Analysis, and a Microcomputer Analysis Package among others.

- The basic service allows users to select up to five score types to be reported such as raw scores, scaled scores, national percentile ranks and stanines, grade equivalents, and NCEs.

- The current Fall Norms Booklet provides Percentile Ranks, Stanines, Grade Equivalents and Scaled Scores Corresponding to Raw Scores.

Obtaining Standard Scores and NCEs

RS to SS:	Raw Score to Scaled Score
RS to NCE:	Raw Score to Scaled Score
	Scaled Score to Percentile
	Percentile to NCE

Scoring Cautions

The multi-level norms booklets include raw score to scaled score and scaled score to other derived score tables for all levels.

5. OUT-OF-LEVEL TESTING

Name of Expanded Standard Score

Scaled Score

Range of Out-of-Level Testing

The publisher suggests that tests should not be administered more than one level out-of-level.

Locator Test

None

6. EQUIVALENT TESTS

MAT6 will be equated to the 1982 edition of the Stanford Achievement Test and the 1978 edition of the Metropolitan Tests.

7. TEST DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	Teacher's Manual for Interpreting one for each level)
NORMING SAMPLE DESCRIPTION	Technical Manual (in publication)
SCORE CONVERSION TABLES	Norms Booklet (one for each level) and Multi-level Norms Booklet (in publication)
RELIABILITY	Technical Manual (in publication)
VALIDITY	Technical Manual (in publication)

FORMS: 1 and 2

GRADES: K-12

PUBLISHER: Science Research Associates, Inc.
             155 North Wacker Drive
             Chicago, IL   60608

             (800) 621-0664

1. DESCRIPTION

- The SRA Achievement Series surveys general academic progress from grades K through 12.

- No reading is required of students at Level A.

2. TEST STRUCTURE

TEST LEVELS AND IN-LEVEL NORMS

Total and Subtest Scores Available	LEVEL A Kg. F,S Gr. 1-F,S		LEVEL B Gr. 1-S Gr. 2-F,S Gr. 3-F		LEVEL C Gr. 2-S Gr. 3-F,S Gr. 4-F		LEVEL D Gr. 3-S Gr. 4-F,S Gr. 5-F	
READING:	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
Visual Discrimination	25	15	--	--	--	--	--	--
Auditory Discrimination	25	20	25	25	--	--	--	--
Letters and Sounds	25	22	25	20	20	15	--	--
Listening Comprehension	35	40	20	25	20	25	--	--
Vocabulary	--	--	25	15	25	15	30	18
Reading Comprehension	--	--	24	30	24	30	26	30
Total Reading	110	97	119	115	89	85	56	48
LANGUAGE ARTS:								
Mechanics	--	--	--	--	20	15	30	19
Usage	--	--	--	--	20	15	30	20
Spelling	--	--	--	--	29	15	30	16
Mechanics and Usage	--	--	--	--	--	--	--	--
Total Language Arts	--	--	--	--	69	45	90	55
MATHEMATICS:								
Concepts	30	23	30	30	35	30	35	29
Computation	--	--	27	20	27	30	35	35
Concepts and Computation	--	--	--	--	--	--	--	--
Problem Solving	--	--	--	--	--	--	--	--
Total Mathematics	30	23	57	50	62	60	70	64

TEST LEVELS AND IN-LEVEL NORMS

Total and Subtest Scores Available	LEVEL E Gr. 4-S Gr. 5-F,S Gr. 6-F		LEVEL F Gr. 5-S Gr. 6&7-F,S Gr. 8-F		LEVEL G Gr. 7-S Gr. 8&9-F,S Gr. 10-F		LEVEL H Gr. 9,10,11 & 12-F,S	
READING:	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
Visual Discrimination	--	--	--	--	--	--	--	--
Auditory Discrimination	--	--	--	--	--	--	--	--
Letters and Sounds	--	--	--	--	--	--	--	--
Listening Comprehension	--	--	--	--	--	--	--	--
Vocabulary	40	15	40	15	40	15	40	20
Reading Comprehension	50	35	50	35	50	35	50	40
Total Reading	90	50	90	50	90	50	90	60
LANGUAGE ARTS:								
Mechanics	30	15	30	15	30	15	--	--
Usage	40	20	40	20	40	20	--	--
Spelling	30	15	30	15	30	15	30	30
Mechanics and Usage	--	--	--	--	--	--	40	15
Total Language Arts	100	50	100	50	100	50	70	45
MATHEMATICS:								
Concepts	30	25	30	25	30	25	--	--
Computation	40	35	40	35	40	35	--	--
Concepts and Computation	--	--	--	--	--	--	40	35
Problem Solving	30	30	30	30	30	30	40	40
Total Mathematics	100	90	100	90	100	90	80	75

3. NORMING INFORMATION

Recommended Testing Periods

	Grades	Levels	Earliest Administration	Midpoint	Latest Administration
Fall	K-12	A-H	Sept. 17	Oct. 1	Oct. 15
Mid-Year			Not Applicable		
Spring	K-12	A-H	April 8	April 22	May 6

Special Notes on Norming

- "Answer Keys, Norms and Conversion Tables" (one booklet for each test level) have been published. These booklets provide fall and spring empirical norms. The "Preliminary Norms and Conversion Tables" should not be used.

- Standardization sample included 2% of Native Americans, 2.8% of Hispanics, .7% of Asians and .6% "other."

4. SCORING INFORMATION

Method

Hand and machine

Scoring Services

- The basic full service includes List Report of Scores which provides: GSVs, NCEs, national percentile rank and stanines, grade equivalents, and one optional score, for pupils within a class or grade within a school.

1-55

71

● Optional Pre/Posttest Program service provides matched pupil achievement data from two independent test administrations (in reading, mathematics, and language arts) that occur within a school year. The basic scoring service provides separate pretest and posttest reports, and the Pre/Posttest Program service provides merged pre/posttest reports.

● Scores provided in the Pre/Posttest Program service are raw score (RS), Growth Scale Value (GSV), percentile rank (PR) and the normal curve equivalent (NCE). The average PR corresponds to the average GSV. However, the average RS and NCE are calculated from individual scores. Differences in pre-posttest are reported in RS, GSV, NCE, and standard deviation of the NCE.

Obtaining Standard Scores and NCEs

RS to SS: Raw Score to Growth Scale Value

RS to NCE: Raw Score to Percentile
                 Percentile to NCE

Scoring Cautions

● "Answer Keys, Norms and Conversion Tables" (one booklet for each test level) have been published. These booklets provide fall and spring empirical norms. The "Preliminary Norms and Conversion Tables" should not be used.

● For schools that hand-score their tests and test beyond the range of norms available, special expanded norms and conversions tables have been prepared and are available from SRA. These special norms tables enable the user to look up the in-level percentiles. The tables contain percentiles from first-semester kindergarten through second-semester 12th grade for each test score at each test level. For ɛ   y of these expanded norms tables, contact the Test Marketing Cʊ ɔinator at SRA Chicago.

5. OUT-OF-LEVEL TESTING

Name of Expanded Standard Score

Growth Scale Level

Range of Out-of-Level Testing

See information under Locator Test

Locator Test

- Publication "Selecting Appropriate Test Levels in the SRA Achievement Series," (Test Talk, 79-1) is available. Publisher recommends that both test content and difficulty be considered when selecting test levels.

6. EQUIVALENT TESTS

- The expanded standard score (Growth Scale Value) is continuous with the Iowa Test of Educational Development which was restandardized in 1978.

- The Iowa Tests of Educational Development were designed for college bound students in grades 10-12. Level H of the Achievement Series for grades 9-12 is easier than the ITED and is more appropriate for less academically oriented students.

7. TEST DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	User's Guide: pp. 52-64; Skill Areas and Objectives, Form 1 and 2, Levels A-H
NORMING SAMPLE DESCRIPTION	Technical Report #1, 1978: pp. 2-8

Test Information	Source Document/Page
SCORE CONVERSION TABLES	Answer Keys, Norms, and Conversion Tables (one for each test level)
RELIABILITY	Technical Report #1, 1978: pp. 9-14; additional reports in progress
VALIDITY	Reports in progress

74

STANFORD ACHIEVEMENT TEST, 1982

```
          FORMS:    E and F

          GRADES:   1-9

PUBLISHER:  The Psychological Corporation
            555 Academic Court
            San Antonio, TX  78204-0952

          (512) 299-1061
```

1. DESCRIPTION

- A nationally-normed battery of achievement tests in reading, mathe-
  matics, language arts, listening, science and social science.

- Designed to provide both norm-referenced and diagnostic/prescriptive
  information for use in instructional planning and administrative
  decision-making.

- Shares an expanded scale score with the 1982 Stanford Early School
  Achievement Test and the Test of Academic Skills on specific tests
  and subtests.

75

2. TEST STRUCTURE

Total and Subtest Scores Available	TEST LEVELS AND IN-LEVEL NORMS					
	PRIMARY 1 Gr. 1-M,S Gr. 2-F,S		PRIMARY 2 Gr. 2-S Gr. 3-F,S		PRIMARY 3 Gr. 3-S Gr. 4-F,S	
READING:	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
Word Reading	33	20	33	20	--	--
Reading Comprehension	40	25	40	25	60	30
Reading	73	45	73	45	--	--
Word Study Skills	36	20	48	20	54	30
Total Reading	109	65	121	65	114	60
LISTENING:						
Vocabulary	38	20	35	20	38	20
Comprehension	28	20	30	20	40	30
Total Listening	66	40	65	40	78	50
LANGUAGE ARTS:						
Language	--	--	--	--	46	30
Spelling	30	20	30	20	36	15
Total Language Arts	--	--	--	--	82	45
MATHEMATICS:						
Concepts of Number	34	25	34	20	34	20
Computation & Applications	45	45	--	--	--	--
Computation	--	--	38	30	42	35
Applications	--	--	36	25	38	35
Total Mathematics	79	70	108	75	114	90

76

Total and Subtest Scores Available	TESï LEVELS AND IN-LEVEL NORMS					
	INTERMEDIATE 1 Gr. 4-S Gr. 5-F.S		INTERMEDIATE 2 Gr. 5-S Gr. 6&7-F.S		ADVANCED Gr. 7, 8 & 9-F.S	
READING:	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
Word Reading	--	--	--	--	--	--
Reading Comprehension	60	30	60	30	60	30
Reading	--	--	--	--	--	--
Word Study Skills	60	35	60	35	--	--
Total Reading	120	65	120	65	--	--
LISTENING:						
Vocabulary	36	20	36	20	40	20
Comprehension	40	30	40	30	40	30
Total Listening	76	50	76	50	80	50
LANGUAGE ARTS:						
Language	53	30	53	30	59	30
Spelling	40	15	50	15	50	15
Total Language Arts	93	45	103	45	109	45
MATHEMATICS:						
Concepts of Number	34	20	34	20	34	20
Computation & Applications	--	--	--	--	--	--
Computation	44	40	44	40	44	40
Applications	40	35	40	35	40	35
Total Mathematics	118	95	118	95	118	95

77

● An optional writing test provides an assessment of writing skills through holistically scored writing samples.

● A separate "Using Information" score is derived from items embedded in the Mathematics, Language, Science, and Social Studies subtests.

3. NORMING INFORMATION

Recommended Testing Periods

	Grades	Levels	Earliest Administration	Midpoint	Latest Administration
Fall	2-9	Primary 1 -Advanced	Sept. 23	Oct. 7	Oct. 21
Mid-Year	1	Primary 1	Jan. 20	Feb. 3	Feb. 17
Spring	1-9	Primary 1 -Advanced	April 21	May 5	May 19

Special Notes on Norming

● The norms for the Stanford Achievement Test Series, Forms E/F, are published in two formats: regular Norms Booklet (by test level) and Multilevel Norms Booklet.

● The regular Norms Booklet is available for each level of SESAT and Stanford, and for combined Levels 1 and 2 of TASK. Each booklet contains only empirical norms tables for the specific grades and times of the year the test was standardized. The norms tables provide direct conversions from raw scores to percentile ranks, stanines, and grade equivalents. Other score conversion tables included in the regular norms booklets are percentile ranks to normal curve equivalents (percentile to NCE) and raw scores to content cluster performance categories.

78

● The Multilevel Norms Booklet, intended for use in districts which use functional level testing, includes norms tables for all grades and levels of the Stanford series. The norms tables provide direct conversions from raw scores to scaled scores, and from scaled scores to national percentile ranks, stanines, and grade equivalents. A percentile to NCE table is included in the Multilevel Norms Booklet.

● National norm group included 6% of Hispanics and 4% "other." Final test forms were reviewed by a "bias" panel.

4. SCORING

Methods

Hand and machine

Scoring Services

● The Pupil Skills Analysis provides for each subtest and total: raw score, national percentile rank and stanine, and a choice of two additional scores -- scaled score, national NCE, grade equivalent, local percentile rank and stanine, or local NCE. Also, for each content cluster, the number of items correct, number of items possible, and performance category (below average, average, or above average) is provided.

● The Group Skills Analysis for class, school, or district, summarizes the score types appearing on the Pupil Skills Analysis.

● A Pre/Posttest Matching Service is available for use in evaluation.

● The publication <u>Stanford Information Services -- Scoring and Reporting System</u>, provides a detailed description of the score reports available from the publisher's scoring service.

79

Obtaining Scale Scores and NCEs

```
RS to SS:    Raw Score to Scale Score

RS to NCE:   (1) Raw Score to Percentile
                 Percentile to NCE
                 (Norms Booklets)

             (2) Raw Score to Scale Score
                 Scale Score to Percentile
                 Percentile to NCE
                 (Multi-Level Norms Booklets)
```

Scoring Cautions

None

5. **OUT-OF-LEVEL TESTING**

Name of Expanded Scale Score

Scaled Score

Range of Out-of-Level Testing

The advanced level of the SAT '82 is more difficult than Level 1 of the TASK '82. For low-achieving students, the publisher recommends that the SAT '82 Intermediate II level (grades 5-7) be followed by the TASK '82 Level 1.

6. **EQUIVALENT TESTS**

Scaled scores for the 1982 Stanford Achievement Series (SESAT-SAT-TASK) have been equated to scale scores for the 1973 Stanford Achievement Series.

80

7. TEST DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	Stanford Index of Instructional Objectives
NORMING SAMPLE DESCRIPTION	Norms Booklet*: pp. 8-9 (Primary 1 & 2, Intermediate 2, Advanced); pp. 7-8 (Primary 3 & Intermediate 1)
SCORE CONVERSION TABLES	Norms Booklet*: pp. 24-67 (Primary 1); pp. 24-55 (Primary 2); pp. 20-61 (Primary 3 & Intermediate 1); pp. 24-87 Intermediate 2); pp. 26-85 (Advanced)
RELIABILITY	Norms Booklet*: pp. 19-20 (Primary 1 & 2); pp. 16-17 (Primary 3 & Intermediate 1); pp. 19-21 (Intermediate 2); pp. 18-22 (Advanced)
VALIDITY	Stanford Index of Instructional Objectives

* Information also available in Multilevel Norms Booklets.

STANFORD EARLY SCHOOL ACHIEVEMENT TEST, 1982

```
          FORMS:    E

          GRADES:   K-1

PUBLISHER:  The Psychological Corporation
            555 Academic Court
            San Antonio, TX  78204-0952

            (512) 299-1001
```

1. DESCRIPTION

- A group-administered test designed to measure children's achievement upon entrance to kindergarten and during kindergarten through first grade.

- Designed to provide both norm-referenced and diagnostic/prescriptive information for use in instructional planning and administrative decision-making.

- Shares an expanded scale score with the 1982 Stanford Achievement Test and the 1982 Test of Academic Skills on specific tests and subtests.

2. TEST STRUCTURE

Total and Subtest Scores Available	TEST LEVELS AND IN-LEVEL NORMS			
	LEVEL 1 Kg.-F,M,S		LEVEL 2 Kg.-S Gr. 1-F,M,S	
READING:	ITEMS	TIME	ITEMS	TIME
Word Reading	30	15	38	20
Sentence Reading	--	--	30	20
Reading	--	--	29	40
Sounds and Letters	44	30	45	25
Total Reading	74	45	113	65
LANGUAGE ARTS:				
Listening to Words and Stories	45	3C	45	30
Total Language	--	--	--	--
MATHEMATICS:				
Mathematics	40	30	50	30
Total Mathematics	--	--	--	--

83

3. NORMING INFORMATION

Recommended Testing Periods

	Grades	Levels	Earliest Administration	Midpoint	Latest Administration
Fall	K-1	1-2	Sept. 23	Oct. 7	Oct. 21
Mid-Year	K-1	1-2	Jan. 20	Feb. 3	Feb. 17
Spring	K-1	1-2	April 21	May 5	May 19

Special Notes on Norming

- The norms for the Stanford Achievement Test Series, Forms E/F, are published in two formats -- regular Norms Booklet (by test level) and Multilevel Norms Booklet.

- The regular Norms Booklet is available for each level of SESAT and Stanford, and for combined Levels 1 and 2 of TASK. Each Booklet contains only empirical norms tables for the specific grades and times of year the test was standardized. The norms tables provide direct conversions from raw scores to percentile ranks, stanines, and grade equivalents. Other score conversion tables included in the regular norms booklets are percentile ranks to normal curve equivalents (PR-NCE) and raw scores to content cluster performance categories.

- The Multilevel Norms Booklet, intended for use in districts which use functional level testing, includes norms tables for all grades and levels of the Stanford series. The norms tables provide direct conversions from raw scores to scaled scores, and from scaled scores to national percentile ranks, stanines, and grade equivalents. A percentage rank to NCE table is included in the Multilevel Norms Booklet.

- Standardization sample included 6% of Hispanics and 4% "others."

84

4. SCORING

Methods

Hand and machine

Scoring Services

● The Pupil Skills Analysis provides the following for each subtest and total test: raw score, national percentile rank and stanine, and a choice of two additional scores -- scaled score, national NCE, grade equivalent, local percentile rank and stanine, or local NCE. Also, for each content cluster, the number of items correct, number of items possible, and performance category (below average, average, or above average) is provided.

● The Group Skills Analysis for class, school, or district, summarizes the score types appearing on the Pupil Skills Analysis.

● A Pre/Posttest Matching Service is available for use in evaluation.

● The publication Stanford Information Services -- Scoring and Reporting System, provides a detailed description of the score reports available from the publisher's scoring service.

Obtaining Scale Scores and NCEs

RS to SS:	Raw Score to Scale Score
RS to NCE:	(1) Raw Score to Percentile Percentile to NCE (Norms Booklets)
	(2) Raw Score to Scale Score Scale Score to Percentile Percentile to NCE (Multi-Level Norms Booklets)

85

Scoring Cautions

None

5. OUT-OF-LEVEL TESTING

Name of Expanded Scale Score

Scaled Score

Range of Out-of-Level Testing

Level 1 was designed for children who have not had any prior formal instruction. Level 2 was designed for children who have had prior formal instruction.

Locator Test

None

6. EQUIVALENT TESTS

Scaled scores for the 1982 Stanford Achievement Series (SESAT-SAT-TASK) have been equated to scale scores for the 1973 Stanford Achievement Series.

86

7. TEST DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	Stanford Index of Instructional Objectives
NORMING SAMPLE DESCRIPTION	Norms Booklet* (one for each of Levels 1 and 2): pp. 6 & 7
SCORE CONVERSION TABLES	Norms Booklet*: pp. 18-26 (Level 1); pp. 17-23 (Level 2)
RELIABILITY	Norms Booklet*: p. 14
VALIDITY	Stanford Index of Instructional Objectives

*Information also available in Multilevel Norms Booklet.

87

THE 3-R'S TEST, 1982

```
FORMS:      A and B

GRADES:     K-12

PUBLISHER:  The Riverside Publishing Company
            8420 Bryn Mawr Avenue
            Chicago, IL 60031

            (800) 323-9540
            (312) 693-0240 (IL, HI, AK)
```

1. DESCRIPTION

- A nationally norm-referenced battery of tests designed to measure achievement in reading, language arts, and mathematics. In grades K-2 (test levels 6-8), however, only reading and mathematics achievement are tested.

- In addition to the Achievement Edition, the 3-R's is published in an Achievement/Abilities Edition (additional verbal and quantitative abilities subtests, with in-level norms for grades 3-12) and a Class-Period Edition (a shorter version of the Achievement Edition available in Form A only).

- Machine-scorable test booklets are used for tests levels 6-9 to facilitate the test-taking ability of younger students (grades K-3); separate answer sheets are not available for these grades. Answers must be recorded on separate answer sheets for the remaining test levels 10-18 (grades 4-12).

- The 3-R's Test is usually administered in two sessions, preferably on separate days. With levels 6-8, the Practice Test and the Reading Test are administered in the first session whereas the Mathematics Test is administered in the second testing period. With levels 9-18, the Reading Test and Language Test may be administered in the first session with the Mathematics Test administered in a second session.

2. TEST STRUCTURE

Subject Area Test Scores Available	TEST LEVELS AND IN-LEVEL NORMS											
	6 Kg.- F,S		7 Gr. 1- F,S		8 Gr. 2- F,S		9 Gr. 3- F,S		10 Gr. 4- F,S		11 Gr. 5- F, S	
	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
READING:												
Reading	30	50	30	50	35	55	50	45	50	45	50	45
LANGUAGE ARTS:												
Language	--	--	--	--	--	--	40	35	40	35	40	35
MATHEMATICS:												
Mathematics	29	40	29	40	30	40	35	35	35	35	35	35

Subject Area Test Scores Available	TEST LEVELS AND IN-LEVEL NORMS									
	12 Gr. 6- F,S		13 Gr. 7- F,S		14 Gr. 8- F,S		15/16 Gr. 9&10- F,S		17/18 Gr. 11&12- F,S	
	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME	ITEMS	TIME
READING:										
Reading	50	45	50	45	50	45	50	45	50	45
LANGUAGE ARTS:										
Language	40	35	40	35	40	35	40	35	40	35
MATHEMATICS:										
Mathematics	35	35	35	35	35	35	35	35	35	35

Note: Times are estimates for total administration of a test including the distribution of materials, administering a practice test if neces- sary, and actual student test-taking time.

3. NORMING INFORMATION

Recommended Testing Periods:

	Grades	Levels	Earliest Administration	Midpoint	Latest Administration
Fall	K-12	6-18	Sept. 15	Oct. 27	Dec. 8
Mid-Year			Not Applicable		
Spring	K-12	6-18	March 17	April 28	June 9

Special Notes on Norming

- The Teacher's Manual informs users that the fall norms tables should be used if the test was administered before February 1 and that the spring norms tables should be used if the test was given on February 1 or after. Use of the norm-referenced model requires that tests be administered within two weeks on either side of the midpoint norming date or, with interpolation, up to an additional four weeks in either direction.

- Separate norms tables for out-of-level testing are available from the publisher.

- Standardization sample included 5.8% of Hispanics, 1% of Native Americans and .9% of Asians. Review panel screened test items; items were also analyzed for statistical bias.

4. SCORING INFORMATION

Method

Hand and machine

Scoring Services

- The norms tables provide six types of scores: raw score, grade development score, expanded standard score, percentile rank, normal curve equivalent, and stanine.

- Basic scoring service (Plan A) provides three copies of the Student List Report for each class tested, along with one copy per student of the Student Press-On Label.

- Additional scoring services include 13 optional reports/records such as the Classroom Report, Item Analysis Summary, School Administrator's Report, Evaluation Service, etc.

- Scores provided by the basic scoring service on the Student List Report include percentile ranks, stanines, and one additional score (either expanded standard score or grade development score). Optional scores available from the publisher's service include raw scores, NCEs and one of the additional scores listed above.

Obtaining Standard Scores and NCEs

RS to SS: Raw Score to Expanded Standard Score

RS to NCE: Raw Score to Normal Curve Equivalent

Scoring Cautions

For districts doing hand-scoring, separate norms tables for out-of-level testing are available from the publisher.

1-75

91

5. OUT-OF-LEVEL-TESTING

Name of Expanded Standard Score

Expanded Standard Score

Range of Out-of-Level Testing

Publisher recommends testing no more than two levels out-of-level.

Locator Test

None

6. EQUIVALENT TESTS

LaPrueba is the Spanish edition of the 3-R's Test, Form A, Levels 6-14.

7. TEST DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	Teacher's Manual; one each for combined levels 6-8, 9-12, and 13-18
NORMING SAMPLE DESCRIPTION	Technical Manual; each of the three Teacher's Manuals
SCORE CONVERSION TABLES	Norms tables contained in each of the three Teacher's Manuals
RELIABILITY	Technical Manual; each of the three Teacher's Manuals
VALIDITY	Technical Manual; each of the three Teacher's Manuals

WIDE RANGE ACHIEVEMENT TEST, 1978

```
          FORM:    One only

          AGES:    5-Adult

PUBLISHER:    JASTAK Associates, Inc.
              1526 Gilpin Avenue
              Wilmington, DE  19806

              (302) 652-4990
```

1. DESCRIPTION

- A two-level wide range test comprised of three subtests: Reading (recognizing and naming letters and pronouncing words out of context); Spelling (copying marks resembling letters, writing the name, and writing single words to dictation); and Arithmetic (counting, reading number symbols, solving oral problems, and performing written computations). The test, basically a clinical type test, consists of one four-page test booklet which includes both levels.

- Level I is intended for children between the ages of 5 years 0 months to 11 years 11 months.

- Level II is intended for persons from 12 years 0 months to adulthood.

- The Reading subtest must be administered individually. Spelling and some Arithmetic subtests can be administered in a group except for very young limited subjects.

2. TEST STRUCTURE

TEST LEVEL AND IN-LEVEL NORMS

Total and Subtest Scores Available	LEVEL I Age: 5 Years 0 Months to 11 Years 11 Months		LEVEL II Age: 12 Years 0 Months to Adult	
	ITEMS	TIME--Seconds per Item*	ITEMS	TIME--Seconds per Item*
READING: Reading	100	10 seconds	89	10 seconds
LANGUAGE ARTS: Spelling	65	15 seconds	51	15 seconds
MATHEMATICS: Arithmetic	63	10 seconds	57	10 seconds

*Total test administration time is approximately 15-20 minutes.

3. NORMING INFORMATION

Recommended Testing Periods

Not Applicable

Special Notes on Norming

● Midpoints are not applicable; norms are based on age groupings of subjects included in the norming sample.

● The grade ratings that correspond to the raw scores for each subtest are printed on the cover of the test booklet. The grade ratings are not normative scores. The grade ratings must be used to enter the appropriate norms tables which are based on age levels.

● The norms tables provide grade rating score conversions to percentile rank and standard (not expanded) scores. Grade rating scores for pupils can be converted to percentiles for age levels.

● The grade ratings for the 1978 edition are not comparable to those of the 1965 and 1976 editions. The basis for calculating grade ratings was the same for the 1965 and 1976 editions but was different for the 1978 edition.

4. SCORING INFORMATION

Method

Hand

Scoring Services

None

Obtaining Standard Scores and NCEs

RS to SS:	Raw Score to Grade Rating Grade Rating to Standard Score
RS to NCE:	Raw Score to Grade Rating Grade Rating to Percentile Percentile to NCE

Scoring Cautions

The standard scores in the age norms tables are not expanded standard scores and should not be averaged.

5. OUT-OF-LEVEL TESTING

Name of Expanded Standard Score

None

Range of Out-of-Level Testing

Not applicable

Locator Tests

None

6. EQUIVALENT TESTS

None

7. TEST DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	Not Available
NORMING SAMPLE DESCRIPTION	Manual of Instructions: pp. 43-45
SCORE CONVERSION TABLES	Manual of Instructions: pp. 16-42
RELIABILITY	Manual of Instructions: pp. 46-48
VALIDITY	Manual of Instructions: pp. 49-62

STANDARDIZED

PLACEMENT/DIAGNOSTIC

TESTS

THE BILINGUAL

TEST

INFORMATION

SYSTEM

97

STANDARDIZED PLACEMENT/DIAGNOSTIC TESTS:

2-1

KEYMATH DIAGNOSTIC ARITHMETIC TEST
1971 - 1978

FORMS:	One only
GRADES:	K-7
PUBLISHER:	American Guidance Series Publisher's Building Circle Pines, MN 55014
	(612) 786-4343

1. DESCRIPTION

- An individually administered, wide range arithmetic test comprised of 14 subtests organized into three major areas: Content, Operations, and Applications. It has only one level which is administered across grades. Procedures for establishing basal and ceiling levels to determine the number of test items to administer to pupils are included in the manual.

- Keymath provides four levels of diagnostic information: total test performance, subtest performance, and item performance.

- Although the test spans grades K-7, empirical percentiles and NCEs are available only for grades 2-6.

2. TEST STRUCTURE

	TEST LEVELS AND IN-LEVEL NORMS
Total and Subtest Scores Available	Wide Range Level Gr. 2-6: F, S
MATHEMATICS Mathematics	ITEMS 209

Time of administration varies by student depending on number of items administered. Administration and scoring time is usually 30-40 minutes.

3. NORMING INFORMATION

Recommended Testing Periods (1978 only)

	Grades	Levels	Earliest Administration	Midpoint	Latest Administration
Fall	2-6	only one level	Oct. 1	Oct. 15	Oct. 29
Mid-Year	Not Applicable				
Spring	2-6	only one level	April 1	April 15	April 29

2-3

Special Notes on Norming

● The 1971 norms tables included in the Manual provide only raw score to grade equivalent.

● The publisher restandardized the test in 1977-1978. The test content was not revised. The 1977-1978 Supplementary Norms Tables for grades 2-6 are available from the publisher.

● Publisher's directions for using the 1977-1978 norms tables when testing more than two weeks from the empirical midpoint are not in compliance with Model A requirements.

4. SCORING INFORMATION

Scoring Method

Hand and machine

Scoring Services

● Basic scoring services include NCE scores.

● Normal curve equivalent scores corresponding to total test scores are available in separate tables for October and April (1978 Supplementary Norms).

Obtaining Standard Scores and NCEs

RS to SS: Not Applicable

RS to NCE: Raw Score to NCE
(Based upon 1978 norms, grades 2-6;
total test score only.)

Scoring Cautions

None

5. OUT-OF-LEVEL-TESTING

Name of Expanded Standard Score

Not Applicable

Locator Test

Not Applicable

6. EQUIVALENT TESTS

None

7. TEST DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	Manual: Appendix A, pp. 36-48
NORMING SAMPLE DESCRIPTION	KeyMath Manual: pp. 25-27. Supplementary Norms Tables: pp. 1

Test Information	Source Document/Page
SCORE CONVERSION TABLES	Manual: pp. 51-52 Supplementary Norms Tables (1977-1978)
RELIABILITY	Manual: pp. 27-28
VALIDITY	Manual: pp. 29-30

ORAL LANGUAGE

PROFICIENCY

DEVICES

THE BILINGUAL

TEST

INFORMATION

SYSTEM

ORAL LANGUAGE PROFICIENCY DEVICES:

BILINGUAL SYNTAX MEASURE I (BSM I) 1975, 1978

```
FORMS:      BSM-English and
            BSM-Spanish

GRADES:  K-2

PUBLISHER:  The Psychological Corporation
            555 Academic Court
            San Antonio, Texas 78204-0952

            (512) 299-1061
```

1. DESCRIPTION

- Designed to measure student's oral proficiency in English and/or Spanish syntactical structures.

- Indicates the degree to which a student is structurally dominant in English or in Spanish when the results of BSM-S (Spanish) and BSM-E (English) for the same student are compared.

- BSM-E measures English structural proficiency of students whose native language is Spanish or any other language besides English.

- BSM-S measures the Spanish structural proficiency of native and non-native speakers of Spanish.

- BSM-S is designed to measure the degree of maintenance or loss of certain basic Spanish grammatical structures in students who acquired Spanish as a first language; BSM-E measures the same attributes in students who acquired English as a first language.

- BSM is based on the theory that children acquire a second language by a process of creative construction in which they gradually construct rules for the speech they hear. It was found that children acquire syntactic structures in a systematic manner in all languages and that there is a common order of acquisition of certain English grammatical morphemes by children of diverse language backgrounds.

2. TEST STRUCTURE

The BSM-E consists of twenty-five items which measure the following
syntactical structures: word order, pronoun case, progressive -ing,
copula singular, short plural, auxiliary singular, article, copula plu-
ral, auxiliary plural, past regular, present indicative, progressive
-s, long plural, conditional auxiliary, past irregular, perfect auxil-
iary, and past participle.

The BMS-S consists of twenty-five items which measure the following
syntactical structures: word order, progressive, auxiliary estar, cop-
ula ser, copula estar, present indicative, infinitive, adjective
gender, possessive, indirect object pronoun, conjunction que, refle-
xive, article, present subjunctive, direct object pronoun, and past
subjunctive (perfect).

Underlying the syntactic structures, according to the theory upon which
the test is based, is a psycholinguistic continuum of syntactic devel-
opment.

3. LANGUAGE SKILLS CHART FOR BSM I

SKILLS

CONTENT	Listening	Speaking	Reading	Writing
Sound				
Words				
Syntax		X		
Pragmatics				

4. TEST ADMINISTRATION INFORMATION

Administration Method

Individual

Testing Technique

BSM approximates a real conversation between the examiner and student about seven cartoon-type pictures. The examiner asks the student questions about the pictures which the student answers orally. The examiner records the student's oral responses in a booklet.

Administration Requirements

Examiners must be speakers of the language of the test version they administer (English and/or Spanish). Training is needed by examiners in administering the test questions, in recording the student's oral responses, and in scoring test items.

Administration Time

Approximately 10-15 minutes per child.

5. NORMING INFORMATION

How Instrument Was Normed

BSM was not normed. The test research sub-samples drawn from several ethnic groups were not selected to be representative of the entire U.S. and were not matched for length and type of exposure to English and/or Spanish. Therefore, the groups are not comparable.

Number and Type of Students in Test Research Sample

The test research sample numbered 1,603 students. Included in the sample were students whose first language was Spanish and students whose first language was English. The sample is judgmental rather than random and was selected to include pupils from urban, suburban and rural school districts in four U.S. geographic regions. A broad spectrum of Spanish speaking pupils was selected from the three major Hispanic groups in the U.S.: Mexican-American, Cuban and Puerto Rican. The target group consisted of students in grades K-2 in school districts with large concentrations of Spanish speakers.

Scoring Levels: Derivation and Validation

The three upper Proficiency Levels (3-4-5) were empirically established.

6. SCORING INFORMATION

Scoring Method

Hand

How Scores Are Reported

Student scores are reported as one of five Proficiency Levels:

        Proficiency Level 1: No English/Spanish
        Proficiency Level 2: Receptive English/Spanish only
        Proficiency Level 3: Survival English/Spanish
        Proficiency Level 4: Intermediate English/Spanish
        Proficiency Level 5: Proficient in English/Spanish

7. DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	Technical Handbook: pp. 15-17
FIELD TEST SAMPLE DESCRIPTION	Technical Handbook: pp. 34-39; Manual-English: p. 4; Manual-Spanish: p. 4
SCORING - PROFICIENCY LEVEL	Manual-English: p. 12; Child Response Booklet: p. 7; Manual-Spanish: p. 12; Respuesta del nino: p. 7
RELIABILITY	Technical Handbook: pp. 40-46
VALIDITY	Technical Handbook: pp. 30-33

IDEA ORAL LANGUAGE PROFICIENCY TEST I - ENGLISH (IPT I)
1979

```
          FORMS:    A and F

          GRADES:   K-6

PUBLISHER:  Ballard & Tighe, Inc.
            480 Atlas Street
            Brea, California  92621

            (714) 990-4332
```

1. DESCRIPTION

- Screening device to determine level of oral language mastery for placement into <u>IDEA Oral Language Management Program</u>.

- Used to classify students as non-English speaking (NES), limited English speaking (LES), and fluent English speaking (FES).

- Tests four areas of English oral language proficiency: vocabulary, comprehension, syntax and verbal expression which includes articulation.

- Based on view that children pass through four stages of language acquisition: babbling (meaningless sounds); echolalic (repeating); graphic (one-or-more word sentences); and, syntactic (expanded communication, exploration of language and world around).

2. TEST STRUCTURE

The IPT I consists of eighty-three items measuring English listening-speaking skills in four areas as follows:

AREA	ITEMS
Vocabulary	20
Comprehension	19
Syntax	18
Verbal Expression*	26

*Verbal Expression includes articulation skills.

3. LANGUAGE SKILLS CHART

SKILLS

CONTENT	Listening	Speaking	Reading	Writing
Sounds	X	X		
Words	X	X		
Syntax	X	X		
Pragmatics				

3-8   112

4. TEST ADMINISTRATION INFORMATION

Administration Method

Individual

Testing Techniques

Testing techniques vary within the levels of the test. Techniques used are: question-and-answer; commands which expect an oral or physical response; sentence repetition; listening to a story/answering oral questions; characterizing spoken pairs of words as being the same or different; and story-retelling.

Administration Time

14 minutes (average)

Administration Requirements

● It is strongly recommended that examiners be bilingual in English and the language of the student.

● Inservice training in the use of the test is needed by examiners as well as access to the Examiner's Manual.

5. NORMING INFORMATION

How Instrument Was Normed

Data were collected from research samples in several studies. These were used to develop a normative interpretation of student language proficiency Score Levels.

Number and Type of Students in Norming Group

Combined research samples consisted of 2,061 K-8 students in ten California rural, urban and suburban school districts from bilingual classrooms with identified NES/LES students available for testing. Specifically, the scores of monolingual English and fluent English-speaking students in those classrooms were used. These data were combined with those of 364 students who took part in the Spring, 1981 Construct Validity Study.

Scoring Levels: Derivation and Validation

Scoring levels (NES/LES/FES) were derived by correlating IPT I scores with NES/LES/FES classification levels established by the California State Department of Education. The validity of the NES/LES/FES designations were assessed by correlating them with teacher ratings of students.

6. **SCORING INFORMATION**

Scoring Method

Hand

How Scores Are Reported

- IPT Score Levels A, B, C, D, E, F, M

- English Language Proficiency Levels:

    NES - Non-English Speaking
    LES - Limited English Speaking
    FES - Fluent English Speaking

7. TEST DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	Examiner's Manual: pp. 13-14
FIELD TEST SAMPLE DESCRIPTION	Technical Manual: pp. 7-10, 24, B-2 to B-3, C-2 to C-3
SCORING-PROFICIENCY LEVEL	Examiner's Manual: pp. 9-11
RELIABILITY	Technical Manual: pp. 28-35
VALIDITY	Technical Manual: pp. 11-27

IDEA ORAL LANGUAGE PROFICIENCY TEST - SPANISH, IPT (SPANISH)
1980

```
            FORMS:     One only

            GRADES:    K-8

PUBLISHER:  Ballard & Tighe, Inc.
            480 Atlas Street
            Brea, California 92621

            (714) 990-4332
```

1. DESCRIPTION

● Screening device to determine level of oral language mastery for placement into Spanish IDEA Oral Language Program.

● Used to classify students as non-Spanish speaking (NSS), limited Spanish speaking (LSS), and fluent Spanish speaking (FSS).

● Tests four areas of Spanish oral language proficiency: vocabulary, comprehension, syntax and verbal expression which includes articulation.

● Based on the view that children pass through four stages of language acquisition: babbling (meaningless sounds); echolalic (repeating); telegraphic (one-or-more word sentences) and, syntactic (expanded communication, exploration of language and world around).

2. TEST STRUCTURE

The IPT (Spanish) consists of eighty-three items measuring Spanish listening-speaking skills in four areas as follows:

AREA	ITEMS
Vocabulary	16
Comprehension	20
Syntax	21
Verbal Expression*	26

*Verbal Expression includes articulation skills.

3. LANGUAGE SKILLS CHART FOR IPT (Spanish)

SKILLS

CONTENT	Listening	Speaking	Reading	Writing
Sounds	X	X		
Words	X	X		
Syntax	X	X		
Pragmatics				

4. TEST ADMINISTRATION INFORMATION

Administration Method

Individual

Testing Techniques

Testing techniques vary within the levels of the test. Techniques used are: question-and-answer; filling in blanks, orally; sentence repetition; questions about pictures; transformations; directed responses; oral comprehension questions on an oral story; reading of time; picture/sentence matching; reciting of information; verbal commands for verbal/physical responses; requests to infer outcomes to stories and to events in pictures; sound matching; and, story retelling.

Administration Time

14 minutes (average)

Administration Requirements

• Examiners must speak Spanish.

• It is recommended that examiner practice by giving the test at least once before actually administering it for designation of language proficiency.

5. NORMING INFORMATION

How Instrument Was Normed

Data were collected from a research study in spring 1980. These were used to develop a normative interpretation of student language proficiency Score Levels.

Number and Type of Students in Norming Sample

Research sample consisted of 614 students, grades K-8, ages 5-16, male and female divided evenly. Students were recruited from three school districts in California and one in Texas. All students were from recognized bilingual classrooms with identified NSS/LSS students available for testing.

Scoring Levels: Derivation and Validation

Scoring levels (NSS/LSS/FSS) were derived by correlating IPT (Spanish) scores with school district and teacher NSS/LSS/FSS classifications.

6. SCORING INFORMATION

Scoring Method

Hand

How Scores Are Reported

● IPT (Spanish) Score Levels A, B, C, D, E, F, M

● Spanish Language Proficiency Levels

    NSS - Non-Spanish Speaking
    LSS - Limited Spanish Speaking
    FSS - Fluent Spanish Speaking

7. DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	Examiner's Manual: pp. 14-15
FIELD TEST SAMPLE DESCRIPTION	Technical Manual: pp. 7-10
SCORING - PROFICIENCY LEVEL	Examiner's Manual: p. 11; Technical Manual: p. 18
RELIABILITY	Technical Manual: pp. 11-19
VALIDITY	Technical Manual: pp. 20-22

IDEA ORAL LANGUAGE PROFICIENCY TEST II - ENGLISH (IPT II)
1983

```
            FORMS:    A and B

            GRADES:   7-12

PUBLISHER:  Ballard & Tighe, Inc.
            480 Atlas Street
            Brea, California 92621

            (714) 990-4332
```

1. DESCRIPTION

- Assesses phonological structure, morphology and lexical items, syntax, comprehension, oral production, and pragmatics (the appropriate use of language to obtain specific goals).

- Measures adolescent students' degree of "control" over meaningful language forms and "velocity of growth" in elaborating language.

- Results are used to: determine the level of English oral language proficiency of secondary school students who are native speakers of other languages; classify students as non-English proficient (NEP), limited English proficient (LEP), and fluent English proficient (FEP); and, provide diagnostic information for instruction.

- Based on view that children pass through four stages of language acquisition: babbling (meaningless sounds); echolalic (repeating); telegraphic (one-or-more word sentences); and, syntactic (expanded communication, exploration of language and world around).

2. TEST STRUCTURE

The IPT II consists of ninety-one items measuring Spanish listening-speaking skills in four areas as follows:

AREA	ITEMS
Vocabulary	26
Comprehension	19
Syntax	25
Verbal Expression*	21

*Verbal Expression includes articulation skills.

3. LANGUAGE SKILLS CHART FOR IPT II ᐟ

SKILLS

CONTENT	Listening	Speaking	Reading	Writing
Sounds	X	X		.
Words	X	X		
Morphemic Structure		X		
Syntax	X	X		
Pragmatics		X		

122

4. TEST ADMINISTRATION INFORMATION

Administration Method

Individual

Testing Techniques

Testing techniques vary within the levels of the test. Techniques used
are: filling in of blanks, orally; question-and-answer; commands which
expect oral/physical responses; answering questions about pictures;
structural exercises and transformations; reading of time; telling
short stories/retelling short stories; answering questions about oral
stories; reciting information from memory; saying whether words sound
the same or different; requests to infer outcomes of oral stories;
determine mood and retell main ideas; requests to infer mathematical
information from a graph; and, requests to put events in sequence.

Administration Time

15 minutes (average)

Administration Requirements

● It is strongly recommended that the examiner be bilingual in English
  and in the language of the student.

● Examiners must be given pr: tice in administering the test.

5. NORMING INFORMATION

How Instrument Was Normed

Data were collected from research samples in several field studies.
These were used to develop a normative interpretation of student lan-
guage proficiency Score Levels.

Number and Type of Students in Test Research Sample

Research samples consisted of 306 native English speakers and 153
minority students classified as non-, limited and fluent English speakers. All were from school districts in California. Fourteen language
groups were represented, the majority of which was from Spanish and
Vietnamese speaking background. Some other languages represented were
Korean, Cambodian, Laotian and Chinese.

Scoring Levels: Derivation and Validation

Scoring levels (NEP/LEP/FEP) were derived from validity studies involving English-only students and those who are native speakers of languages other than English. The IPT II test data of the language minority groups were correlated with the amount of time spent in the U.S.,
teacher predictions of scores, and district and teacher classification
of language minority students. Data from teacher and district
NEP/LEP/FEP classifications were then correlated with established classification levels established by the California State Department of
Education.

6. SCORING INFORMATION

Scoring Method

Hand

How Scores Are Reported

- IPT Score Levels A, B, C, D, E, F, M

- English Language Proficiency Levels

    NEP - Non-English Proficiency
    LEP - Limited English Proficiency
    FEP - Fluent English Proficiency

7. DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	Examiner's Manual: pp. 13-17; Technical Manual: pp. 73-77
FIELD TEST SAMPLE DESCRIPTION	Technical Manual: pp. 14, 32, 34-35, 37
SCORING - PROFICIENCY LEVEL	Examiner's Manual: p. 10; Technical Manual: p. 55
RELIABILITY	Technical Manual: pp. 56-64
VALIDITY	Technical Manual: pp. 15-55

LANGUAGE ASSESSMENT SCALES (LAS)
1977, 1986

FORMS: A and B in English
           A in Spanish
           LAS Short Form

GRADES: K-5 (Level I)
           6-12 (Level II)

PUBLISHER: Linguametrics Group
           P.O. Box 3495
           San Rafael, CA 94912-3495

           (415) 459-5350

1. DESCRIPTION

● Assesses the following English and Spanish language skills:
  phoneme discrimination and production; knowledge of concrete nouns;
  oral sentence comprehension; oral production; written production
  (optional, Level II only); and pragmatics (ability to use language
  for one's own ends, optional).

● Determines whether student is a proficient, limited or non-speaker
  of English and Spanish.

● Based on the view that language consists of four linguistic
  subsystems: phonology (phonemics, stress, rhythm, intonation); the
  lexicon (the "words" of the language); syntax (the meaningful
  utterances); and pragmatics (the appropriate use of language to
  obtain specific goals).

2. TEST STRUCTURE

Total and Subtest Scores Available	Test Levels - Long Form - English & Spanish	
	Level I Grades K-5	Level II Grades 6-12
	Number of Items	Number of Items
Minimal Pairs	30	24
Lexical	20	20
Phonemes	36	36
Sentence Comprehension	10	10
Oral Production (story retelling)	1 story	1 story
Total	97	91
Written Production (story rewriting; optional)		1 story
Observation (optional)	10	10

Total and Subtest Scores Available	Test Levels - Short Form - English & Spanish	
	Level I Grades K-5	Level II Grades 6-12
	Number of Items	Number of Items
Lexical	20	20
Sentence Comprehension	10	10
Oral Production (story retelling)	1 story	1 story
Total	31	31

● The three LAS Short Form scales (Lexical, Sentence Comprehension and Oral Production) are identical to those of the LAS Long Form.

● LAS Short Form may be used as a preliminary screening instrument. LAS Short Form should be used for K-1 students unless the Pre-LAS is available to assess students ages 4-6.

3. **LANGUAGE SKILLS CHART FOR LAS, LONG FORM**

SKILLS

CONTENT	Listening	Speaking	Reading	Writing
Sounds	X	X		
Words		X		X
Syntax	X	X		X
Pragmatics		X		

LANGUAGE SKILLS CHART FOR LAS, SHORT FORM

SKILLS

CONTENT	Listening	Speaking	Reading	Writing
Sounds				
Words		X		
Syntax	X	X		
Pragmatics				

4. TEST ADMINISTRATION INFORMATION

Administration Method

Individual

Testing Techniques

Each LAS subscale employs a different testing technique. For Minimal Pairs, student is presented with a pair of spoken words and indicates whether they are the same or different; for Lexical, the examiner points to a picture and the student says what it is; for Phonemes, the student repeats words and sentences; for Sentence Comprehension, the student matches a spoken sentence with one of three drawings; for Oral Production, the student listens to a story and retells it to the examiner; for Observations (optional), the examiner completes an observation form which yields student pragmatic language information; and, for Written Production (optional), the student writes down the story (s)he has retold for Oral Production.

Administration Time

Approximately 10 minutes per student for LAS Short Form; 20 minutes for LAS Long Form.

Administration Requirements

● Examiners should be teachers or other school personnel who are qualified to work with students, proficient speakers of English and Spanish, completely familiar with administration of the test, and able to discriminate correct from incorrect responses.

● Examiners must be trained to administer the test.

5. NORMING INFORMATION

How Instrument Was Normed

A random sample of students was used as a research sample to develop the Oral Proficiency Levels and to develop a normative interpretation for each Oral Proficiency Level.

Number and Type of Students in Test Research Sample

The research sample included 2300 students from California, Texas, District of Columbia, Florida, New York, New Mexico, Puerto Rico, Mexico, Guatemala and Honduras. Included were English speakers in the U.S. of Anglo-American and other ethnic backgrounds, and Spanish speakers from within and without the U.S. Socio-economic backgrounds included low, lower-middle and middle class. On the English form, students were from grades K-12; on the Spanish form, from ages 6-17.

Scoring Levels: Derivation and Validation

The five Oral Proficiency Levels were derived empirically through a series of reliability and validity studies.

6. SCORING INFORMATION

Scoring Method

Hand, computer programs (AUTOLAS), and machine.

How Scores Are Reported

Student scores are reported as one of five Oral Proficiency Levels:

    Proficiency Level 1: Non-speaker
    Proficiency Level 2: Non-speaker
    Proficiency Level 3: Limited English Speaker
    Proficiency Level 4: Proficient Speaker
    Proficiency Level 5: Proficient Speaker

7. TEST DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	A Convergent Approach to Oral Language Assessment: pp. 8-61
TEST RESEARCH SAMPLE DESCRIPTION	LAS Technical Supplement B-1: pp. 1-3; LAS Technical Supplement B-11: pp. 1-3; Norms for the LAS (1983): p .1; A Convergent Approach to Oral Language Assessment: pp. 22, 25-29, 31-32, 39-42, 55-59; LAS Short Form Validity and Reliability, pp. 2-3
SCORING-PROFICIENCY LEVELS	Scoring and Interpretation Manual for Language Assessment Scales - LAS 5th Edition: pp. 6-11
RELIABILITY	LAS Technical Supplement B-I: pp. 14-18; LAS Technical Supplement B-II: pp.3-12; A Convergent Approach to Oral Language Assessment: pp. 62-78; LAS Short Form Validity and Reliability: pp. 3-18
VALIDITY	LAS Technical Supplement B-I: pp. 18-25; LAS Technical Supplement B-II: pp. 3-17; A Convergent Approach to Oral Language Assessment: pp. 78-128; the LAS Short Form Validity and Reliability: pp. 3-18

132

PRE-LAS ENGLISH
1985, 1986

FORMS: A and B

AGES: 4-6

PUBLISHER: Lingametrics Group
            P.O. Box 3495
            San Rafael, CA 94912-3495

            (415) 459-5350

1. DESCRIPTION

● Measures children's expressive and receptive abilities in oral
  language components of morphology, syntax and semantics.

● Determines whether student is a proficient, limited or non-speaker
  of English and Spanish.

● Based on the view that language consists of four linguistic
  aspects: phonology (phonemics, stress, rhythm, intonation); the
  lexicon (the "words" of the language); syntax (the meaningful
  utterances); and, pragmatics (the appropriate use of language to
  obtain specific goals).

2. TEST STRUCTURE

Total and Subtest Scores Available	PRE-LAS, Forms A and B Number of Items
Simon Says (following directions)	10
Choose a Picture (match oral stimulus to one of two pictures)	10
What's in the House? (naming of labels for household objects)	10
Say What You Hear (sentence imitation)	10
Finishing Stories (supply spoken clause to complete a compound or complex sentence)	5
Let's Tell Stories (story retelling)	2 stories

3. LANGUAGE SKILLS CHART FOR PRE-LAS ENGLISH

SKILLS

CONTENT	Listening	Speaking	Reading	Writing
Sounds				
Words		X		
Syntax	X	X		
Pragmatics	X			

4. TEST ADMINISTRATION INFORMATION

Administration Method

Individual

Testing Techniques

Each Pre-LAS subscale employs a different testing technique. For Simon Says, student is given an oral command and responds by performing it; for Choose a Picture, student points to one of two pictures which matches what examiner says; for What's in the House, student names object pointed to; for Say What You Hear, student repeats a spoken sentence; for Finishing Stories, examiner starts a sentence orally and student completes it orally; and, for Let's Tell Stories, student retells a story about a set of pictures (s)he views.

Administration Time

Approximately 10 minutes per student.

Administration Requirements

● Examiners should be teachers or school personnel who are qualified to work with 4- to 6 year-olds, proficient speakers of standard English, and able to discriminate correct from incorrect responses.

● Examiners must be trained to administer the test.

5. NORMING INFORMATION

How Instrument Was Normed

A random sample of students was used as a research sample to develop the Oral Proficiency Levels and to develop a normative interpretation for each Level.

Number and Type of Students in Test Research Sample

846 preK-1 students in Texas, California, Oregon, Connecticut and Hawaii participated. Background variables for selection included: language majority/minority, male/female, and ages 3-8.

Scoring Levels: Derivation and Validation

Reported in Pre-LAS Technical Report #1.

6. SCORING INFORMATION

Scoring Method

Hand

How Scores Are Reported

Student scores are reported as one of five Oral Proficiency Levels:

    Proficiency Level 1: Non-speaker
    Proficiency Level 2: Non-speaker
    Proficiency Level 3: Limited English Speaker
    Proficiency Level 4: Proficient Speaker
    Proficiency Level 5: Proficient Speaker

7. TEST DOCUMENTATION

Test Information	Source Document/Page
ITEM OBJECTIVES OR CONTENT CLASSIFICATION	How to Administer Pre-LAS English, Forms A and B: p. 2
TEST RESEARCH SAMPLE DESCRIPTION	Pre-LAS Norms and Reliability on Forms A and B (#521): pp. 1-2
SCORING-PROFICIENCY LEVELS	Scoring and Interpretation Manual for Pre-LAS (2nd edition): pp. 9-10
RELIABILITY	Pre-LAS Norms and Reliability on Forms A and B (#521): pp. 3, 6, 11, 16, 21, 26, 29
VALIDITY	Pre-LAS Technical Report