

Administration Practices for Standardized Assessments

• • • • • • • • • •

Sasha Zucker

With assistance from:

Margarita Galindo

Elaine Grainger

Nancy Severance

April 2004

Administration Practices for Standardized Assessments

Note: This report is based on “Appropriate Administration Practices with Standardized Tests” published in Testing Information Bulletin No. 3 by Harcourt Brace Educational Measurement, June 1996.

Introduction

Pearson Inc. (Pearson) develops and distributes a variety of assessments for educational and clinical purposes. To meet the goal of producing highly valid, reliable results for test users, each of these products is developed according to strict guidelines. For standardized tests, the way in which the test is administered is equally important. By carefully adhering to administration practices, a test user can contribute to the validity and reliability of each student’s results. Moreover, following appropriate administration practices satisfies widely accepted ethical standards and codes of responsibilities required for test administrators. This report discusses general and documented administration practices for standardized assessments, especially with regard to Pearson’s educational assessment products, such as the *Stanford Achievement Test Series*, Tenth Edition (Stanford 10).

Assessment Standardization

The way in which an assessment is administered to students is especially important for standardized assessments. During the development process, *standardization* establishes uniform procedures for using the assessment so that observation, administration, equipment, materials, and scoring rules will be the same for every student to whom it is administered (Millman and Greene, 1993). The goal of this uniformity is to make the assessment’s results as objective as possible so that they can be considered valid and meaningful when used to compare the assessed qualities of students. If an assessment’s results are perceived to have been significantly affected by external factors that occur before, during, or after its administration, they can feasibly be considered subject to different interpretations, thereby decreasing their informational value (Nitko, 2004). Standardization attempts to control these external factors to the greatest degree possible so that the assessment is a valid measurement tool that produces meaningful results.

Administration Practices for Standardized Assessments

For a *norm-referenced assessment*, such as Stanford 10, standardization is especially important. A valid scoring system for a norm-referenced assessment—such as percentile ranks, stanines, or normal curve equivalents—can only be developed using large-scale trials with a well-defined, national sample of students (the norm group). The standardization process is conducted under highly controlled conditions, including the time limits (if specified) for each test in the assessment’s battery, the materials the students may use during the assessment (such as scratch paper or calculators), and the directions for administering (see Figure 1).

When the published version of the assessment is administered to students, the conditions under which the test was normed must be replicated as closely as possible. Deviation from the norming conditions may invalidate scores that measure a student’s achievement compared to the norm group and can make the results less meaningful (Nitko, 2004). Therefore, the administrator of a norm-referenced test should carefully follow the directions for administering.

Specific Directions for Administering

To administer the Practice Test, read aloud the bold-faced directions next to the word **SAY**. Instructions for the test administrator (e.g., the teacher) are printed in plain text and should not be read aloud to the students. The number of the test question being discussed is shown next to the directions.

SAY Today we are going to do some interesting activities. I am going to give each of you a booklet. Leave your booklet closed until I tell you what to do.

Distribute the Practice Test booklets, with the front covers facing up. Be sure each student gets a booklet, a pencil, a marker, and an eraser.

SAY Open your booklet to page 2.

Demonstrate and point. Then walk around the room to make sure that all students have page 2 showing.

SAY Look at the group of words at the top of the page.

Demonstrate and point.

SAY

1 Which one is the word *herself*? *Herself*. Which word is the correct answer?

Pause for responses.

SAY Yes, that’s right. The second word is the word *herself*, isn’t it? The first word is *knowing*, and the third word is *please*. I will show you how to mark the right answer.

Figure 1. Directions for administering the Stanford 10 Primary 1 Practice Test.

Administration Practices for Standardized Assessments

Pearson's rigorous attention to detail is reflected in the extensive directions for administering created for each assessment. The directions for administering for each assessment product provide specific instructions for procedures to follow before, during, and after test administration. They are each developed and evaluated prior to test standardization. During standardization, the materials are verified for accuracy and utility. Care should be taken in following a test's directions for administering. As discussed below, certain accommodations may be necessary.

Assessment Accommodations and Modifications

It is important to note that in certain cases changes to the administration of a standardized assessment may be appropriate. The *No Child Left Behind Act* of 2001 (*NCLB*), as well as the 1997 reauthorization of the *Individuals with Disabilities Act* of 1975 (*IDEA*), requires changes to an assessment to accommodate the needs of students with disabilities and English language learners.

Assessment Accommodations

An assessment *accommodation* can be a change in format, response, setting, timing, or scheduling that does not alter in any significant way what the test measures or the comparability of scores (Phillips, 1993). A change to an assessment qualifies as an accommodation when it has been demonstrated as not invalidating a student's results. Accommodations that are appropriate are often detailed in the *individual education plan* (IEP) for a student with a disability and should already be present in the classroom environment (Nitko, 2004; Thurlow, Elliott, and Ysseldyke, 1998; Tindal and Fuchs, 1999). For further discussion of the types of accommodations that are permitted for Pearson's assessment products, see Case, 2003a and Case, 2003b.

Assessment Modifications

In contrast, an assessment *modification* refers to a change in administration that alters the constructs assessed or the comparability of scores (Phillips, 1993). When making changes to an assessment to meet a student's needs, policy makers (at state education agencies or large school districts) must determine whether the change is an accommodation or a modification. If the change is an accommodation, then the scores can be compared to the scores of other students (and aggregated with the norm group, if applicable). If the change is a modification, the student's results may need to be reported separately from those of the other students and can be compared only to the results of students who were administered the test with the same modification.

Professional Responsibilities in Assessment Administration

Anyone administering an educational assessment to students is expected to assume certain responsibilities documented in the codes and standards of professional education organizations. Some general principles of appropriate conduct for persons administering assessments include protecting the safety, health, and welfare of all examinees; knowing and complying with relevant state and federal laws; maintaining and improving competence in educational assessment; and performing all responsibilities with honesty, integrity, due care, and fairness (Nitko, 2004). The *Code of Professional Responsibilities in Educational Measurement* (the National Council on Measurement in Education, 1995) offers guidance for those who administer assessments, including the responsibilities to “understand the procedures needed to administer the assessment prior to administration”; to “administer standardized assessments according to the prescribed procedures and conditions”; and to “avoid any conditions in the conduct of the assessment that might invalidate the results.” These standards of conduct are widely recognized and apply to anyone who takes the professional responsibility of administering an assessment to students.

Qualifications for Using Assessments

For certain Pearson assessment products, purchasing eligibility is restricted to individuals with specific training and experience in a relevant area of assessment. Pearson complies with the *Standards for Educational and Psychological Testing* (American Educational Research Association et al., 1999) and with the professional and ethical standards of several professional organizations, including the American Counseling Association (ACA), the American Psychological Association (APA), the Council of Exceptional Children (CEC), the National Association of School Psychologists (NASP), and the previously mentioned National Council on Measurement in Education. To uphold these standards, Pearson may require proper credentials from the purchaser of an assessment and retains the right to withhold or withdraw approval for a test purchase where there is evidence of violation of commonly accepted testing practices. See Pearson’s catalogs of assessment products for detailed information about purchasing qualifications.

Qualifications for Using Educational Assessments

Although some assessment instruments require specialized training to ensure their appropriate professional use, Stanford 10 is designed to be administered by a teacher or counselor without the need for additional specialized qualifications or training. For elementary school students, the administrator is usually the students’ teacher. However, the test administrator is expected, in accordance with the professional responsibilities discussed above, to become familiar with the Stanford 10 directions for administering and strictly follow examination procedures.

Administration Practices for Standardized Assessments**Qualifications for Using Psychological Assessments**

Pearson publishes psychological assessment products for a wide variety of users, including paraprofessionals, education specialists, and individuals with a high level of training in administering and interpreting results from specific tests. Some tests, such as intelligence tests, are restricted to individuals with this specific training, while others are not. Pearson indicates who is qualified to use the psychological assessments in its catalog by classifying them as A-, B-, or C-level instruments.

Table 1. Instrument Level Classifications and Qualified Users.

Instrument Level	Qualified Users
A	Individuals without advanced training in assessment administration and interpretation of assessment results.
B	Individuals with more expertise than users qualified for an A-level instrument as the administration of the assessment and interpretation of its results are more complex. Individuals certified by an organization recognized by Pearson and with relevant training and experience in administering and interpreting the assessment may qualify.
C	Individuals with a doctorate in psychology, education or a related field, or licensure. Individuals with certification by certain agencies and national organizations may also qualify based on training and expertise.

For additional information on qualifications for purchasing and using Pearson assessment products, contact the Pearson Qualifications Group at 1-800-211-8378.



Procedures Before Testing

To ensure that a standardized educational assessment produces results that accurately assess student achievement, Pearson recommends that, in addition to the product-specific directions for administering, examiners follow the procedures below.

Maintaining Assessment Security and Integrity

Upon purchasing and receiving assessment materials from Pearson, the test administrator must protect their security and value by guarding them against improper use and distribution. Materials should be stored in locked files or storage cabinets that are accessible only to authorized personnel. No reproduction of test materials is allowed in any form or by any means, electronic or mechanical. The materials should be released only to persons who are qualified to interpret and use them properly, who have a professional, responsible interest in their use, and who agree to safeguard their security and integrity.

Preparing Students for Testing

The test administrator may also be responsible for preparing the students for the assessment. For students to give their maximum performance on an educational assessment, those who are responsible for administering the assessment must provide them with basic information that they require, including when the assessment will be administered, the content and abilities that will be assessed, what the assessment will emphasize, the standard or level of performance expected, how the assessment performance will be scored, and how the results of the assessment will be used to make decisions (Mehrens & Lehmann, 1991; Nitko, 2004).

It is Pearson's view that the best preparation that teachers can provide to their students for a standardized educational achievement test is to focus instruction on the local curriculum rather than on the format and content of a specific test. Schools may also prepare students for a standardized assessment by teaching general test-taking skills, such as paying attention to oral and written directions, writing responses or marking answers neatly, and using assessment time wisely to complete all the required tasks (Nitko, 2004; Oosterhof, 2001).

To assist students in preparing for Stanford 10, Pearson publishes Stanford 10 practice tests that were factored into the assessment's standardization process. Materials for students who require further assistance with test-taking skills are published by Pearson Achieve under the Rigby and Steck-Vaughn imprints (<http://www.pearsonassess.com>).

Stanford 10 Practice Tests

Stanford 10 incorporates a number of design innovations that set it apart from traditional standardized achievement tests, including full-color layout, use of place markers to guide early-grade-level students, and no time limits. To help familiarize both test administrators and students with this new test-taking experience in a way

Administration Practices for Standardized Assessments

that does not invalidate results, practice tests for the SESAT, Primary, Intermediate, and Advanced levels of Stanford 10 are available. The Advanced level practice test may also be used for the TASK levels. Stanford 10 practice tests are very brief and are designed to familiarize the test administrators and students with the test's content and format.

If available, this practice test should be administered one week before the actual assessment is scheduled to occur. This practice allows both students and test administrators to learn to follow the directions and fill in the answer spaces correctly. Moreover, the practice test enables students to understand how to respond to the different types of Stanford 10 items. Finally, because norm groups used practice tests during the standardization process, norm-referenced scores from Stanford 10 are most accurate when students have taken the practice tests prior to taking the actual Stanford 10.

Preparing Test Administrators for Testing

Persons administering a standardized assessment should familiarize themselves with the materials and procedures well in advance of the testing periods. If permitted by the school district, they should carefully read and examine the directions for administering, student test booklets, answer documents, and any other materials provided for testing.

Scheduling Testing Periods

Testing periods should be scheduled at a time that encourages maximum student performance. Scheduling test periods immediately before or after a vacation, on the days of school events, on Monday mornings, or on Friday afternoons is discouraged. It is recommended that no more than two subtests be administered in one sitting and that a brief rest period be given between subtests. The Stanford 10 directions for administering recommends rest periods of approximately 10 minutes.

One of the features of Stanford 10 is that it is an untimed test. A time schedule for each subtest based on research from standardization is provided for planning purposes. Test administrators should plan for enough time for students to complete the test.

Selecting a Testing Environment

The environment selected for administering Pearson's educational assessments should parallel the classroom environment as closely as possible. Standardized educational assessments should be administered to classroom-sized groups of about twenty-five students. In larger groups, the test administrator should have one assistant for every 25 students (see Table 2). Because students seem to perform better when they are familiar with the test administrator, the persons administering the tests should, if possible, be the students' own classroom teachers. Any people who assist the test administrator should also be familiar to the students.

Table 2. Administrators and Proctors Required for Test Administration.

Number of Students	Administrators Required
1–25	1 administrator
26–50	1 administrator and 1 proctor
51–75	1 administrator and 2 proctors
76–100	1 administrator and 3 proctors

Practices During Testing

During testing, the test administrator should follow the directions for administering exactly as written. The directions for administering typically include information about the following.

Testing Materials

Persons administering educational assessments may be required to supply certain materials that students need to complete the assessment. For example, most educational assessments require test booklets, answer sheets, soft-lead pencils, and scratch paper. Certain levels of Stanford 10 allow students to use place markers to help keep their place on the page or in the test booklet, rulers, mathematics formula reference sheets, and calculators, if allowed for mathematics subtests.

The directions for administering may also specify materials required by the persons administering the assessment. Extra materials required by the students should be kept on hand and distributed as necessary. A “DO NOT DISTURB” sign should be available to hang on the door to the testing room. Although Stanford 10 is standardized as an untimed test giving students the opportunity to respond to every item without a time limit (Case, 2003c), the directions for administration recommend that a digital watch or a clock with a second hand be available for the administrator.

The Testing Room Environment

The physical conditions of the room should be carefully controlled during testing. Students should be seated in an arrangement that prevents them from seeing the work of other students. The students’ desks and work spaces should be clear with sufficient room for the test booklets and answer sheets. Cover any materials and information on the walls, bulletin boards, and chalkboards that might help students to answer

Administration Practices for Standardized Assessments

questions. The testing room should have sufficient light and ventilation and be free from noise.

Reading Directions

While administering the assessment, the test administrator should read the directions exactly as written, using a natural, deliberate tone of voice and manner while enunciating clearly. The directions should not be repeated unless a mistake is made in the initial reading or the directions for administering say to do so. The test administrator may attempt to clarify any directions that students do not understand and repeat the instructions, if necessary. Except in the case of sample and practice items, the test administrator should not answer questions about the correct responses to specific test items.

Observing Testing

During testing, the administrator and proctors should move around the room to observe the testing situation and to make certain that the students are following the directions, working on the correct tests, and marking the responses correctly on the answer sheet. The test administrator and assistants should be visibly engaged in the testing by observing all aspects of the process.

If a student has trouble understanding what to do, the test administrator should try to explain the instructions quickly and quietly, but must not assist the student with responding to any specific items. Moreover, students must not help one another.

Random Responses and Changing Answers

Random marking of answer documents should be discouraged. Instead, students should be encouraged to respond to items for which they think they know the correct response. When changing answers, students should be reminded to erase the original mark thoroughly.

Attemptedness Rules

Students should be encouraged to answer all the questions and finish the test. However, certain criteria are used to determine whether a student actually attempted to complete a subtest. Attemptedness rules require students to respond to a specific number of items to receive a valid score. If students do not meet the criterion for attemptedness for a subtest, they receive the designation DNA (“Did Not Attempt”) for that subtest, and the subtest does not contribute results to any summary reports.

The attemptedness rules for the standardized achievement tests published by Pearson vary by state or program. However, each rule checks that a student has responded to a minimum number of items selected from throughout the subtest. The specifications for the attemptedness rules also depend on whether the subtest is completed in a single sitting or multiple sittings.

Administration Practices for Standardized Assessments

Collecting Testing Materials

At the end of a test sitting, test administrators should immediately collect all of the test booklets, answer sheets, and other testing materials from the students. Doing so is essential to maintaining the security of the testing materials. Moreover, this practice ensures that the testing materials for individual students are kept together so that they can be more easily accounted for and checked after test administration.

Procedures Following Test Administration

After testing, the persons administering the assessment should take the following steps to preserve the integrity and validity of student results. Because Pearson's assessment products and individual subtests may require different post-test procedures, test administrators should follow the information provided in the directions for administering accompanying each specific test.

Checking Answer Documents

The test administrator should examine each answer document and make certain that all erasures are complete and that there are no stray marks. Faint or poorly marked answers should be darkened so that they can be read during scanning. Marks must be made with a soft-lead pencil rather than by ballpoint pen, felt-tip pen, hard pencils, or colored pencils. Test administrators should neither alter nor change student answers. Check for stray marks in bar code or sheet registration areas, as these marks can cause test documents to be misidentified or rejected.

The identifying information on each student's answer documents should be checked for accuracy and completed with required information where necessary as specified by the directions for administering.

If students have been administered a Braille or large-print test, their responses must be transferred to machine-scoreable answer documents. Test administrators should also follow any local directions that they have been given.

Packaging Answer Documents for Machine Scoring

Pearson answer documents and other machine-scoreable testing materials, such as test booklets, should be organized and assembled according to the directions for administering with the proper Scoring Service Identification Sheets. Documents should be carefully packaged and shipped according to the Instructions for Packaging and Shipping Answer Documents. The shipping method should be prepaid with proof of mailing to minimize transit time.

Administration Practices for Standardized Assessments



Other steps for preparing answer documents for shipment include:

- checking for damage to the machine-scoreable testing materials, such as bent corners;
- using paper bands to group documents, rather than rubber bands or string (which may inadvertently cut them and make them potentially unscorable);
- checking that the correct number of documents is returned for scoring; and
- helping to eliminate the possibility of a shipment being lost or split by minimizing the number of packages sent to Pearson.

Additional specific instructions for packaging answer documents are provided by Pearson's product catalogs. For further assistance, contact Pearson Customer Service by phone at 1-800-211-8378 or e-mail at Customer_Service@PearsonAssess.com.

Conclusion

The test administration practices discussed in this report are based on the accepted codes and standards of leading educational organizations, the latest findings of education researchers, and the experience accumulated by Pearson since the beginning of the previous century in developing standardized assessments. By following these practices as closely and completely as possible, a test administrator enables students to demonstrate accurately what they know and are able to do.

However, obtaining accurate measures of a student's academic achievement is not merely an end in itself. In the current era of accountability in education, data from standardized achievement tests frequently serve as the principal source of evidence for evaluating the success of districts, schools, and even the teaching methods of individual educators. To accurately evaluate our nation's students—and thereby contribute to improving our nation's education systems—it is incumbent on test administrators to prepare for each specific assessment that they administer and follow test administration practices to the best of their ability.

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1999). *Standards for education and psychological testing*. Washington, DC: American Educational Research Association.
- Case, B. J. (2003a). *Accommodations on Stanford 10 for limited English proficient (LEP) students*. San Antonio, TX: Pearson Inc.
- Case, B. J. (2003b). *Accommodations on Stanford 10 for students with disabilities*. San Antonio, TX: Pearson Inc.
- Case, B. J. (2003c). *It's about time*. San Antonio, TX: Pearson Inc.
- Mehrens, W. A. & Lehmann I. J. (1991). *Measurement and evaluation in education and psychology*. Fort Worth, TX: Harcourt Brace College Publishers.
- Millman, J. & Greene, J. (1993). In Linn, R. L. (Ed.), *Educational measurement* (pp. 335–366). Phoenix, AZ: The Oryx Press.
- National Council on Measurement in Education. (1995). *Code of professional responsibilities in educational measurement*. Retrieved from http://www.natd.org/Code_of_Professional_Responsibilities.html on February 27, 2004.
- Nitko, A. J. (2004). *Educational assessments of students*. Englewood Cliffs, NJ: Prentice Hall.
- Oosterhof, A. (2001). *Classroom applications of educational measurement*. Upper Saddle River, NJ: Merrill Prentice Hall.
- Phillips, S. E. (1993). *Legal implications of high-stakes assessments: What states should know*. Oak Brook, IL: North Central Regional Educational Laboratory.
- Thurlow, M. L., Elliott, J. L., & Ysseldyke, J. E. (1998). *Testing students with disabilities*. Thousand Oaks, CA: Corwin Press, Inc.
- Tindal, G., & Fuchs, L. (1999). *A summary of research on test changes: An empirical basis for defining accommodations*. Lexington, KY: University of Kentucky, Mid-South Regional Resource Center.

**Additional copies of this and related documents are available from:
Pearson Inc.
19500 Bulverde Road
San Antonio, TX 78259
1-800-211-8378
1-877-576-1816 (fax)
<http://www.pearsonassess.com>**