Stanford Diagnostic Reading Test 4

The 1995 Stanford Diagnostic Reading Test (Karlsen & Gardner, 1996) is the fourth edition of a test originally published in 1966. The SDRT4 is a group-administered diagnostic test designed to identify specific strengths and weaknesses in reading. It provides detailed coverage of skills in phonetic analysis, vocabulary, comprehension, and scanning. Because the test is intended for use with low achievers, it contains easier questions than most achievement tests do.

An interesting feature of this test is the way in which the items are ordered. Most diagnostic tests have the items within a subtest arranged in order of difficulty, from the easiest to the most difficult. The items in the SDRT4 are arranged so that difficult items are interspersed among easier items. The easier items cushion student performance and help alleviate the frustration experienced when students fail many items in a row.

There are six levels of the SDRT4, with one form at each of the first three levels and two alternative and equivalent forms at each of the upper three levels. The levels and the grades for which they are intended are shown in Table 22.1.

Sampled Behaviors

The SDRT4 can be group-administered by a classroom teacher. Four skill domains are sampled by the test, although not all domains are sampled at all levels. Subtests and skill domains sampled by the SDRT4 are reported in Table 22.2. Behaviors sampled are as follows.
**Phonetic Analysis**  This domain is sampled only at the red, orange, and green levels of the test (prior to grade 4.5). Items in this domain sample students' skill in associating letters and word segments with consonant and vowel sounds.

**Vocabulary**  This subtest appears at all levels of the test and is a measure of listening vocabulary, skill in identifying synonyms, and verbal classification. Knowledge of synonyms is assessed by asking students to choose the printed word that means the same as the tested word. Knowledge of the relationships among words is assessed by having students classify words into categories.

**Comprehension**  Comprehension is assessed using three kinds of materials: recreational, informational, and functional text. Four kinds of comprehension are assessed:

1.  **Initial Understanding.**  Students must recognize ideas and relationships that are directly stated.

2.  **Interpretation.**  Students must show skill in making inferences and predictions based on what is read.

3.  **Critical Analysis.**  Students must show skill in evaluating what is read.

4.  **Reading Strategies.**  Students are asked to show skill in recognizing and using reader strategies, text structures, and types of text.

**Scanning**  This subtest measures students' skill in scanning a text for important information. The new fourth edition of the SDRT includes three optional informal assessment instruments: a Reading Strategies Survey, a Reading Questionnaire (a
measure of attitudes toward reading, reading interests, and familiarity with concepts that appear in the comprehension subtest), and a measure of Story Retelling.

Scores

The SDRT4 is both norm-referenced and criterion-referenced. It can be used to assess a pupil's performance relative to the performance of others, and it can be used to pinpoint individual pupils' strengths and weaknesses in specific reading skills.

Students respond either in the test booklets or on machine-readable answer sheets. The test can, therefore, be either hand scored or machine scored. Six kinds of scores can be obtained: (1) Raw scores are obtained for each subtest and can be transformed into (2) "progress indicators," (3) percentile ranks, (4) stanines, (5) grade equivalents, or (6) scaled scores. Which scores are useful depends on the purpose for which the test has been administered. Progress indicators are criterion-referenced scores, whereas the other four derived scores are norm-referenced. Progress indicators are + or - indications as to whether a pupil achieved a predetermined cutoff score in a specific skill domain; they show whether a pupil demonstrates mastery of specific skills that are important to the various stages in the process of learning to read effectively. The manual reports that "in setting the Progress Indicator cutoff scores, the SDRT4 authors were guided by the relative importance of each skill to the reading process, by the location of these skills in the developmental sequence, and by the performance of students at different achievement levels on the items measuring these skills" (Karlsen & Gardner, 1996, p. 15). The manual
for each level of the SDRT4 includes an appendix that lists specific instructional objectives assessed by each level of the test.

The norm-referenced scores obtained by administering the SDRT4 can be used for a variety of purposes. The authors provide a detailed table in the test manual showing the recommended uses of each of the kinds of scores and the extent to which scores are comparable across subtests, forms, levels, and grades. The scores on the test can be used to make setting decisions, identify reading strengths and weaknesses, evaluate pupil progress, and identify trends in reading achievement at the class, school, and district level.

A number of reports can be generated from the SDRT4 by making use of the publisher's computer-scoring service. Examiners can obtain an individual diagnostic report, which contains a detailed analysis of the performance of a single pupil. They can also obtain a class summary report, which shows the average scores earned by the pupils on each of the subtests. It also provides an analysis of skill development for the class by indicating the number of students in the class who obtained a progress indicator of + and of -. Examiners can obtain a master list report, which consists of a listing of scores for all students in a class. They can obtain a parent report, classified specifically for sending test results home to parents. Additionally, they can obtain a pupil item analysis, showing the raw scores earned by a particular student on each subtest and cluster, as well as the student's response to each item.

Norms
In preparing this fourth edition of the SDRT, the authors wrote or rewrote all items. In selecting the standardization sample for the SDRT4, the authors used a stratified random-sampling technique. Socioeconomic status, urbanicity, ethnicity, and geographic region were the stratification variables. School-system data were obtained from the U.S. Office of Education's 1990 census tapes. Age and gender were not controlled in standardizing the SDRT4.

School districts within each of the stratified cells were invited to participate in standardization of the test. A random sample of consenting districts within each cell was selected. The SDRT4 was standardized during the fall of 1994 and spring of 1995. Four hundred school systems participated in the fall standardization (33,000 students). The test was standardized on about 60,000 students. The authors provide a table in their manual showing the relationship of sample characteristics to census characteristics. The numbers are a close match. There is no report of cross-tabulations. Thus, we don't know how many low-SES students were from urban versus suburban settings.

Reliability

Two kinds of reliability data are provided for the SDRT4. Data on internal consistency are provided for all students in the standardization sample. Data on alternate-forms reliability are reported for three levels of the test at which there are alternate forms (purple, brown, and blue). All but one of the internal-consistency coefficients exceed .80. The reliability of the Vocabulary subtest is .79 at grade 1. Alternate-forms reliabilities are generally lower. These range from .62 to .88.
Validity

Data are provided on content validity and criterion-related validity. As test items were written, they were reviewed by content experts, who made sure that the items were actually assessing the content objectives they were intended to assess. Measurement experts reviewed items for appropriate test-item properties, and more than twice as many items were tried as were retained. The item tryout phase of the standardization program was on 150 districts from 32 states; about 16,000 students participated. The authors appropriately indicate that judgments about content validity must ultimately be made at the local school level by comparing the content of the test to the content of the local curriculum. Criterion-related validity was established by correlating performance on subtests of the SDRT4 with performance on their counterparts on the SDRT3. There is a strong relationship between performance on the two editions.

Summary

The SDRT4 is a group-administered device that is both norm-referenced and criterion-referenced. The device was exceptionally well standardized and is reliable enough to be used in pinpointing specific domains of reading in which pupils demonstrate skill-development strengths and weaknesses. Validity for the SDRT4, as for any achievement measure, must be judged relative to the content of local curricula. The test is one of the more carefully designed and developed diagnostic reading measures available.