Test Review and Evaluation

Woodcock Diagnostic Reading Battery (WDRB)

A. Demographic Information

   Address of Publisher: 425 Spring Lake Dr Itasca IL 60143-2079

   Cost (1999 price data): $273 per complete kit; $37 per 25 recording forms; $45 per examiner's manual; $40 per norms manual; $175 per scoring and interpretation program (Windows or Macintosh)

   Purpose of Test: Provides a diagnostic test that assesses reading achievement and important related abilities.


B. Specific Judgments
   a. Standard conditions of administration and scoring: The WDRB is made up of 10 separate tests. Administration of all batteries takes approximately 60 minutes, while administration of the reading tests takes about 20-25 minutes. The test is administered in a standardized easel format, and a computer program for scoring is included. Both the presentation and the scoring program for this test are derived from the Woodcock-Johnson Psycho-Educational Battery-Revised (WJ-R), which was first developed in 1977. Many parts of the WDRB are taken from this test, and WDRB is aimed at educators interested in focusing specifically on the reading aspects of the WJ-R. Overall reliability in terms of administration and scoring seem high due to this pedigree.

   b. Normative Data used: 6,026 individuals ranging from ages 4 to 95 were administered all tests. The sample population included preschoolers not enrolled in kindergarten, K-12 students, college students, and an adult nonschool sample 14 to over 90 years of age. The sample population closely matches U.S. Census data and were gathered over a span of two years. This extensive sample indicates well-tested normative data.

   c. Reliability estimates: The WDRB was tested in several ways for stability and internal consistency, through split-half internal consistency procedures and test-retest techniques. The resulting correlation coefficients ranged from a high of .936 to a low
of .787. Clearly, even the lower coefficient is well above the cutoff of .60, which is generally seen as the minimum needed for reliability.

d. **Validity Estimates**: Evidence for the validity of the WDRB is presented for content, or criterion-related, and construct validity. Evidence for content validity lacks specific references, but is given in general terms such as, “expert opinion” (Rudman). Criterion-related validity includes thorough studies for elementary and secondary students, and correlation coefficients are given for the WDRB and other standardized tests such as the Stanford-Binet and the WISC-R. These coefficients range from .43 to .91, which raises questions about the criterion-related validity. Construct validity is demonstrated primarily in reference to other Woodcock tests, such as the WJ-R from which the WDRB is constructed. Overall, estimates given for validity raise some questions and need to be explored further.

C. **Overall Judgment**

This test does seem to have merit, although further investigation regarding validity needs to be done. The test is very reliable as seen from several key measures, but the lack of similarly compelling validity correlations is troubling. The test would best be used to measure specific reading traits and changes over time. An example of this would be pre-and post-test administrations to determine the effectiveness of a new curriculum or other teaching intervention.

**REFERENCES**: All information regarding this test is drawn from the following source:

Test Review and Evaluation

Graduate Record Examination – General Test

A. Demographic Information

   Address of Publisher: Educational Testing Service, Rosedale Road, Princeton NJ 08541

   Cost: Not given. $140 fee to take test.

   Purpose of Test: Designed to assess the Verbal, Quantitative and Analytical reasoning abilities of graduate school applicants

   Scores: 3: Verbal Reasoning, Quantitative Reasoning, Analytical Reasoning.

B. Specific Judgments
   a. Standard conditions of administration and scoring: The GRE is administered under tightly controlled conditions. Subjects must travel to authorized centers for administrations, and conditions there are strictly supervised to ensure standardization of test-taking. There are no verbal instructions, so all participants receive the same information. Breaks are structured, and there are no outside materials such as calculators permitted in the test-taking rooms. Scoring is computerized and centralized. Based on the information presented as well as my personal experience as a test-taker, standardization of both administration and scoring is as great as can be practically accomplished.

   b. Normative Data used: Not given. However, the GRE is a well-known and generally accepted test that is required for admission in virtually all major graduate programs throughout the United States. It can be assumed that normative data has been collected from tens of thousands of participants. In addition, scores are ranked relative to the performance of others, and so can always potentially change. Additionally, the form of the test itself is constantly updated, and the computerized version even adapts according to the test-takers responses. In essence, normative data is constantly updated and revised from session to session, and the lack of data given should not be cause for concern.

   c. Reliability estimates: Internal consistency using the Kuder-Richardson 20 generates reliability coefficients ranging from .90 for the verbal and quantitative sections to .86 for the analytical section. This indicates a strong degree of internal consistency.
d. **Validity Estimates:** The validity of the GRE as a predictor of graduate student success (as demonstrated by graduate GPA) is questionable. Validity coefficients relating GPA to GRE scores range in the area of .20 to .30. When combined with undergraduate GPAs, these coefficients increase to .32 to .56. A recent meta-analysis of validity studies reports coefficients of .30 to .36 (depending on which section of the test is being evaluated) when combined with the undergraduate GPA, which suggest, “moderate correlations between GRE scores and important criterion measures, including GGPA, comprehensive examination scores, and faculty ratings of student competence” (Kuncel, Hezlett), & Ones; 2001). The authors then conclude there is no evidence, “to support the position that admission decisions that rely on the GRE or UGPA will result in inferior and limited graduate students.” This misses the point of the main criticism of the GRE. The point is not that relying on the GRE only will dilute the pool of graduate students including “inferior and limited students,” but that relying only on the GRE will exclude potentially successful graduate students.

**C. Overall Judgment**

The question of graduate student admissions is difficult. It appears as though the GRE has little merit as a predictor of graduate student success and provides marginally useful information in attempting to resolve this difficult question. Given the high cost to applicants, its utility diminishes. It is essentially the best thing available at this time, and as such is in widespread use. On their own website, the ETS states: “Test scores should always be used along with other sources of information such as course grades, letters of recommendation, personal statements, samples of academic work or professional experience.” Given the low correlation coefficients between GRE scores and GPA, this is a prudent recommendation, to say the least.

**REFERENCES:**


