Woodcock Johnson-III: Tests of Achievement
(WJ-III)

**Publisher/Date:**

**Purpose:**
- Individually-administered, comprehensive test of academic achievement and scholastic aptitude for individuals ages 2-00 to 90-00. The achievement test includes two parallel-forms (Forms A & B). A new brief version (Form C) was developed and available.

**Provides:**
- Measures of academic achievement in broad areas of reading, math, written language, oral language, and knowledge. The test breaks these areas down into more focused performance clusters addressing basic reading skills and reading comprehension, math calculation and applied math reasoning, basic writing skills and written expression, oral expression and listening comprehension, and academic knowledge. Additional clusters assess general academic skills, fluency, and applications. Additional measures of phoneme/grapheme knowledge and several supplemental/diagnostic subtests (Story Recall—Delayed, Handwriting Legibility Scale, Writing Evaluation Scale) are also available. The test affords multiple interpretive scores for analyzing and reporting student performance.

**Standardization Issues:**
- The cognitive and achievement measures were co-normed on 8,818 individuals consisting of 1,143 preschoolers, 4,784 school-age children (Kindergarten-12th grade), 1,165 college students, and 1,843 adults. The norm group included students attending public, private, and home-schooling, students with disabilities, and English language learners who had at least one year or more of experience in English-speaking classes. In all, 13 socioeconomic-status variables were accounted for, as well as 10 specific community and individual variables. Normative recalculation of the WJ III’s existing norm base was matched to the 2005 U.S. Census projections (Normative Update, or NU). The US Bureau of the Census had reported changes in the general population characteristics and these changes were reflected in the newly-revised normative statistics. Early reports of changes to the U.S. population—such as geographic shifts, increased urbanization, and greater percentages of young children and certain minorities in the overall population, to name only a few—provided the impetus to evaluate how such changes would impact the calculation of
norms. Based on the 2005 census data specifications, 8,782 of the original standardization subjects were selected for inclusion in the WJ-III NU norms. The 2005 norms are reflected in the *WJ III Compuscore NU*® and *Profiles Program* and the *WJ III NU Technical Manual*. Care in the norming and standardization of the instrument were reported as a “centerpiece” of the technical revision from its predecessor, the WJ-R.

**Reliability and Validity Issues:**

- Because of the sheer statistical complexity and number of tests, composite scores, and age groupings involved for which reliability data could be reported, this review will focus on providing an overall evaluation of the test’s reliabilities. Internal consistency reliabilities are high for both individual tests (.80s and .90s) as well as clusters (.90s). The extensive and detailed Technical Manual provides reliability estimates for any individual tests or clusters that the user may wish to examine. Internal reliability for Total Achievement was very high (.98). Several stability studies (test-retest) for reliability of scores over time included time-intervals of less than 1 year, 1-2 years, and 3-10 years and yielded acceptable median reliabilities (.70s to .90s). Another stability study using a 1-year interval yielded encouraging results, in the .80s and .90s. Interrater correlations for even the most subjective measures in the total battery (Writing Samples, Writing Fluency, and Handwriting from the Achievement Battery) were very high (upper .90s), suggesting the confidence in clarity of administrative and scoring procedures. Reliability between forms A & B is high across the age groups (.80-.96). A variety of validation procedures and correlational studies confirmed the structure of the test. The cognitive clusters intended to predict achievement correlate well with associated achievement cluster scores on the Tests of Achievement (.70-range), which is higher than those typically found between other ability measures, indicating good predictive validity. The Total Achievement score correlates well with other achievement tests, for both normal and clinical groups.

**Additional Points:**

- Users of the WJ-III are urged to obtain and use the Normative Update program. In addition to updated and more representative norms, the *WJ III NU Compuscore* and *Profiles Program* offers a number of new interpretive procedures and report options.

- The WJ-III is a complex instrument and many test users may benefit from advanced training, as many may find themselves overwhelmed with the information provided. Users of the WJ-III may find Nancy Mather and Lynne E. Jaffe’s reference publication, “*Woodcock-Johnson III: Reports, Recommendations, and Strategies*” (John Wiley & Sons, Inc., 2002) very helpful in their understanding and interpretation of the test. Their 516-page
resource also links cognitive and achievement test performance to empirically-based intervention strategies and instructional suggestions.

• Users may also find ASB#8, “Educational Interventions Related to the Woodcock-Johnson III Tests of Achievement” to be a helpful resource. This 40-page document is aligned with the Normative Update, and provides evidence-based intervention strategies linked to performance on the WJ-III Tests of Achievement. The document can be downloaded from: www.assess.nelson.com/test-ind/wj-3.html

• The technical manual’s analysis with regard to gender did not find any significant differences between males and females.

• Only a few subtests of the achievement battery require subjectivity in scoring (Writing Samples, Writing Fluency, and Handwriting), and it may be helpful to score these tests jointly by two raters.

• The test may very appropriate for assessing gifted students because of the lack of penalties for time and the very wide item difficulty range.