Woodcock Johnson-III: Tests of Cognitive Abilities
(WJ-III)

Publisher/Date:

Purpose:
• Individually-administered test of general and specific cognitive functioning, scholastic aptitude, and oral language for individuals ages 2-00 to 90-00.

Provides:
• Measures of general intellectual ability including a brief intellectual ability estimate (BIA), and standard- and extended-battery versions of general intellectual ability (GIA). In addition, numerous broad and more-specific cognitive ability and diagnostic cluster scores can be obtained. The test is aligned with the Cattell-Horn-Carroll (CHC) stratified model of intellectual abilities. The cognitive battery is designed to work in tandem with the parallel-developed, “Tests of Achievement,” which allows for direct comparisons and analysis of performance. Numerous clinical comparisons, analyses of performance, and reporting of performance options are possible, and an enormous amount of information is available from the tests.

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®.
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. 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. 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.

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 offer 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.

- Other cognitive measures may be more appropriate and used with more confidence for younger (preschool-age) children. The validity evidence for this group may be less strong as for the school-age population and the effects of normalizing score distributions—and extremes of ability—may make inferencing for this group somewhat tentative.

- A study by Edwards & Oakland investigating the WJ-III’s General Intellectual Ability (GIA) and cognitive factors, and validity between Caucasian and African American students indicated that while WJ-III scores appear to have comparable meaning for both groups, their mean
scores did differ. The mean GIA for Caucasian children was 104.2, vs. 93.3 for the African American children (about 11-points).

- Edwards & Oakland found mean scores on subtests from the Standard Battery (Tests 1-7) to range from 101.2 (Visual Matching) to 103.7 (Sound Blending) for the Caucasian group, and 90.9 (Verbal Comprehension) to 98.5 (Visual-Auditory Learning) for the African American group. Not surprisingly, the greatest difference between the groups was seen on the measure of Crystallized ability (Verbal Comprehension), where the mean Caucasian score was 103.3 vs. the African American mean of 90.9 (12.4 points).

- Edwards & Oakland’s factor-analysis further suggested that each individual test in the standard battery contributed to the overall intellectual estimate in about the same weighting for both groups. The correlations between groups was remarkably similar, with the only exception pertaining to the test measuring short-term memory (Numbers Reversed), suggesting that it may not be “as strong” a measure of general ability (“g”) for African Americans (.59) as for Caucasian Americans (.69).

- Finally, their findings suggested that correlations between the GIA and the achievement clusters for reading, math, and written language (as well as for each academic test on the Achievement Battery) do not seem to differ significantly by ethnicity.

- Rizza et al., conducted work similar to Edwards and Oakland in doing a profile analysis of gifted students vs. non-gifted students and found that while the gifted group scored higher on average across CHC factors than the non-gifted group, patterns of variance on the seven cognitive factors between the groups was not present. Their findings did support the use of the WJ-III as a measure for gifted students on the basis of these findings as well as the wide ability-range span of items on the test. The groups in the study only comprised of Caucasian-American and Asian-American children, however, and in light of the findings of Edwards & Oakland, more information on it’s applicability for gifted-identification with African American students would be helpful.

- Flanagan & Ortiz (2001, *Essentials of Cross Battery Assessment*) have analyzed and classified individual WJ-III Cognitive subtests by degree of cultural loading and linguistic demand (Culture-Linguistic Interpretive Matrix, or “C-LIM”), to aid in nondiscriminatory interpretation of these test scores. Use of the C-LIM assists practitioners in selecting tests that provide more accurate or fairer estimates of true ability as well as to systematically evaluate the influence of cultural and linguistic factors on test performance. Users of the WJ-III Tests of Cognitive Ability with culturally and linguistically diverse populations are strongly encouraged to
familiarize themselves with the work of Flanagan, Ochoa, and Ortiz and the appropriate use of the C-LIM.

- Scoring procedures do allow for responses to be scored as correct on certain tests when expressed in languages other than English (though the examiner must understand the dialect to ensure that the response is correct). The manual also outlines recommended accommodations for a variety of situations (handicaps). Examiners should be aware that the practicality and validity associated with these procedures has not been established.