**[16073134]**

**Woodcock-Johnson® III Diagnostic Supplement to the Tests of Cognitive Abilities.**

**Purpose**: Designed to expand "the diagnostic capabilities of the [Woodcock-Johnson III Tests of Cognitive Abilities] for educational, clinical, or research purposes."

**Population**: Ages 2-90+.

**Publication Dates**: 1977-2003.

**Acronym**: WJ III(r) Diagnostic Supplement to the Tests of Cognitive Abilities.

**Scores, 11**: Memory for Names, Visual Closure, Sound Patterns-Voice, Number Series, Number Matrices, Cross Out, Memory for Sentences, Block Rotation, Sound Patterns-Music, Memory for Names-Delayed, Bilingual Verbal Comprehension-English/Spanish.

**Administration**: Individual.

**Price Data**: Available from publisher.

**Time**: (5-10) minutes per subtest.

**Comments**: Tests from the Diagnostic Supplement can be combined with other tests from the Woodcock-Johnson III (15:281) Tests of Cognitive Abilities Standard and Extended Batteries and the WJ III Tests of Achievement to provide 14 additional interpretive clusters.

**Authors**: Richard W. Woodcock, Kevin S. McGrew, Nancy Mather, and Fredrick A. Schrank.

**Publisher**: Riverside Publishing.

Review of the Woodcock-Johnson(r) III Diagnostic Supplement to the Tests of Cognitive Abilities by TIMOTHY SARES, Director of Testing and Psychometrics, American Registry of Diagnostic Medical Sonography (ARDMS), Rockville, MD:

DESCRIPTION. The Woodcock-Johnson III (WJ III(r)) Diagnostic Supplement to the Tests of Cognitive Abilities consists of 11 testing measures intended to augment the WJ III Tests of Cognitive Abilities (WJ III COG). The supplemental tests are numbered 21-31 to coincide with tests numbered 1-20 on the WJ III COG. The Supplement is intended to provide further diagnostic capability for the standard and extended batteries of the WJ III COG and is especially useful as a complement in the assessment of bilingual persons and young children. The fundamental criteria for developing cognitive abilities in the Supplement are derived from the Cattell-Horn-Carroll (CHC) theory of cognitive abilities and are described in the WJ III COG examiner's manual (Mather & Woodcock, 2001) The broad CHC abilities covered on one or more of the supplemental tests are Long-Term Retrieval (Glr), Auditory Processing (Ga), Fluid

Reasoning (Gf), Processing Speed (Gs), Short-Term Memory (Gsm), Visual-Spatial Thinking Comprehension (Gv), and Comprehension-Knowledge (Gc).

Several cognitive ability and broad CHC clusters can be evaluated in combination with the WJ III COG Standard and Extended Batteries. For example, Supplementary Test 21: Memory for Names, Test 22: Visual Closure, and Test 27: Memory for Sentences can be combined with the WJ III COG Standard Battery Test 1: Verbal Comprehension, Test 6: Visual Matching, and Test 8: Incomplete Words to derive a General Intellectual Ability-Early Development (GIA-EDev) scale to assess young children or other individuals functioning at the preschool level. Two additional cognitive ability clusters identified are General Intellectual Ability-Bilingual (GIA-Bil) and Broad Cognitive Ability-Low Verbal (BCA-LV). The GIA-Bil is intended for English-dominant bilingual individuals. The BCA-LV is intended as a broad measure of cognitive ability requiring relatively minimal verbal expression or exchange.

The administration times may vary for the WJ III Diagnostic Supplement Tests as they adapt to the individual test-taker's progress. The manual states that supplemental GIA-EDev tests for preschool-age children generally require about 5-10 minutes per test to administer. Some tests in the supplemental battery may be suitable for assessing individuals in situations where attention span or response style is a concern.

The tests can be administered to individuals varying in ages from 2 to 90-plus years and is suitable for preschool, K-12, college and university undergraduate and graduate student educational levels, and also for adults over 14 years old not attending school. The following tests: Test 23: Sound Patterns-Voice, Test 27: Memory of Sentences, and Test 29: Sound Patterns-Music, require audio recording and audio equipment for administration. Test 26: Cross Out requires a stop watch, digital watch, analog watch, or clock with a second hand. Test 31: Bilingual Verbal Comprehension-English/Spanish is intended for administration by an examiner fluent in Spanish.

DEVELOPMENT. Items were written and reviewed by experienced teachers and psychologists serving as subject matter experts. Their overriding objective was to measure the narrowly defined applicable CHC ability construct. Support for structural validity is reported through intercorrelations of test scores for the Diagnostic Supplement Tests. Norms for the WJ III Diagnostic Supplement batteries are developed on test data from a national sample of over 8,000 individuals ranging in age from 2 years to over 90 years. The WJ III Diagnostic Supplement manual directs questions regarding the norming sample to the WJ III technical manual (McGrew & Woodcock, 2001). The norming sample was also administered the WJ III COG and the WJ III Tests of Achievement (WJ III ACH) and selected to resemble the demographic and community characteristics of the general U.S. population relative to geographic region, community size, gender, race, Hispanic origin, and type of school or college.

The test manual refers very briefly to the item selection criteria that follow stringent statistical fitting criteria of the Rasch model but provides very little useful basic information about the process, which leaves the reader questioning its sufficiency. As fit analysis criteria could signify item bias as well as extraneous content, no information is provided on what factor analysis, if any, was conducted to coincide with the fit analysis.

Test score intercorrelations and associated sample sizes are presented for the following age group ranges: 2-3, 4-5, 6-8, 9-13, 14-19, 20-39, and 40 and over. A rationale for selecting correlations for these age ranges is not provided in the testing manual. The correlations are based on sample sizes ranging from 42 (2-3 years) children to 1,984 (9-13 years). The intercorrelation of Block Rotation and Visual Closure in the smallest sample (n = 42, 2-3-year-olds) is fairly low (r = .17) and is even lower (r = .10) for the largest sample (n = 1,984) of 9-13-year-olds.

Approximate Age Equivalent (AE) scores and Grade Equivalents (GE) scores are shown in the Diagnostic Supplement on total number correct for all tests except Test 30: Memory for Names-Delayed. AE and GE scores for Test 30 can only be derived from the software-scoring program. AE and GE score estimates for the other batteries are shown alongside the score sheets for each test.

TECHNICAL. Overall internal construct validity is weakest for the four lower age groups (2-13 years) and strongest for the higher age groups (14 and older). For instance, low correlations are recorded between Delayed Memory for Names and Visual Closure (r = .24) at ages 40+, (r = .06) at ages 9-13, and Block Rotation (r = .06) at ages 6-8. The intercorrelations for Visual Closure produced somewhat inconsistent results across age groups. For age groups above 9 years old, the intercorrelation coefficients on visual closure are lowest for the 9-13-year age group. However, the intercorrelations for Visual Closure on the three youngest age groups reveal stronger relationships (than those shown for the 9-13-year-old age group) for all measures except Memory for Sentences (6-8 years) and Delayed Memory for Names (4-5 years). The cross-correlations between scores of the 2-3-year-old group on the five supplemental tests reported are for Memory for Sentences and Visual Closure (r = .52), which are much higher than those reported in the 9-13-year-old group (r = .12).

The reliability estimates are derived from Rasch (Wright & Stone, 1979) standard error of measurement person measures, and split-half computational methods corrected for test length using the Spearman-Brown correction formula. The median reliability values are reported for each diagnostic test across all age level samples. A Rasch method is implemented for computing the reliability estimates for Test 26: Cross Out, Test 27: Memory for Sentences, and Test 28: Block Rotation, as split-half computational procedures are inappropriate for these tests. The reliability estimates for these tests are computed from unique estimates of the standard errors of measurement (SEM) for each person's raw score using the Rasch method as described in the WJ-R technical manual (McGrew, Werder, & Woodcock, 1991, p. 95).

The fact that median reliability coefficients are consistently high comes as no great surprise as the process of inputting all incorrect scores for measures above the ceiling or correct ones below the basal eliminates potential error variance for those responses. Reliability statistics are reported for as few as 17 individuals at age 4 for the number series construct (r11 = .89). The median reliability coefficient for Test 25: Number Matrices is .91 but is lowest for 5-year-olds (r11 = .63, n = 173). The reliability estimate for Test 26: Cross Out (median r11 = .76) is lowest for 5-year-olds and 18-year-olds (r11 = .63).

For ages 2 through 10, the standard score deviations are highest for the Number Series construct. For ages 11 and above, score standard deviations are about equally high for both the Number

Matrices and Number Series constructs. The technical supplement of the testing manual reports two standard errors of measurement, one for the W-scale centered around 500 and a second for the standard score (SS) scale centered on 100 with standard deviation (SD) of 15.

Generally speaking, the confidence intervals are widest for Test 22: Visual Closure, Test 24: Number Series, Test 25: Number Matrices, and Test 26: Cross Out. The relatively wide confidence intervals for these measures increase as observed scores move out further from the scale centers for these measures. This occurs because Rasch standard errors of measurement are generally smaller near the scale centers. Although the median SEM and SD values are not presented with the median reliability coefficients, their values appear fairly large in regard to the score distributions. Overall, the SEM values differ greatly by test and by age group. For example, the SEM values range from 5.94 to 12.29 for Visual Closure and 3.53 to 12.77 for Number Series.

SUMMARY. The WJ III Diagnostic Supplement provides additional measures of CHC cognitive abilities intended for those seeking greater diagnostic assessment of an individual's cognitive skills. The tests are normed across a wide range of ages and educational levels. Test developers have given consideration for measurement biases on women, individuals with disabilities, and cultural and linguistic minorities. The tests appear as supplements to the WJ III COG and can be administered in as little as 3 minutes for Test 26: Cross Out, and range in length from 24 items on Test 28: Block Rotation, to 72 items on Test 21: Memory for Names, which can be cut short at one of two cutoff points.

The test battery provides additional diagnostic instruments intended to supplement the original WJ III COG test battery and provides support for their construct validity. The diagnostic supplemental tests also provide another mechanism for examining both narrow and broad CHC cognitive abilities that are applicable to a wide range of ages and educational backgrounds.

REVIEWER'S REFERENCES

Wright, B. D., & Stone, M. H. (1979). Best test design. Chicago: MESA Press.

Mather, N., & Woodcock, R. W. (2001). Examiner's manual: Woodcock-Johnson III Tests of Cognitive Abilities. Itasca, IL: Riverside Publishing.

McGrew, K. S., Werder, J. K., & Woodcock, R. W. (1991). WJ-R technical manual. Itasca, IL: Riverside Publishing.

McGrew, K. S., & Woodcock, R. W. (2001). Technical manual: Woodcock-Johnson III. Itasca, IL: Riverside Publishing.

Review of the Woodcock-Johnson(r) III Diagnostic Supplement to the Tests of Cognitive Abilities by DONALD L. THOMPSON, Professor of Counseling and Psychology (Adjunct), Troy University-Montgomery, Montgomery, AL:

DESCRIPTION. This review will provide an overview of the Diagnostic Supplement (DS) and examine its utility and empirical merits in terms of the test administration and interpretation process, and evidence for the test validity and reliability. The manual and various promotional materials for the Diagnostic Supplement (DS) indicate that it is intended to extend the diagnostic capabilities of the Woodcock-Johnson III Tests of Cognitive Abilities. The DS includes 11 new tests, several new clusters, and new interpretive procedures that are intended to improve and enhance the diagnostic usefulness of the cognitive battery. By using the additional tests and clusters, the DS can provide a more complete evaluation of an individual's relative cognitive strengths and weaknesses. The DS is intended for use in educational, clinical, and research settings. It can be used with individuals who are age 2-90+. The Diagnostic Supplement is particularly useful for the assessment of bilingual individuals and young children. It is also recommended for use in any situation when a language-reduced cognitive ability score is required. In most cases, the DS will be used as a supplement to the Tests of Cognitive Abilities rather than a stand-alone test.

The complete WJ III contains two separate and distinct, conormed batteries: the WJ III(r) Tests of Achievement and the WJ III Tests of Cognitive Abilities. Although this review focuses only on the DS, it is important to know the characteristics of the complete WJ III battery to get a full understanding of the DS. Because this is beyond the scope of this review, the reader is referred to the most recent Mental Measurements Yearbook reviews of the Woodcock-Johnson III battery available in the 15th MMY (Cizek, 2003; Sandoval, 2003).

The descriptive materials available on the website for the WJ III Tests of Cognitive Abilities indicates that the test "is based on the Cattell-Horn-Carroll (CHC) theory of cognitive abilities, which combines Cattell and Horn's Gf-Gc theory and Carroll's three-stratum theory" (Riverside Publishing, 2004). The Diagnostic Supplement includes 11 new tests that increase the breadth and depth of coverage for several broad and narrow abilities as defined by CHC theory. The subtests include: Memory for Names, Visual Closure, Sound Patterns-Voice, Number Series, Number Matrices, Cross Out, Memory for Sentences, Block Rotation, Sound Patterns-Music, Memory for Names-Delayed, and Bilingual Verbal Comprehension-English/Spanish.

The subtests of the DS can be combined with various subtests of the Standard Battery and/or Extended Battery of the WJ III Tests of Cognitive Abilities for selective testing to examine specific factors in three broad areas including intellectual ability (e.g., General Intellectual Ability-Bilingual), broad CHC clusters, and narrow CHC clusters. The manual identifies three intellectual ability clusters, two broad CHC clusters, and seven narrow CHC clusters. Logically, it would appear that test users might selectively combine various subtests to address specific questions for a particular client.

The administration of the 11 subtests is generally easy and straightforward. All subtests are presented using an easel from which both the examiner and examinee work. The examiner easel sheets provide the administration and scoring instructions for each subtest item. Three subtests are administered using a prerecorded tape provided with the test package. The test materials are of the highest professional quality, and the administration instructions and scoring procedures are described in excellent detail.

Scoring of the 11 subtests is done by hand during the test administration. Many subtests require that a ceiling level be established for the examinee, and several subtests require the establishment of a basal level. Although the subtests are scored by hand, the interpretation requires the use of the Compuscore(r) and Profiles software program that is included. This process is required because there are complicated scoring factors that, according to the manual, cannot be built into the printed score tables. The test interpretation is based principally on the relative clusters of tests because (as the online promotional material states), "cluster interpretation results in higher validity because scores are based on a broad, multifaceted picture of each ability instead of on a single, narrow ability" (Riverside Publishing, 2004).

DEVELOPMENT. As was noted earlier, the DS is an extension of the WJ III Tests of Cognitive Abilities, and as such was developed using the Cattell-Horn-Carroll (CHC) theory that was used in developing the other two components of the WJ III battery. The subtests were developed to provide measures that address the cognitive abilities conceptualized by that theory. The manual that accompanies the DS does not provide much detail regarding the test development process or how items were selected; however, it appears that the test development occurred concurrently with the work done in preparing the latest version of the WJ III Tests of Cognitive Abilities published in 2001. Other reviewers have described the development process in very positive terms, emphasizing that the effort used sound psychometric practices in the process (Cizek, 2003; Sandoval, 2003).

TECHNICAL. The standardization section of the manual provides information describing how the norm sample matches the intended user population. The DS was developed and normed using the same normative data that were used for the latest edition of the Woodcock-Johnson III (2001), and includes 8,818 individuals from more than 100 geographically diverse communities in the U.S. that theoretically represent the U.S. population from ages 2-90+. The norm sample includes college and university undergraduate and graduate students and a number of adults (nonstudents). The two MMY reviews cited earlier are generally very positive with regard to the norming and standardization of the WJ III, and this reviewer concurs that the development effort appears to represent good psychometric practices in test development.

The overall reliability and validity data provide empirical support for the DS when it is used in appropriate situations. Reliability data were generated using two primary means. Split-half reliabilities were calculated by age for eight of the subscales, whereas the reliability estimates for three subscales were calculated (also by age) using the Rasch analysis procedures. Split-half reliability coefficients for the eight subscales were in the acceptable range of .80 or above (five were <.90), and the Rasch procedures suggested acceptable reliability of the three scales examined in this manner. However, it was noted that some reliability coefficients for younger test subjects (ages 2 to 5) were lower. The reliability of 15 defined clusters was also examined, and all of the coefficients were above .80 (many were above .90).

Test validity was examined using a variety of methods including content, concurrent, and construct methods. Data reported in the test manual describe factor analyses that support the theoretical (CHC) factor structure indicating that hypotheses based on the theory were supported by the test results. Although it focuses on the Tests of Cognitive Abilities and does not specifically address the DS, one recent external study by Lohman (2003) provides additional

empirical support for both the concurrent and construct validity of the WJ III Tests of Cognitive Abilities and the CHC theoretical model. Concurrent validity evidence is presented in the manual indicating the correlations between the DS subscales and clusters scores and other established measures of cognitive ability (e.g., WPPSI-R, WISC III, and WAIS III). The overall reliability and validity data provide empirical support for this instrument when it is used in appropriate situations.

COMMENTARY. Although the Woodcock-Johnson Tests of Achievement has long been a favorite of educators and clinicians, the Tests of Cognitive Abilities has never achieved the popularity of the Wechsler and Stanford-Binet for assessing general intellectual ability/IQ. However, with the addition of the Diagnostic Supplement to the Tests of Cognitive Abilities, this may change. Clearly the utility of the DS and its empirical qualities make it a test that should be considered by those who work with special populations/situations including bilingual assessment, early childhood assessment, or other purposes when a language-reduced cognitive ability score is needed. The administration and scoring procedures are easy to learn and simple to use. Because all parts of the WJ III are based on the same theoretical foundation, it potentially provides for more accurate and valid comparisons between different parts of the WJ III battery, and between and among the Cattell-Horn-Carroll (CHC) abilities than is possible if separately normed instruments are used for assessing ability and achievement (manual, p. 4). As with the other parts of the WJ III battery, the DS user may choose to administer only specific subtests in order to address particular questions about the test taker. Because of the flexible structure of the DS, test users have the option of choosing specific subtests to meet the needs of the client being assessed. Although some parts of the DS must be scored on-the-fly as the test is administered in order to establish basal and/or ceiling levels, the overall test interpretation must be done using the Compuscore and Profiles computer program. The scoring software works well and calculates scores using several variables, including time limits, test sections, age-equivalents, grade equivalents, and delay times.

There are some minor factors that detract from the overall excellence of the DS. Although the norm sample includes a significant number of college students and nonstudent adults, the nature of the stratification/sampling for these groups leaves questions as to whether they are truly representative of all college students/adults in the U.S. population. The manual that accompanies the DS provides little in the way of details regarding the test development process or how items were selected, but related information is included in the technical manual (McGrew & Woodcock, 2001) for the Woodcock-Johnson III. Although the DS can be purchased separately from other parts of the WJ III battery, and it is possible to use certain DS subtests for specific limited purposes independent of other parts of the battery, generally the DS should not be viewed as a stand-alone test. It is intended to be used as the title indicates, as a supplement to the Tests of Cognitive Abilities. A computer-based report program for the WJ III battery (Report Writer for the WJ III) is also available for purchase separately. This program generates reports that are comprehensive and will be useful in many counseling and placement situations; unfortunately, the results obtained from the DS are not supported by the program.

SUMMARY. The 2001 standardization of the Woodcock-Johnson III battery and the development of the Diagnostic Supplement have resulted in a well-developed and empirically sound instrument. It is easy to use and score and it has solid psychometric properties. One of the

DS subtests can be used to test bilingual individuals. Although this represents a major strength of the test, it also requires bilingual proficiency on the part of the examiner (or an ancillary examiner). The manual is well written and comprehensive and fairly detailed regarding the standardization process as well as the test administration and scoring procedures. The software for computer-based scoring is a major plus. The software works well and makes for convenience and readily available results. Overall, this reviewer finds the DS used in conjunction with the Tests of Cognitive Abilities to be a particularly good choice for assessing clients who are younger, bilingual, or have some language-related deficiencies.

REVIEWER'S REFERENCES

Riverside Publishing. (2004). The Woodcock-Johnson III product details. Itasca, IL: Author. Available online at: http://www.riverpub.com/products/clinical/wj3/product.html

Cizek, G. J. (2003). [Review of the Woodcock-Johnson III.]. In B. S. Plake & J. C. Impara (Eds.), The fifteenth mental measurements yearbook (pp. 1020-1024). Lincoln, NE: Buros Institute of Mental Measurements.

Lohman, D. F. (2003). The Woodcock-Johnson and the Cognitive Abilities Test (Form 6): A concurrent validity study. The University of Iowa. Available online at http://faculty.education.uiowa.edu/dlohman/pdf/cogat_WJIII_final_2col%202r.pdf

McGrew, K. S., & Woodcock, R. W. (2001). Technical manual. Woodcock-Johnson III. Itasca, IL: Riverside Publishing.

Sandoval, J. (2003). [Review of the Woodcock-Johnson III.]. In B. S. Plake & J. C. Impara (Eds.), The fifteenth mental measurements yearbook (pp. 1024-1027). Lincoln, NE: Buros Institute of Mental Measurements.