20 Assessing Diverse Populations With Nonverbal Measures of Ability in a Neuropsychological Context

Jack A. Naglieri and Tulio M. Otero

MEASUREMENT OF ABILITY

Assessment of ability for diverse populations of children and adults has been and continues to be one of the most important problems facing our profession. Typical IQ tests have used the familiar verbal, quantitative, and nonverbal format since Binet (in 1905) and Wechsler (in 1939) published their influential tests. The division of items by content was not based on a theory of verbal and nonverbal types of intelligences, but in fact the division into verbal and nonverbal scales was a practical partition to meet the need of testing individuals with different levels of education and English language skills. In fact, Yoakum and Yerkes (1920) wrote that the nonverbal (Army Beta) tests were used because a person could fail verbal and quantitative (Army Alpha) tests due to limited skills in English. To avoid "injustice by reason of relative unfamiliarity with English" (Yoakum & Yerkes, 1920, p. 19), these persons were tested with the Arm Beta (e.g., nonverbal tests) to ensure accurate measure of their ability. Rather than attempting to measure verbal and nonverbal intelligences, the Alpha and Beta tests were used to measure general ability.

GENERAL ABILITY

Traditional tests such as the Wechsler and Binet scales measure *general ability* using verbal, spatial, or quantitative test questions. The spatial tests (e.g., arranging blocks to match a simple design or assembling puzzles to make a common object) have been described as nonverbal even though Wechsler did not have any intention to measure a construct called nonverbal ability. In fact, Wechsler did not view verbal and nonverbal tests as measures of two types of intelligence despite the fact that for years his tests yielded Verbal and Performance (nonverbal) IQ scores. He argued that nonverbal tests help to "minimize the overdiagnosing of feeblemindness that was, he believed, caused by intelligence tests that were too verbal in content . . . and he viewed verbal and performance tests as equally valid measures of intelligence and criticized the labeling of performance [nonverbal] tests as measures of special abilities" (Boake, 2002, p. 396). Wechsler stated that "the subtests are different measures of intelligence, not measures of different kinds of intelligence" (1958, p. 64) and he "viewed verbal and performance tests as equally valid measures of intelligence" (Wechsler & Naglieri, 2006, p. 1). Furthermore, Naglieri (2003a, 2008a) wrote that the term *nonverbal* refers to the *content* of the test, not a type of ability, and that the goal is to measure general ability. The problem is, however, if tests of general ability require knowledge of language and quantitative skills, can these factors pose threats to the internal validity of a test of ability?

Those who have not had the chance to acquire verbal and quantitative skills due to limited opportunity to learn from a developmental or acquired neurological condition will likely do poorly in school but verbal and quantitative tests may not be a good reflection of their ability to learn after having had ample instruction. Traditional IQ tests require a considerable amount of information and skills. For example, some *Stanford-Binet Intelligence Scales–Fifth edition* (*SB-5*; Roid, 2003) Quantitative Reasoning items require examinees to calculate the total number of circles on a page. This is similar to items on the *Wechsler Individual Achievement Test–Second edition* (*WIAT-II*, Wechsler, 2001). Likewise, a *Woodcock-Johnson Tests of Achievement–Third edition* (*WJ-III ACH*; Woodcock, McGrew, & Mather, 2001a), Applied Problems subtest item asks the child to count the number of objects pictured. Some Standford-Binet items require the child to complete simple math problems just as the *WJ-III ACH* Math Fluency and the *WIAT-II* Numerical Operations tests do. Although it seems reasonable that math skills should be part of a test of achievement, if fair and equitable assessment of diverse populations is important, it does not seem reasonable that math skills should be used to measure ability because such acquired skills are influenced by instruction, ability, and the underlying neurocognitive processes related to doing the task efficiently.

Traditional IQ tests include a measure of word knowledge, as do tests of achievement. For example, students are required to define simple words on subtests included in the SB-5 or WISC-IV intelligence tests and the WJ-III Achievement test. The *Woodcock-Johnson Tests of Cognitive*

Abilities–Third edition (*WJ-III COG*; Woodcock, McGrew, & Mather, 2001b) battery contains a Verbal Comprehension subtest that has an item similar to "Tell me another word for small," and the WJ-III Achievement contains a Reading Vocabulary question like "Tell me another word for little." Included in the WJ-III Achievement Reading Vocabulary test is something like "Tell me another word for (examiner points to the word big)" and in the *WJ-III COG*, the examiner asks something like "Tell me another word for tiny." In addition, the WJ-III Cognitive Verbal Comprehension test contains 23 Picture Vocabulary items and the WJ-III Achievement includes 44 Picture Vocabulary questions. This overlap in content artificially increases the correlation between these tests of ability and achievement and raises important questions about the utility of measuring ability with questions that are clearly achievement laden.

It is particularly important that the role of knowledge and skills be recognized when ability tests are given to diverse populations, especially during neuropsychological assessment procedures. One way to assess ability without the confounding variables of language and knowledge is to use a nonverbal test of ability. These tests provide a way to assess individuals from diverse linguistic groups, especially those who have limited language skills as well as children with language impairments. In addition, children who cannot tolerate a lengthy test battery, such as some autistic children and others who are significantly inattentive, hyperactive, or children who easily fatigue secondary to traumatic brain injury, are more easily evaluated using nonverbal tests, especially those that are brief. Importantly, nonverbal tests provide the neuropsychological practitioner a way to conduct an evaluation on an individual who on other intelligence tests would fare poorly due to poor language skills.

Nonverbal tests can be particularly important when Hispanic children are assessed, as these children are more likely to have varying histories of educational opportunity and vary with respect to academic English language proficiency. To equitably evaluate the level of ability of Hispanics (the largest minority group in the United States; Ramirez & de la Cruz, 2002), tests that do not gauge intelligence on the basis of verbal and quantitative skills are necessary.

If it is accepted that verbal and quantitative questions in traditional IQ tests can be useful for prediction of achievement but are problematic for assessment of diverse populations, we may need to consider the following: What does such a test measure? Is a nonverbal test sufficient? Would a nonverbal test assess only a portion of intelligence? To address these questions, a reexamination of the history of the concept of general ability using verbal, quantitative, and nonverbal questions, and the view that these might be three separate "intelligences," should be placed within a more accurate historical perspective.

GENERAL ABILITY MEASURED USING NONVERBAL TESTS

The essence of a nonverbal test of general ability is that it measures general ability without verbal and quantitative test questions. The test questions evaluate general ability nonverbally via subtests with strong spatial requirements such as assembly of blocks to make a design or progressive matrices. The essential concept behind these tests is that they measure general ability nonverbally. How this is accomplished varies considerably. For example, some authors argue that the entire test must be administered nonverbally leading to the use of pantomimed instructions (e.g., the Universal Nonverbal Intelligence Test; Bracken & McCallum, 1997). Others suggest that nonverbal test directions for administration may be spoken (e.g., Naglieri Nonverbal Ability Test; Naglieri, 1997). Another method is to use pictorial directions as found in the *Naglieri Nonverbal Ability Test–Second edition* (*NNAT-2*; Naglieri, 2008a) and the Wechsler Nonverbal Scale of Ability (WNV; Wechsler & Naglieri, 2006). These nonverbal tests of general ability also differ in the diversity of the tests used. For example, some nonverbal tests are comprised of one type of item, the progressive matrix (e.g., *NNAT-2*) given in a group format or individual format (*Naglieri Nonverbal Ability Test–Individual Form*; Naglieri, 2003b). Another method is to use several different types of nonverbal subtests as found in the WNV (as well as the Universal Nonverbal Intelligence Test; Bracken & McCallum, 1997). Despite the differences in administration approach and subtest composition, tests measure general ability nonverbally and provide a way to fairly assess a wide variety of individuals regardless of their educational or linguistic backgrounds and/or disabilities. In the remainder of this chapter we will illustrate the advantages of an individually administered measure of general ability using the WNV.

THE WECHSLER NONVERBAL SCALE OF ABILITY

The WNV is comprised of subtests that measure general ability using tasks with a strong visual-spatial requirement, demand recall of spatial information or recall of the sequence of information, and paper-and-pencil skills. The multidimensionality of these tasks distinguishes the WNV from tests such as the *NNAT-2* (Naglieri, 2008a), which use only progressive matrices. Most of the WNV subtests have appeared in previous editions of the Wechsler scales and have an established record of reliability and validity for the nonverbal measurement of general ability. Adaptation of the subtests was necessary to accommodate the new pictorial directions format, identify items that were most appropriate for the specific ages, and provide directions in several languages.

STRUCTURE, ADMINISTRATION, AND SCORING

WNV raw scores are converted to *T*-scores (mean of 50 and standard deviation [*SD*] of 10) for each subtest. A Full Scale score is calculated for each battery that has a mean of 100 and a *SD* of 15. There are separate WNV norms tables based on standardization samples collected in the United States and Canada. There are 4- and 2-subtest batteries for each age band, 4:0 to 7:11 and 8:0 to 21:11. The subtests are briefly described below:

Matrices

The Matrices (MA) subtest requires the examinee to discover how different geometric shapes are spatially or logically interrelated. The multiple-choice items are constructed of geometric figures such as squares, circles, and triangles using some combination of the colors black, white, yellow, blue, and green. MA are always administered (i.e., it is given to examinees in both age bands and is included in both the 4- and 2-subtest batteries).

Coding

The Coding (CD) subtest requires the examinee to copy symbols (e.g., two vertical lines, a dash) that are paired with simple geometric shapes or numbers according to a key provided at the top of the page. Form A is used in the 4-subtest battery for ages 4:0 to 7:11 and Form B is used in the 4-subtest battery for ages 8:0 to 21:11.

Object Assembly

The Object Assembly (OA) subtest is comprised of items that require the examinee to complete pieces of a puzzle to form a recognizable object such as a ball or a car. OA is included in the 4-subtest battery of the WNV for examinees aged between 4:0 and 7:11.

Recognition

The Recognition (RG) subtest was created for use in the WNV and is included in both the 4- and 2-subtest batteries for examinees aged between 4:0 and 7:11. It requires the examinee to examine a stimulus (e.g., a square with a small circle in the center) for 3 seconds and then choose which option is identical to the stimulus that was just seen. The figures are colored black, white, yellow, blue, and/or green to maintain interest and minimize the likelihood that impaired color vision will influence the scores.

Spatial Span

The Spatial Span (SSp) subtest requires the examinee to touch a group of blocks arranged in an irregular pattern on an 8 × 11-inch board in the same and reverse order demonstrated by the examiner. SSp is included in both the 4- and 2-subtest batteries for ages 8:0 to 21:11.

Picture Arrangement

The Picture Arrangement (PA) subtest involves cartoon-like illustrations that must be put into a sequence that is logical and makes sense. PA is included in the 4-subtest battery for examinees aged between 8:0 and 21:11.

The WNV administration begins with short standardized introductions that tells examinees to look at the pictorial directions to understand what to do and that they can ask the examiner questions if necessary. The verbal instructions are provided in English, French, Spanish, Chinese, German, and Dutch. Actual administration procedures follow carefully scripted directions designed to ensure that the demands of the tasks are completely understood. Pictorial directions provide a standardized method of communicating the demands of the task by illustrating a scene like the one the examinee is currently in. The frames of the directions show the progression of an examinee being presented with the question, then thinking about the item, and finally, choosing the correct solution.

Examiner instructions include actions that must be carefully followed. Gestures are used to direct the examinee's attention to specific portions of the pictorial directions and to the stimulus materials and sometimes to demonstrate the task itself. Sometimes simple statements are also included because they convey the importance of both time and accuracy to the examinee. These are standardized simple sentences and gestures for communicating the requirements of the task. When the examinee is in need of further assistance, an opportunity to provide help is allowed, which allows the examiner flexibility. Examiners are given the opportunity to communicate in whatever manner they think will best explain the demands of the subtest based on their judgment. This could include providing further explanation or demonstration of the task, restating or revising the verbal directions, or using additional words to describe the requirements of the task. At no time, however, is it permissible to teach the examinee how to solve the items.

When using an interpreter to assist with administration, it is important that the interpreter has training about what is and what is not permitted. This interpreter should translate an explanation of the testing situation for the examinee, including the introductory paragraph at the beginning of chapter 3 in the *WNV Administration and Scoring Manual* before administration begins. The interpreter must recognize the boundaries of his or her role in administration. See Brunnert, Naglieri, and Hardy-Braz (2008) for more information about working with interpreters and especially when testing those who are deaf or hard of hearing.

Scoring the WNV is uncomplicated. Five of the six subtests (i.e., MA, CD, RG, SSp, and PA) are scored by summing the number of points earned during administration. The sixth subtest (i.e., OA) has time bonuses for some items that might be part of the raw score. The raw scores are converted to *T*-scores. The sum of *T*-scores is converted to a Full Scale score, with corresponding percentile rank and confidence interval included in the conversion table. The WNV Scoring Assistant provides computer scoring program that obtains all derived scores based on the United States as well as the Canadian normative sample comparisons. The report writing feature of the software provides reports that are appropriate for clinicians as well as parents. The parent report is available in English, French, and Spanish. The software also provides links between the *WNV* and the *WIAT-II* and all the ability comparisons to achievement.

WNV STANDARDIZATION SAMPLE

The WNV was standardized in the United States and Canada. The U.S. sample consisted of 1,323 examinees stratified across five demographic variables: age (4:0–21:11), sex, race/ethnicity (Black, White, Hispanic, Asian, and Other), Education Level (8 years or less of school, 9–11 years of school, 12 years of school [high school degree or equivalent], 13–15 years of school [some college or associate's degree], and 16 or more years of school [college or graduate degree]), and Geographic Region (Northeast, North Central, South, and West). Education Level was determined by the parent education for examinees aged between 4:0 and 17:11 and by the examinee's own education for ages from 18:0 to 21:11. Approximately 4% of the U.S. normative sample was comprised of individuals with limited English skills.

The Canadian sample consisted of 875 examinees stratified across five demographic variables: age (4:0–21:11), sex, race/ethnicity (Whites, Asians, First Nations, and Other), Education Level (less than a high school diploma; high school diploma or equivalent; college/vocational diploma or some university, but no degree obtained; and a university degree), and Geographic Region (West, Central, and East). In addition, the Canadian sample consisted of 70% English speakers, 18% French speakers, and 12% speakers of other languages. See the *WNV Manual* (Wechsler & Naglieri, 2006) for more details.

RELIABILITY OF THE WNV

WNV coefficients are provided by subtest and Full Scale scores by age and overall ages for the U.S. and Canadian normative samples, and for all the special groups in the WNV *Technical and Interpretive Manual*. The reliability estimates for the U.S. normative sample ranged from 0.74 to 0.91 for the subtests and were .91 for both Full Scale scores across ages. The reliability estimates for the Canadian normative sample ranged from 0.73 to 0.90 for the subtests, were 0.90 for the Full Scale score: 4-subtest battery, and 0.91 for the Full Scale score: 2-subtest battery. The reliability estimates for the studies with examinees that were diagnosed with or classified as being Gifted, Mild Mental Retardation, Moderate Mental Retardation, Reading and Written Expression Learning Disorders, Language Disorders, English Language Learners, Deaf, and Hard of Hearing are provided in the manual. Other information such as the standard error of measurements, confidence intervals, and test–retest stability estimates for both the U.S. and the Canadian normative samples is provided in the WNV Technical and Interpretive Manual and Administration and Scoring Manuals.

INTERPRETATION METHODS

The WNV test results should always be interpreted within context; past and present. Perhaps the most important are issues such as the behaviors observed during testing, relevant educational and environmental backgrounds, and physical and emotional status, all in relation to the reason for referral. In order to obtain the greatest amount of information from the WNV, there are methods of interpretation that warrant discussion that are the same for the 4- and 2-subtest batteries as well as others that are unique to each version. In this chapter, the issues that apply to both batteries will be covered first and then the finer points of interpretation within a neuropsychological assessment will be examined.

Interpretation of the Two WNV Versions

Both versions of the WNV are comprised of subtests (set at a mean of 50 and *SD* of 10) that are combined to yield a Full Scale score (set at a mean of 100 and *SD* of 15) based on either the 4- or 2-subtest batteries. This score provides a nonverbal estimate of general ability that has excellent reliability and validity. In addition, even though the WNV subtests have different demands—that is, some are spatial (e.g., MA or OA), others involve sequencing (PA and SSp), require memory (e.g., RG and SSp), or use symbol associations (CD)—they all measure general ability. General ability, as represented by the Full Scale standard score, provides an estimate for predicting how well a person, for example, will be able to understand spatial as well as verbal and mathematical concepts, remember visual relationships as well as quantitative or verbal facts, and work with sequences of information of all kinds. The content of the questions may be visual or verbal and require memory or recognition, but general ability (sometimes referred to as *g*) underlies performance on all these kinds of tasks.

WNV Interpretation

Step 1: The Full Scale score should be reported with its associated percentile score, categorical description (Average, Above Average, etc.), and confidence interval. The following illustrates how this information could be included in a written document:

Sally obtained a WNV Full Scale score of 91, which is ranked at the 27th percentile and falls within the Average classification. This means that she performed as well as or better than 27% of examinees her age in the normative sample. There is a 90% chance that her true Full Scale score falls within the range 85–99.

Step 2: Examine the subtests' *T*-scores, taking into consideration the lower reliability of these scores. Examination of the four WNV subtests should also be conducted with consideration that even though the subtests are all nonverbal measures of general ability they do have unique attributes (i.e., some involve remembering information, others spatial demands, etc.). In addition, statistical guidelines should be followed to ensure that differences interpreted are beyond those that could be expected by chance. The values needed for significance when comparing a WNV subtest for an examinee to that examinee's mean *T*-score are provided in the WNV Administration and Scoring Manual (Table B.1) and in more detail by Brunnert et al. (2008), and should be used when examining subtest variability. The following steps should be used to compare each of the four WNV subtest *T*-scores to the child's mean subtest *T*-score:

1. Calculate the mean of the four subtest *T*-scores.
2. Calculate the difference between each subtest *T*-score and the mean.
3. Subtract the mean from each of the subtest *T*-scores (retain the sign).
4. Find the value needed for significance using the examinee's age group and the desired significance level in Table 12.3 of the WNV Manual.
5. If the absolute value of the difference is equal to or greater than the value in the table, the result is statistically significant.
6. If the subtest difference from the mean is lower than the mean, then the difference is a weakness; if the subtest difference from the mean is greater than the mean, then the difference is strength.

When *there* is significant variability in the WNV subtests, it is also important to determine whether a weakness relative to the examinee's overall mean is also sufficiently below the average range. Determining whether a child has significant variability relative to his or her own average score is a valuable way to determine strengths and weaknesses relative to the child's mean score, but Naglieri (1999) cautioned that a relative weakness could also be significantly below the normative mean. He recommended that any subtest score that is low relative to the child's means should also fall below the average range to be considered a noteworthy weakness (e.g., <1 *SD* below the normative mean).

TESTING IN NEUROPSYCHOLOGICAL CONTEXTS

Forward and Backward Span

The WNV SSp subtest Forward and Backward scores can be interpreted separately, particularly when this test is embedded within the greater context of a comprehensive assessment. The sizes of the differences required for statistical significance by age and for the U.S. and Canadian samples are 11 and 13 for the 0.10 and 0.05 levels for the United States and 10 and 13 for the Canadian standardization samples for the combined ages 8:0 to 21:11. The comparisons are accomplished using Table C.1 of the WNV Administration and Scoring Manual, which provides a way to convert the raw scores to *T*-score equivalents for SSp Forward and SSp Backward. A difference of 9 *T*-score points is needed at the 0.15 level (13 at the 0.05 level) for significance. (Note, base-rate data by the direction of the difference is provided in the WNV Manual.)

Information about SSp Forward and Backward *T*-scores may provide useful information, but it should be integrated within the greater context of a comprehensive assessment. For example, if a difference between SSp Forward *and* Backward *was* found, it would be expected that other similar test score results, such as WISC-IV Digit Span Forward vs. Digit Span Backward would also be found. The Backward scores could be related to the Planning Scale of the Cognitive Assessment System (see Naglieri, 1999) and may suggest that the examinee has difficulty with development and utilization of strategies for reversing the order of serial information. This finding suggests difficulty in one component on executive functions. In neuropsychological paradigms, digit or SSp forward are considered measures of initial registration and sequencing, whereas Digit Span or SSp Backward may be considered as measures of sustained concentration and verbal or spatial working memory, respectively (Miller, 2007, 2009). Working memory is subserved by several brain regions; in particular the frontal lobes are intimately involved (Goldberg, 2009).

Qualitative Interpretation

When conducting a pediatric neuropsychological assessment, a child's performance on the WNV can be very informative and was one of several recommended methods encouraged by the late premier neuropsychologist, academician, test developer, and researcher Edith Kaplan (E. Goldberg, personal communication, June 2009). Take for example the CD subtest; it is similar to the task on other Wechsler scales in which a series of numbers is paired with symbols. The examinee is required to draw the symbol that is arbitrarily associated with the corresponding number quickly and correctly. The numbers are not arranged in any particular order and the subject has to approach the task in a sequential manner without skipping any. The brain–behavior relationship related to this task has not been specifically identified and the neuroanatomical substrates continue to be explored. Koziol and Budding (2009) hypothesized that a task such as CD places greater demands on working memory functions because there are numbers and symbols, and quick performance might be facilitated by "holding this information online in working memory in the course of performing the task" (p. 261). So for example, the associations between symbols and numbers would be maintained within reciprocal prefrontal-cortical circuits. A short-term plan of action is activated through these representations and associations. If the number–symbol associations are made quickly it is assumed that less conscious effort is required for the task and that the cortical-subcortical loops of the prefrontal-basal ganglia are engaged efficiently. This observation is consistent with Gabrieli, Stebbins, Singh, Willingham, and Goetz (1997) formulation that working memory capacity facilitates fast performance and the attainment of procedural learning.

The child who completes the CD subtest accurately but very slowly is approaching the task differently from the child who completes the task quickly but with many errors (making the wrong number–symbol association, skipping) and is different from the child who completes the task quickly and accurately. One useful way to interpret this kind of performance could be in terms of attention to the instructions (e.g., the instruction to do the task as fast as you can was ignored) or conscious effort or concentration. The child who works faster for the 120 seconds gives more responses than the subject who worked more slowly and gives fewer responses have approached the task very differently. These subjects completed the tasks at different rates over the same time interval. From this we can hypothesize that the child who worked more slowly had to put forth greater conscious control and effort, which may be related to recruitment of more brain area (Saling & Phillips, 2007). The child who worked more slowly had to concentrate harder and the child who worked quickly likely expended less effort. In this way, the subtest scale score for CD may be viewed as a measure of level of efficiency of concentration.

Some speculate that the speed with which CD is performed can be considered as one measure of executive control. Speed of task execution and executive control are related to the frontal-striatal system (Rabbit et al., 2007). Poor performance on a task such as CD can occur for many reasons, and the reason is not necessarily evident in the objective data or summary sores. The point here is that processing speed could be related to cognitive control. Therefore the clinician always needs to be observant and consider all the observations and objective data carefully.

The MA subtest is similar to others that have a long history as good measures of general ability as measured by high "*g*" loadings (Jensen, 1998; Williams, Wiess, & Rolfhus, 2003). Within a neuropsychological framework, the MA subtest of the WNV can be viewed as a test of visual perceptual reasoning. Viewed this way, the test score reflects one subcomponent of visual-spatial processes typically assessed within neuropsychological evaluations. MA can also be considered a test of Simultaneous processing; a mental activity by which a person integrates stimuli into interrelated groups or a whole (Naglieri, 1999). Simultaneous processing tests typically have strong visual-spatial aspects. The cognitive demand of the task requires the integration of information. This ability to identify patterns as interrelated elements is made possible by the parieto-occipital-temporal brain regions (Naglieri & Otero, in press).

Children who have great difficulty integrating letters to make a whole word, are overfocused on details of their work, and have great difficulty seeing the relationships between the details may have difficulty with a subcomponent of visual-spatial processes referred to as part–whole relationships. In cases of brain injury where simultanagnosia is present—impaired recognition of the meaning of whole pictures or objects, but intact ability to describe the parts/objects, or in cases of intergrative Agnosia—failure in integrating the parts of a picture or object into a whole, the MA subtest may be one measure, among others, to use in the evaluation of these types of impairments.

SS can be considered the nonverbal variant of the of the Digit Span subtest. For SS Forward, the child repeats a sequence of tapped blocks in the same order as demonstrated by the examiner. For SS Backward, the child repeats a sequence of tapped blocks in the reverse order of that demonstrated by the examiner. SS is considered to tap into visual-spatial working memory. Goldberg (2009) defines working memory "as the selection of task–relevant information" (p. 94), and it is the selection process incorporated into the task that engages the frontal cortex. In SS Backward, the child has to select (repeat) the correct sequence of block tapping shown by the examiner. Observing the child's performance on SS Forward can reveal information about how well the child initially commits sequenced visual-spatial information to memory and visual-spatial capacity. SS Backwards allows the examiner to observe visual working memory capacity and efficiency in the selection of the sequence executed. Normative information for comparisons of SSp Forward and Backwards as well as normative sample base rates are found in the WNV Manual.

Recently, one of the authors (T.O.) evaluated an 11-year-old Hispanic male named Arturo who has a history of uneven academic progress in math and writing since a soccer-related head injury over 1 year ago. The young man was left with significant expressive language and attention concerns postinjury. His neurologist has indicated that Arturo has recovered nicely in spite of his ongoing difficulties in both of these areas. As part of a larger assessment and because of expressive language impairment, Arturo was administered the WNV. In this case, the WNV allowed the examiner to estimate Arturo's ability more reliably than if he had assessed him with a more traditional IQ test with verbally laden items. Arturo scored within the average range (FS = 105). His SSp Backward *T*-score was well below average (T = 34), while his SSp Forward was average (T = 53). The difference between these subtest conditions is significant and may be considered a pathognomonic finding consistent with working memory difficulties when also viewed in relation to his behavior during the test. Arturo was slow in initiating his response on SS Backwards and asked for repetition of the taping sequence and seemed to be satisfied with his response even when falling short of completing the entire sequence. Arturo's teachers indicated he seems to lose track of what he is doing in the middle of working on math problems and lose track of what he is writing about. He can read well but has difficulty summarizing what he reads. These behaviors may be considered consistent with poor working memory.

WNV IN CLINICAL NEUROPSYCHOLOGICAL SETTINGS

English as a Second Language

The United States continues to become more diverse as the number of individuals whose primary language is not English continues to increase. The large number of immigrants in the United States makes clear the need for neuropsychological tests that are appropriate for those who come to the assessment with limited academic and English language skills. Nonverbal tests of general ability such as the WNV are, therefore, a particularly useful way to assess minority children because they yield smaller race and ethnic differences (which is attributed to the difference in content) while these instruments retain good correlations with achievement and can help identify minority children for gifted programs (Bracken & McCallum, 1997; Naglieri & Ford, 2003; Naglieri & Ronning, 2000a, b).

In the WNV Manual, Wechsler and Naglieri (2006) provide studies of the utility of the WNV for individuals who are learning English. The findings were presented for those who speak English as a second language in comparison to a matched sample from the WNV standardization sample. This included 55 examinees aged 8–21 years whose native language was not English, spoke a language other than English at home, and whose parents had resided in the United States less than 6 years. There were 27 Hispanics and 28 examinees whose primary language was Cantonese, Chinese, Korean, Russian, Spanish, or Urdu. This sample earned very similar scores to their matched counterparts from the normative sample with negligible effect sizes for the Full Scale scores from both WNV test batteries. Additional information about this sample is available in the WNV *Technical and Interpretive Manual*.

Gifted

The WNV can be used to address the underrepresentation of minority children in classes for the gifted (Ford, 1998; Naglieri & Ford, 2005) because the test measures general ability without verbal and quantitative content. Naglieri and Ford (2005) stated that the verbal and quantitative content of some of typical IQ tests are inconsistent with the characteristics of culturally, ethnically, and linguistically diverse populations and therefore a nonverbal measure is more appropriate. That is, because IQ has traditionally been defined within a verbal/quantitative/nonverbal framework, students with limited English language and math skills earn lower scores on the Verbal and Quantitative scales these tests include because they do not have sufficient knowledge of the language or training in math and not because of low ability (Bracken & Naglieri, 2003; Naglieri, 2008a).

There is evidence that gifted children earn high scores on the WNV in the test manual. The WNV was administered to gifted children who were carefully matched to control subjects included in the standardization sample on the basis of age, race/ethnicity, and education level. The study included 41 examinees, all of whom had already been identified as gifted using a standardized ability measure where they performed at 2 *SD*s above the mean or more. The students in the gifted programs performed significantly better than their matched counterparts from the normative sample with effect sizes that were large for the Full Scale score.

Deaf and Hard of Hearing

Wechsler and Naglieri (2006) compared profoundly deaf examinees with cases from the standardization of the WNV who were matched on a number of important demographic variables. The profoundly deaf examinees met the following criteria: no hearing of spoken language after the age of 18 months, no ability to lip read; no cued speech; and severe to profound deafness. These examinees performed minimally differently than their matched counterparts from the normative sample with effect sizes negligible for Full Scale score: 4-subtest battery and Full Scale score: 2-subtest battery. Similarly, a sample of individuals who were hard of hearing was compared to a group from the standardization sample. This study included examinees who had unilateral or bilateral hearing loss or deafness, the age of onset of their inability to hear could be any age, and they could have implants. Again, there were negligible differences between the groups for the WNV Full Scale scores on the 2- and 4-Subtest Batteries.

Specific Learning Disabilities

The WNV provides a score that represents general ability that can be compared to current achievement test scores to help determine if there is an ability/achievement discrepancy. The WNV provides an evaluation of ability that is not influenced by achievement-like verbal and quantitative content (see Naglieri & Bornstein, 2003, or Naglieri, 2008b, for more discussion of the similarity of test questions on ability and achievement tests). The WNV also provides a measure of ability that can be viewed as nondiscriminatory on the basis of race, ethnicity, language, and disability (see the following sections). The WNV is not, however, designed to be a test of basic psychological processes, and other tools should be used for that purpose (e.g., the Cognitive Assessment System; Naglieri & Das, 1997).

Practitioners who wish to compare WNV scores with the *WIAT-II* (Wechsler, 2005) can do so using the predicted-difference and simple-difference methods. The predicted-difference method takes into account the reliabilities and the correlations between the two measures. In this method, the ability score is used to predict an achievement score, and the differences between predicted and observed achievement scores are compared. The values needed to make these comparisons are provided in the WNV *Technical and Interpretive Manual* for both U.S. and Canadian normative samples when using the *WIAT-II*.

SUMMARY AND CONCLUSIONS

The WNV was designed to provide a nonverbal measure of general ability that would be appropriate for a wide variety of culturally and linguistically diverse populations and be useful in a number of clinical assessment settings. The selection of subtests in conjunction with pictorial directions and oral directions in five languages provides a unique approach to measuring general ability nonverbally. The evidence provided in this chapter summarizes some of the validity evidence provided in the test Manual, which supports the utility of the test for fair assessment of cognitive ability especially for those from culturally diverse backgrounds as well as those with language differences, deficiencies, and impairments and those who are deaf or hard of hearing. Within the context of a comprehensive pediatric neuropsychological evaluation, the WNV can be of additional usefulness in going beyond the level of performance in relation to normative data, by making comparisons across and within subtests and by engaging in qualitative analysis as a way to assess children and take notice of pathognomic signs.

REFERENCES

Boake, C. (2002). From the Binet-Simon to the Wechsler-Bellevue: Tracing the history of intelligence testing. *Journal of Clinical and Experimental Neuropsychology, 24*, 383–405.

Bracken, B. A., & McCallum, R. S. (1997). *Universal nonverbal intelligence test.* Itasca, IL: Riverside.

Bracken, B. A., & Naglieri, J. A. (2003). Assessing diverse populations with nonverbal tests of general intelligence. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children* (pp. 243–273). New York: Guilford.

Brunnert, K., Naglieri, J. A., & Hardy-Braz, S. (2009). *Essentials of WNV assessment.* New York: Wiley.

Ford, D. Y. (1998). The underrepresentation of minority students in gifted education: Problems and promises in recruitment and retention. *Journal of Special Education, 32,* 4–14.

Gabrieli, J. D., Stebbins, G. T., Singh, J., Willingham, D. B., & Goetz, C. G. (1997). Intact mirror-tracing and impaired rotary-pursuit skill learning in patients with Huntington's disease: Evidence for dissociable memory systems in skill learning. *Neuropsychology,* 11, 272–281.

Goldberg, E. (2009). *The new executive brain: Frontal lobes in a complex world.* New York: Oxford University Press.

Jensen, A. R. (1998). *The 'g' factor: The science of mental ability.* Westport, CT: Praeger.

Koziol, L., & Budding, D. E. (2009). Subcortical structures and cognition: Implications for neuropsychological assessment. New York: Springer.

Miller, D. C. (2009). *Best practices in school neuropsychology: Guidelines for effective practice, assessment, and evidence-based intervention.* Hoboken, NJ: Wiley.

Miller, D. C. (2007). *Essentials of school neuropsychology assessment.* Hoboken, NJ: Wiley.

Naglieri, J. A. (1997). *Naglieri nonverbal ability test—Multilevel form.* San Antonio: Psychological Corporation.

Naglieri, J. A. (1999). *Essentials of CAS assessment*. New York: Wiley.

Naglieri, J. A. (2003a). Naglieri nonverbal ability tests: NNAT and MAT-EF. In R. S. McCallum (Ed.), *Handbook of nonverbal assessment* (pp. 175–190). New York: Kluwer.

Naglieri, J. A. (2003b). *Naglieri Nonverbal Ability Test—Individual form.* San Antonio: Psychological Corporation.

Naglieri, J. A. (2008a). *Naglieri Nonverbal Ability Test* (2nd ed.). San Antonio: Pearson.

Naglieri, J. A. (2008b). Traditional IQ: 100 years of misconception and its relationship to minority representation in gifted programs. In J. VanTassel-Baska (Ed.), *Critical issues in equity and excellence in gifted education series alternative assessment of gifted learners* (pp. 67–88). Waco, TX: Prufrock Press.

Naglieri, J. A., & Bornstein, B. T. (2003). Intelligence and achievement: Just how correlated are they? *Journal of Psychoeducational Assessment, 21,* 244–260.

Naglieri, J. A., & Das, J. P. (1997). C*ognitive Assessment System*. Austin, TX: ProEd.

Naglieri, J., & Ford, D. Y. (2003). Addressing under-representation of gifted minority children using the Naglieri Nonverbal Ability Test (NNAT). *Gifted Child Quarterly, 47, 15–160.*

Naglieri, J. A., & Ford, D. Y. (2005). Increasing minority children's participation in gifted classes using the NNAT: A response to Lohman. *Gifted Child Quarterly, 49,* 29–36.

Naglieri, J. A., & Otero, T. (2010). Cognitive assessment system: Redefining intelligence from a neuropsychological perspective. In A. Davis (Ed.). *Handbook of pediatric neuropsychology*. New York: Springer.

Naglieri, J. A., & Ronning, M. E. (2000a). Comparison of white, African-American, Hispanic, and Asian children on the Naglieri Nonverbal Ability Test. *Psychological Assessment, 12,* 328–334.

Naglieri, J. A., & Ronning, M. E. (2000b). The relationships between general ability using the NNAT and SAT Reading Achievement. *Journal of Psychoeducational Assessment, 18,* 230–239.

Rabbit, P., Scott., M., Lunn, M., Thacker, N., Lowe, C., Pendelton, N., et al. (2007). White matter leasions account for age related declines in speed but not intellegence. *Neuropsychology, 21,* 363–370.

Ramirez, R. R., & de la Cruz, G. (2002). *The Hispanic population in the United States: March 2002*. *Current Population Reports* (pp. 20–545). Washington, DC: U.S. Census Bureau.

Roid, G. (2003). *Stanford-Binet Intelligence scale* (5th ed.). Itasca, IL: Riverside Publishing.

Saling, L. L., & Phillips, J. G. (2007). Automatic behavior: Efficient not mindless. *Brain Research Bulletin, 73,* 1–20.

Wechsler, D. (2005). *Wechsler Individual Achievement Test* (2nd ed.). San Antonio, TX: Pearson.

Wechsler, D. (2001). *Wechsler Individual Achievement Test* (2nd ed.). San Antonio, TX: Pearson.

Wechsler, D., & Naglieri, J. A. (2006). *Wechsler Nonverbal Scale of Ability*. San Antonio, TX: Pearson.

Woodcock, R. W., McGrew, K.S., & Mather, N. (2001a). *Woodcock Johnson III Test of Achievement*. Itasca, IL: Riverside Publishing Company.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001b). *Woodcock Johnson III Test of Cognitive Ability*. Itasca, IL: Riverside Publishing Company.

Yoakum, C.S., & Yerkes, R.M. (1920). *Army mental tests*. New York: Henry Holt & Company.