Technical Assistance Paper   312662
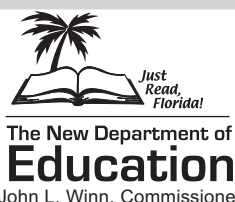
Nonverbal Tests of Intelligence

Background

This technical assistance paper (TAP) provides an introduction to the field of nonverbal assessment of intelligence followed by questions and answers that are relevant to the topic. Specifically, it elaborates on the clinical decision to use nonverbal tests of intelligence for students who may have economic, cultural, or language/communication issues that could introduce significant error to the full scale IQ scores obtained with many comprehensive tests. It is strongly suggested, however, that the principles and recommendations embedded in this technical assistance paper be considered best practice in evaluation and be applied on an individual bases with all students who are being evaluated.

In Florida, there is a disproportionately high percentage of African-American students receiving exceptional education services under the category of mentally handicapped and more specifically, educable mentally handicapped (EMH), suggesting that the impairment is mild rather than profound. Research has indicated that this is not an issue restricted to Florida; it is a prevalent problem throughout most of the United States. In fact, national data suggest that the risk for an African-American student to be classified as mildly mentally handicapped is more than twice the rate of Caucasian students. Research suggests that much of this discrepancy across racial groups can be attributed to the type of assessment conducted in the subsequent psycho-educational evaluation. Selecting a nonverbal test of intelligence for students of diverse backgrounds and students with English language concerns may be the most valid and reliable way to ascertain an estimate of their current level of functioning (National Research Council, 2002); however, other research suggests that a nonverbal test should only be used if existing data clearly support excluding the verbal or language-based assessments (Lohman, 2004). In general, it is recommended that for students who are referred for assessment, all areas related to the area of suspected disability be evaluated and the data obtained should be interpreted within the context of a collaborative problem-solving approach that explores both student and context characteristics.

Given that a student can only be considered eligible for services under the EMH program if the intellectual assessment, the adaptive behavior assessment, and the assessment of levels of achievement are significantly below average (typically interpreted as 2 or more standard deviations below the normative mean), parents and caregivers who share the student's cultural background must be given the opportunity to contribute relevant information to the evaluation team. In an effort to increase parent participation in the data collection process, the Florida

Department of Education developed a technical assistance paper "Measuring Adaptive Behavior" (FY2005-3) that describes the imperative role parents and primary caregivers play in assessing adaptive behavior functioning.

In addition to the over-representation of African-American students in the program for students who are EMH, a lower than expected number of African-American students are identified and receiving exceptional education services through programs for gifted students (FDOE, 2004). In line with the data previously discussed, the results of some language-loaded tests of intelligence (i.e., tests that heavily rely on the use of both receptive and expressive language skills) may underestimate the current and potential functioning of students from culturally diverse backgrounds. This is also true for students from disadvantaged environments where the opportunity for early language development and pre-literacy skills may be limited.

More than 80 languages are spoken in Palm Beach County Schools (Fast Fact, 1996), and at least 54 languages are spoken in Broward County (Donzelli, 1996). Because many tests of intelligence require extensive use of verbal language for administration, educators have felt challenged when measuring the cognitive functioning of individuals who are deficient in English language skills. The demand for evaluation procedures that significantly reduce the impact of language skills on student performance resulted in the development and publication of a large number of nonverbal assessment instruments during the 1990s. These tools increasingly are used to support the decision-making process in determining eligibility for exceptional student education.

Introduction

Nonverbal intellectual assessment is the process of assessing the construct of intelligence *without* placing receptive or expressive language demands on either the examinee or the examiner (McCallum, Bracken & Wasserman, 2001). Most of these tests were designed to measure general cognition without the confounding effect of language ability; they are not designed to be tests of a construct such as "nonverbal intelligence" (DeThorne & Schaefer, 2004). Ideally, verbal test directions and spoken responses should not be required during the administration of a nonverbal test. It should be noted that a measure of intelligence (verbal, nonverbal, or a combination of both) is *only one factor* used in the determination of a disability under the 1997 reauthorization of the Individuals with Disabilities Education Act (IDEA). For students suspected of having a disability, the evaluation should be a comprehensive assessment and may or may not include a standardized test of intelligence. Depending on the specific areas of concern prompting the referral, a comprehensive assessment might include but is not limited to an evaluation of adaptive behavior, a social-developmental history, and an assessment of current academic functioning. Multiple sources of data should be explored by including input from parents, teachers, and others who know the student well. Finally, multiple methods for gathering data should be included in any broad-based assessment. These methods for obtaining information include observations, interviews, review of records, and formal and informal tests.

Additional information on intellectual assessment in the evaluation of students is contained in the following TAPs: "The Use of Part Scores with Tests of Intelligence" (in process); "Assessing Limited English Proficiency (LEP) Students for Eligibility for Gifted Programs" (FY1999-6); "School Psychology Reports" (FY1994-6), and "Standard Error of Measurement" (FY1996-7), available online at http://www.firn.edu/doe/commhome/tap-home.htm.

Purpose

This TAP is designed to assist Florida school districts in conducting appropriate assessments with students who are struggling learners and in meeting the needs of exceptional students. It is intended to provide guidance in determining the supports and services necessary to meet the academic, social, and cognitive goals of these students and has been developed through a review of professional literature; consultation with experts in the fields of language deficiencies, intellectual assessment, and mentally handicapping conditions; and input from school district personnel in exceptional student education and school psychology.

Questions and Answers

1. **How do I decide which test of intelligence, if any, is appropriate for a given student?**

   The psychometric properties (validity and reliability) of a test are important factors to consider. Information on validity and reliability should be available through the published technical manuals accompanying each test. The psychometric strengths and weaknesses of a given test should be evaluated thoroughly before a decision is made to administer it for data collection or eligibility consideration. An examiner deciding which test to administer should ensure to the greatest extent possible that the sample of students used in the norming of the test reflects the age, grade, ethnicity, and other critical features of the student being evaluated. Finally, a student's English language proficiency as well as his or her cultural background should be considered when deciding the type of intellectual assessment tool to administer (i.e., language free or language-reduced or language-loaded).

2. **Are test selection procedures different for a student from a racial or ethnic minority group than for a student with an English language deficiency or a language-processing deficit?**

   No. The *process* for test selection should not vary, but the choice of the test actually used in a comprehensive assessment may be different based on the presenting characteristics of the student. Research indicates that the type of intellectual measure used can influence the rate of exceptional education eligibility for certain racial/ethnic groups (Macmillan et al., 1998). In all cases, if a language-loaded test would introduce bias to the assessment results, then the data should not be interpreted as a unitary construct representing the student's overall intellectual functioning. This is not to suggest that the data should be ignored. All results should be interpreted within the context of the student and the learning environment in which he or she is expected to achieve and progress.

3. **In evaluating students with English language deficiencies or ethnic differences, how might results differ between a broad-based intellectual evaluation (assessing both verbal and nonverbal abilities) and those from a nonverbal assessment?**

   Most tests are designed to measure general intelligence. Similar to the response for the previous question, if a student has a language deficiency, scores generated from tests that require verbal activity from both the examiner and the examinee may introduce more

error to the overall estimate of intelligence than nonverbal tests. In other words, if a language-loaded test is administered to a student with suspected language difficulties (language-processing disorder, English as a second language, limited English proficiency, poor articulation, or cultural differences that could impact a student's communication skills) and the results fall outside normal limits, it is imperative that the examiner review all possible explanations for the discrepancy before drawing conclusions about the student's level of cognitive functioning. In some cases, the global IQ score attained for a student may significantly under-represent his or her intellectual capacity because of specific language processing and communication deficits or cultural and environmental experiences and opportunities. When assessed using a nonverbal measure of intelligence, this same student may achieve substantially higher scores that more appropriately reflect his or her level of functioning.

4. **Which test results should be used when a significant discrepancy exists between the verbal and the non-verbal scores?**

   The individual responsible for conducting the assessment should use his or her professional judgment regarding which test results are most representative of the student's current level of intellectual functioning and identify which variables, if any, have skewed the outcome data. Evaluators, including school psychologists; ESE, LEP, and ESOL teachers; and speech/language pathologists, should consult with each other to explore various hypotheses about inconsistent test results before drawing conclusions about the student's capacity in the learning environment.

   Since in some cases there may be a greater error factor in verbal subtests, direct comparisons between verbal and nonverbal tests may be misleading. Verbal tests given to students with limited English proficiency, for example, can have value as baseline estimates and should not automatically be excluded from an assessment battery. As always, the staffing team should review all relevant data available for a particular student before making decisions about eligibility for special education services. In some cases, additional, more focused assessment may be required for consideration in the problem-solving process.

5. **How do the racial/ethnic or linguistic characteristics of the examiner affect the performance of the examinee?**

   Research to date has not supported the existence of a systematic effect of the race/ethnicity of the examiner on test performance of students from similar or different racial/ethnic groups. However, if the language of the examiner is difficult for the student to understand or vice-versa, the results of an evaluation that is language-loaded may be compromised. Nonverbal measures may be the most appropriate choice under these circumstances.

   The ability to develop and maintain rapport is considered more important than the ethnic or language characteristics of the examiner per se. When testing students who are deaf or hearing-impaired, it is important that the examiner accurately assess his or her own communication skills (i.e., sign language) and the skills of the student (i.e., total language approach) and employ the expertise of a qualified interpreter if needed. Similar practice is recommended when assessing a student who has limited English proficiency. In some

situations, either a translator should be present or, ideally, a bilingual psychologist should administer the assessment. This is particularly important if conclusions regarding the student's intellectual functioning will be drawn from a comprehensive test (verbal and non-verbal) that requires language skills in which the student is deficit or impaired.

6. **How might performance on verbal and nonverbal intellectual evaluations differ for a student who has a language processing deficiency but is English speaking?**

Language impairments may impact a student's performance on a test where receptive and/or expressive language skills are measured. A student with a language processing deficiency will likely score less well on tasks on an intelligence test that requires receptive language skills and in some cases, expressive language skills. If an eligibility team were to place too much emphasis on the overall results (full scale or global score) of a comprehensive test of intelligence without considering the impact the language deficiency had on the student's performance, the interpretation may significantly underestimate the student's current level of functioning.

Some districts in Florida compare a student's IQ score with the results of a language assessment to determine if a significant difference in performance exists (one or more standard deviations). Other districts look at how language scores compare to a normative mean, and still other districts may choose to compare language scores with the results of an intellectual screening test. Specific policy is defined in the district special programs and procedures manual; however, regardless of local practice, if the results of an intellectual screening measure are of questionable validity, the student should be referred to the school-based, problem-solving team (child study team, student support team) to determine if further assessment is necessary. While districts have some autonomy in defining their own practices, it is best practice to review all of the available data to determine if patterns of performance emerge with both standardized assessment and authentic, curriculum-based assessment. Please refer to the technical assistance paper, "The Use of Part Scores with Tests of Intelligence," for more specific information on this issue.

7. **How do personnel other than the school psychologist (i.e., speech/language pathologist, ESE teachers, LEP teachers, and ESOL teachers) contribute to the problem-solving process as it relates to ESE eligibility for students with English language concerns?**

There are many assessment and intervention specialists who can provide unique information to the problem-solving team. For example, an ESOL teacher may provide insight about the cultural variables that can affect student outcomes and can offer guidance as to when an assessment should be administered in the student's first language. Exceptional and general education teachers may offer the team functional achievement data and recommend appropriate intervention strategies based on the student's learning profile. These interventions can provide information regarding the student's rate of learning and skill acquisition when evidence-based instruction is provided with consistency and integrity. For students whose academic functioning may be compromised by a lack of exposure to such instruction, this response to intervention (RtI) method of measuring growth can be data-rich and exceedingly important in determining eligibility for ESE

services. In addition, speech/language pathologists provide insight regarding a student's capacity to process language effectively and efficiently. This type of collaboration designed to identify the primary area of need for the student and generate specific interventions is more effective than analyzing and interpreting test results in isolation.

8. **Who in the state of Florida is qualified to administer a nonverbal test of intelligence?**

Based on Florida statute, if the assessment data are used as a measure of general intellectual functioning, the only qualified personnel are those licensed and/or certified as psychologists or school psychologists (Florida Law 6A-6.0331, FAC). If the data are being used for purposes other than a general intellectual measure (such as a measure of nonverbal performance often assessed for the purpose of determining eligibility for speech and language services,) evaluators should refer to the publisher's technical manual for qualification guidelines.

Appendix A
Specific Assessment Tools Used in Nonverbal Assessment of Intelligence

The Comprehensive Test of Nonverbal Intelligence (C-TONI), the Universal Nonverbal Intelligence Test (UNIT), and the Raven's Progressive Matrices (RPM) are examples of tests of intelligence that are administered in a nonverbal manner. With the exception of a few subtests, the Leiter-Revised is also primarily a nonverbal measure as is the Kaufman Assessment Battery for Children-II (KABC-II) when administering the nonverbal subtests using the pantomimed administration specified by the test authors. These tests are different from instruments that are language-reduced tests where examinees must understand spoken test directions in order to attempt the required tasks, even if those tasks do not require a verbal response from the examinee.   As do all assessment tools, nonverbal tests vary on important characteristics including comprehensiveness, psychometric quality, representation of diverse groups in the standardization sample, and appropriateness for individual or group administration.

In addition, nonverbal tests may be classified as *unidimensional or multidimensional*. Unidimensional tests are those that use progressive matrices to measure a narrow aspect of intelligence. The Test of Nonverbal Intelligence-third edition (TONI-III), C-TONI, Naglieri Nonverbal Ability Test (NNAT), and RPM are all unidimensional tests. In contrast, it is suggested that multidimensional tests are more comprehensive and assess a broader range of cognitive skills such as attention, memory, and reasoning. The UNIT and Leiter-R are typically considered multidimensional tests; however, interpretation of extended cognitive skills from the assessment results should be corroborated with other supporting data.

Appendix B

Table Providing a Brief Summary of the Critical Features of Seven Nonverbal Tests of Intelligence Currently Used by School Psychologists

Examples of Nonverbal Tests of Intelligence

Test Name Author Publisher Publication Date	Administration and Scoring Descriptions	Educational and Clinical Applications
Leiter International Performance Scale-revised *(Leiter-R)* Roid & Miller Stoelting 1997	• individual administration • twenty subtests forming 2 batteries as well as 4 socio-emotional rating scales • ten subtests used to assess visualization and reasoning and 10 used to assess attention and memory • age range: 2 - 20 years • hierarchical model of intelligence with factors changing with age of examinee • measures the ability to perform complex nonverbal mental manipulations related to conceptualization, inductive reasoning, and visualization • two IQ scores available • brief IQ screener and full scale IQ with mean of 100 and standard deviation of 15 • composite scores available for rating scales • visualization/reasoning subtests, which require no verbalization by the examiner, used exclusively to compute the full scale IQ	• designed for early identification of cognitive delays • can assess small increments of change in intelligence • reliable and valid measure of intelligence irrespective of language or motor ability • useful for students who have limited expressive language skills and/or poor motor coordination

10

Test Name / Author / Publisher / Publication Date	Administration and Scoring Descriptions	Educational and Clinical Applications
Universal Nonverbal Intelligence Test (UNIT) Bracken & McCallum Riverside Press 1998	• individual administration with 6 subtest and three possible formats: abbreviated, standard and extended • age range between 5 and 18 • administration time 15-45 minutes depending on format administered • based on two-tier model of intelligence • yields quotients for memory, reasoning, (symbolic and nonsymbolic) and a full scale IQ • mean equal to 100 with standard deviation of 15	• designed to assess students from diverse cultural backgrounds, LEP, language and/or hearing impairments, learning disabilities, mental retardation, and emotional disorders
Comprehensive Tests of Nonverbal Intelligence (C-TONI) Hammill, Pearson & Wiederholt Pro-Ed 1997	• individual administration • forty-five to 60 minute administration time • age range between 6 – 89 years • matrix-based, multiple choice format • allows pointing response by examinee • may be administered orally or with pantomimed directions • PNIQ, GNIQ, and NIQ: mean of 100 and standard deviation of 15	• subtests assess analogical reasoning, categorical classification, and sequential reasoning using pictures of familiar objects and abstract geometric designs
NNAT Naglieri-Nonverbal Abilities Test Naglieri Harcourt Brace Educational Measurement 1997	• revision and extension of matrix analogies test • group or individual administration • administration time is 30 – 45 minutes; grades k – 12 • 7 levels, including fall and spring grade-level norms • four types of matrix reasoning (pattern completion, reasoning by analogy, serial reasoning, and spatial reasoning) • brief verbal directions in both Spanish and English • designed to assess performance independent of stored acquired knowledge • yields a nonverbal ability index (NAI) with mean of 100 and standard deviation of 15	• useful for large-scale assessment or as an intellectual screening measure • co-normed with Stanford-9 and Aprenda-2

Test Name Author Publisher Publication Date	Administration and Scoring Descriptions	Educational and Clinical Applications
RPM *Raven's Progressive Matrices* Raven, Raven, & Court 1998	• three levels • colored pm (lowest level of difficulty); use for students 5-11 years and with students suspected of having cognitive delays; 15-30 minute administration time • standard pm (average level of difficulty); suitable for use with general population ages 6-80; 20-45 minute administration time • advanced pm highest level of difficulty; for use with adults or for gifted assessment; 40-60 minute administration time	• useful for individual or group assessment; measures reasoning and problem solving; APM can be used for gifted assessment if conditions warrant a nonverbal test
TONI-III *Test of Nonverbal Intelligence-III* Brown, Sherburne & Johnsen Pro-Ed 1997	• individually administered • age range between 6 – 89 years • 15 – 20 minutes administration time • pantomimed instructions • multiple choice response format with provisions for alternate responses for individuals with severe disabilities • yields an IQ with mean 100 and standard deviation of 15 • measures abstract reasoning and problem solving	• use as screening tool • not appropriate as measure of overall intellectual functioning or ability
Kaufman Assessment Battery for Children: second edition (K-ABC-II) Kaufman & Kaufman AGS, 2004	• individually administered • provides a nonverbal index score (NVI) • age range for NVI: 3-18 years • 20-30 minute administration time • requires administration 4 or 5 subtests depending on age of examinee • nonverbal index score (NVI): mean of 100 and standard deviation of 15	• co-normed with the K-TEA-II • reduced discrepancy in scores between ethnic groups in normative sample • administration appropriate for deaf or hearing impaired children, children with moderate to severe language disorders, and children who have LEP