

Technical Assistance Paper

12684

Use of Part Scores with Tests of Intelligence**Purpose**

This is an update of a previous technical assistance paper (TAP) on the use of part scores with tests of intelligence (1996). While some of the content is the same, several new intellectual assessment tools have been published in the past year, and the new structure of some of these tests has changed the way the results are interpreted. The purpose of this technical assistance paper is to provide guidance to districts in the use of part scores in intelligence testing when considering students for eligibility for exceptional student education (ESE) programs and services. The information included in this TAP is based on statistical and psychometric research. It is recommended that professionals referring to the TAP collaborate with others in their districts who may have specific training on these topics.

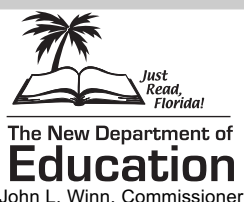
Districts currently engage in a broad range of practices relative to the use of part scores. In many cases, districts have developed guidelines and procedures for the use of part scores.

The intent of this TAP is to bring together professional guidelines and best practices in the use of part scores. No document, however, can address every unique situation that might arise in the intellectual evaluation of a student. The ultimate decision regarding the interpretation of any formal assessment information must be based on the professional training and judgment of the evaluator and the unique characteristics of the student. It is not common or standard practice for factor scores or part composites to be used as broad measures of intellectual functioning, but in unique situations, these scores may be a better indicator of the student's functioning than the global score. It is recommended that districts develop policies and practices that allow for the use of part score interpretation with tests of intelligence when the clinical implications of the assessment data support such practice.

Background Information**Assessment Practice Considerations**

Many intelligence tests provide multiple scores to reflect a student's abilities. An overall score or global score provides a general statement of a student's functioning on the combined tasks that make up the intelligence test. Different subtests or tasks are often combined into part scores (also referred to as factor scores or partial scores) to represent unique skills and deficits in areas with shared common characteristics (e.g., verbal abilities, motor skills, and visual processing). A decision to use part scores rather than the global score to represent a student's intellectual functioning on a test may be made either before or after the test is administered. If the decision is

REFER QUESTIONS TO:
M. Denise Bishop
Student Support Services
310 Blount Street, Suite 215
Tallahassee, FL 32301
Bishop@tempest.coedu.usf.edu
850/922-3727



TECHNICAL ASSISTANCE PAPERS (TAPs) are produced periodically by the Bureau of Exceptional Education and Student Services to present discussion of current topics. The TAPs may be used for inservice sessions, technical assistance visits, parent organization meetings, or interdisciplinary discussion groups. Topics are identified by state steering committees, district personnel, and individuals, or from program compliance monitoring.

BUREAU OF EXCEPTIONAL EDUCATION AND STUDENT SERVICES

made prior to getting test score data, it is considered an *a priori decision*. In this approach, those situations that call for the use of part scores (e.g., students with sensory impairment or non-English speaking students) are known before the assessment, so the evaluator decides prior to assessment to use a part score (e.g., only the performance score is used for students with hearing impairments or non-English speaking students). The basis for this decision is extensive research indicating the inappropriateness of some portions of intelligence tests with students who have unique characteristics that would yield invalid test results. This method of making decisions about the use of part scores prior to testing has a number of advocates in the psychological assessment field.

In contrast to the *a priori* method, the *post hoc* method uses scores obtained from the test to decide whether part or whole (full-scale) scores should be used. In situations where part scores differ markedly from each other, the advocates of *post hoc* decisions assume that it is inaccurate to use a global score to represent functioning. The primary benefit of the *post hoc* approach is that one is less likely to erroneously find a student eligible or ineligible for services or erroneously describe a student's functioning in a global score. The primary concern is that cases in which differences between scores are due to random influences will be incorrectly considered exceptional, which in turn is likely to inflate artificially the number of students declared eligible or ineligible for programs.

One issue that must be considered in the decision to use part versus global scores is a distinction between "reliable" and "rare" score differences. The distinction is irrelevant for *a priori* decisions because the use of part or whole scores is not linked to score differences but to conditions that are known beforehand to influence test performance. However, the distinction is critical for those making *post hoc* decisions. Reliable differences are those that are large enough to assume statistically that they are not due to chance; however, the discrepancy among scores may be quite common in the general population. A reliable difference may exemplify an individual's learning preference or style rather than an area of deficiency. With a rare difference, the discrepancy is both statistically significant and uncommon in the general population. These score differences are referenced in the technical manuals for most of the frequently used intelligence tests; however, the base rate or rarity of differences is not always reported.

On the Wechsler Intelligence Scale for Children; Fourth Edition, (WISC-IV, 2003), for example, when an individual's full scale intelligence quotient (IQ score) falls within the average range, a discrepancy of 10 points between the verbal comprehension index (VCI) score and the perceptual reasoning index (PRI) score occurs in approximately 20 percent of the population or in 1 out of 5 cases. This is a frequently occurring difference between scores. Rare differences, on the other hand, are those that do not occur simply by statistical chance. For example, A VCI -PRI difference of 19 points or more occurs in 8% or less of the population. Similarly, for students who may be gifted learners, discrepancies in index scores are quite common. Statistically, larger discrepancies between index scores occur more frequently as a student's performance falls outside of the normal range of scores (+/- 2 standard deviations). To determine whether a discrepancy is considered to fall within normal limits or conversely, meets criteria as a clinically significant discrepancy, please consult tables B1 and B2 in the Wechsler Administration and Scoring Manual (2003).

Accurate interpretation of assessment data is critical in generating hypotheses about the cause of a student's learning difficulties. Recent research has shown that measuring the severity of a

student's achievement deficit by monitoring his or her response to evidence-based interventions is more educationally informative than using the discrepancy between intellectual functioning and achievement on standardized tests (Reschly, 2003; Prasse, 2004). Information attained from monitoring a student's progress with authentic assessment (i.e., curriculum-based measurement, Dynamic Indicators of Basic Early Literacy: DIBELS) is directly linked to functional academic interventions. With the reauthorization of the Individuals with Disabilities Education Improvement Act (IDEA, 2004), a problem-solving model that analyzes data regarding the student's response to interventions in the general education setting is included as one method of assessing the severity of a student's learning difficulties. While the discrepancy model for determining eligibility in programs for specific learning disabilities remains the standard of practice in Florida, the response to intervention (RTI) method of determining a student's eligibility for exceptional education services is well supported in the fields of special education and school psychology. In fact, the discrepancy formula used in school districts in Florida is a guide and should not supercede the judgment of the evaluation team (OSEP, 1989). Evidence from a student's classroom portfolio, data derived from interventions delivered across settings, and the student's rate of progress toward established academic goals should be considered in all decisions related to eligibility

Sample of Commonly Used Tests and Related Professional References

The Wechsler Intelligence Scale for Children, fourth edition (WISC-IV, 2003)—The WISC-IV provides a measure of general intellectual functioning (FSIQ), as well as four index scores: perceptual reasoning index score (PRI), verbal comprehension index score (VCI), working memory index score (WMI), and processing speed index score (PSI). The dual IQ scores (verbal and performance) and index score structure implemented in the WISC-III is no longer utilized in the WISC-IV, and the four factor structure, initially an optional means of interpretation, has been psychometrically strengthened. The publishers suggest that it is statistically appropriate to use the VCI and the PRI in place of the VIQ and PIQ for part score interpretation. It is not recommended that the WMI or the PSI be used in part score analysis for exceptional student education (ESE) eligibility decisions.

The difference between scores required for significance is computed from the standard error of measurement of the difference (SEdiff). The differences between WISC-IV index scores required for statistical and clinical significance are presented in table B.1 of the *Administration and Scoring Manual* for the WISC-IV. The frequencies of differences between index scores that occurred in the standardization sample for the WISC-IV are presented in table B.2 of the same manual. Even when a score discrepancy is considered to be statistically significant, it may not be clinically significant and as such, may reflect a learning style preference rather than a processing disorder. Keep in mind that the difference required to be clinically significant (rarely occurring) varies depending on the examinee's age and the range of scores obtained.

Harcourt Publishing has also released tables that provide standard scores for an index referred to as the general ability index (GAI). The GAI is calculated using just the VCI and PRI from the WISC-IV. For students considered for gifted eligibility, the publishers have acknowledged that the format of the WISC-IV, with four independent indices contributing to the overall full-scale IQ score, can provide depressed overall scores for students with potentially high cognitive reasoning skills. The working memory and processing speed indices are frequently lower than

the verbal comprehension and perceptual reasoning indices for many intellectually capable students and as such, tend to depress the full-scale score. When eligibility for gifted programs is based in part on the student's IQ score exceeding two standard deviations above the mean, if the WMI or PSI is significantly lower than the VCI and PRI, the GAI score may be used rather than the full-scale score. The tables with the GAI standard scores are included in the text *WISC-IV: Clinical Use and Interpretation* (January, 2005). In addition, the tables can be accessed through Harcourt Assessment, Inc. at <http://www.Harcourtassessment.com>.

There may be cases for which use of an index score on the WISC-IV is used as part of the discrepancy analysis to determine eligibility for ESE services as a student with a specific learning disability. Given that the index scores on this test are comprised of just three subtest scores, it is recommended that the evaluator have additional data to support a decision to use a part score in determining eligibility as a student with a disability.

The Stanford-Binet, 5th Edition (SB5, 2003)—The SB5 provides a 5-factor model. These five factors include fluid reasoning (FR), knowledge (KN), quantitative reasoning (QR), visual-spatial reasoning (VS) and working memory (WM). The five factors then combine to generate an FSIQ, VIQ and NVIQ on the SB5. Use of the verbal and nonverbal IQ scores for part score analysis is acceptable when used with professional discretion; however, factor scores should not be interpreted independently to make eligibility decisions for ESE services. The examiner should use these scores to determine patterns of strengths and weaknesses and to generate hypotheses about possible skill deficiencies. Data that guide the statistical comparison of the verbal and non-verbal IQ scores are published in the technical manual for the SB5 (appendices C and D).

The Kaufman Assessment Battery for Children: Second Edition (K-ABC-II, 2004)—The K-ABC-II does not provide for use of part score analysis, but the K-ABC-II does allow for some a priori decision-making on the part of the evaluator. The structure of the K-ABC-II requires the examiner to determine whether he or she will include the measure of crystallized knowledge (Gc) in the overall cognitive score *before* testing begins. The authors of the test state in the manual, “measures of Gc should be excluded from any score that purports to measure a person's intelligence or overall cognitive ability whenever the measure of Gc is not likely to reflect that person's level of ability.” Therefore, it is recommended that if the examiner suspects, based on a review of existing student data, that including measures of acquired knowledge/crystallized ability would compromise the validity of the global score, he or she should use the Luria-based global score (MPI) rather than the Cattell-based global score (CFI).

Reynolds Intellectual Assessment Scales (RIAS, 2003)—The RIAS does not lend itself to part score interpretation because of the limited number of subtests contributing to the Index scores. Since only two subtests make up the verbal intelligence index and two make up the nonverbal intelligence index, making eligibility decisions based on these indices independent of the composite index is not psychometrically sound practice. If a significant scatter emerges on the RIAS profile, it should be an indicator to the examiner that further assessment is warranted to determine the cause of the variance in scores.

Most test manuals provide special testing considerations for examinees who have limited English proficiency or a language processing disorder, examinees who are deaf or hard-of-hearing, or examinees who are physically or visually impaired. The use of part scores as an estimate of overall intellectual functioning should only be used when supported by clinical data. It is *not*

recommended that districts establish a policy that either denies or accepts the use of part score interpretation under *all* circumstances. Each decision should be made by reviewing the parameters of the individual evaluation situation.

Questions and Answers Related to the Use of Part Scores

1. Which variables should be addressed when considering a part score in eligibility decisions?

The examiner should be aware of any existing factors that may impact the student's performance on the test. For example, some student-centered variables that may affect performance include fine motor difficulties, speech or language concerns, and hearing or vision impairments. In addition, the examiner should be aware of any mitigating conditions in the testing environment such as extraneous noise, time of day, or rapport between the examiner and student. All of these variables can contribute error to the scores attained. If any one of these variables substantially impacts the student's attention and/or response to test items, the results may need to be interpreted with caution. In this case, the use of a part score may provide the best estimate of current intellectual functioning for the student.

2. Are there state guidelines on the use of part scores?

There is a TAP on the use of partial scores that was developed in 1996; however, so many of the tests of intelligence have been updated that the TAP is no longer applicable.

Districts are advised to set up procedures to address the use of part scores in exceptional student education eligibility decisions. The appropriateness of the use of part test scores must be determined on a case-by-case basis by the evaluator, who then makes recommendations to the staffing team. It is not recommended that districts eliminate the consideration of part scores in eligibility decisions. When professional judgment suggests that the part score is a better estimate of the student's current functioning and clinical justification is provided in the psychological report, it is best practice to discuss these data as a team. The decision to accept a part score should be based on procedures that enable accurate assessment and interpretation of data for all students.

3. Do test manuals provide guidance in determining when part test scores are appropriate?

Most test manuals give general guidelines about the use of part scores, and they tend to reference obvious circumstances (e.g., students with sensory impairments, linguistic differences, or physical challenges) when the use of part scores can be determined a priori. When the decision is post hoc, examiners will need to analyze closely the conditions of the testing environment and the possible interference of undocumented areas of disability.

4. What level of occurrence is considered “rare” when interpreting part score differences?

In general, if a difference between scores occurs in less than 10-15% of the normative sample, then it is considered to be clinically rare. Most practitioners interpret a score difference that occurs in less than 15% of the population sampled to be a rare score discrepancy.

5. How should an examiner document that he or she recommends use of a part score?

The psycho-educational report **must** provide written, clinical justification that a part score is the best estimate of the current level of intellectual functioning for that particular student. As always, data that corroborates the decision should be referenced in the report as well. The core issue is to ensure that the best, most reliable estimate of the student’s overall functioning is being used in the decision-making process. In addition, there should be evidence from the student’s learning environment that supports or corroborates the decision to use a part score.

6. How are districts in Florida addressing the issue of part scores and exceptional student education eligibility decisions?

A great deal of variance exists among districts with respect to using part scores as a best estimate of overall ability. With the larger school district, it is likely that a written policy has been established. It is not common or standard practice for factor scores or part composites to be used as broad measures of intellectual functioning, but in unique situations, these scores may be a better indicator of the student’s functioning than the global score.

7. Is it accurate to state that in some cases when a part score is used, a student’s level of cognitive functioning may be determined based on the results of just three subtests?

This is true. If the school psychologist were to use the VCI or PRI on the WISC-IV, the index score is comprised of just three subtests. While the VCI may be the best representation of the student’s functioning on that particular test, it may not provide enough information to generalize to his or her overall functioning. In this case, additional assessment may be necessary to make a reliable statement about intellectual functioning. Once again, clinical judgment is critical.

8. How should districts document their policy and procedures for using part scores in ESE eligibility decisions?

Districts should provide written justification in the psychological report in those instances when a part score is used in lieu of a full scale or composite score. Justifications may include but are not limited to documented sensory or physical impairments, language and linguistic factors, or the administration of another test of intellectual functioning that offers additional support for the use of part scores. To say that the part

score is a better indicator of the student's ability **does not** adequately address the justification requirement because it does not include any clinical information from the individual student assessment.

9. What are some resources that may be used to guide decisions related to the use of part score?

Examiners can reference the professional journals in their field as well as the technical manuals that accompany the tests of intellectual functioning. The following resources may have recent literature that addresses this issue: *The Journal of School Psychology*, *School Psychology Review*, *Psychology in the Schools*, *School Psychology Quarterly*, and *Student Assessment News*. In addition, the National Association of School Psychologists (NASP) and the American Psychological Association (APA) frequently publish technical assistance papers that guide practitioners in best practices in assessment. NASP can be accessed on-line at <http://www.nasponline.org>, and APA can be accessed at <http://www.apa.org>.

10. Can an IEP team apply the standard error of measurement to the score as well as using a part score when determining eligibility for an ESE program?

No. It is not recommended to use both the standard error of measurement and a part score rather than the full scale or global index. If the student profile is significantly varied and the data do not adequately answer the referral questions regarding the student's learning concerns, which extend beyond eligibility determination, then further assessment is recommended.

11. What variables should be considered when writing district policy regarding the use of part scores?

If in conjunction with test manual guidelines a district can address and answer the following questions, it should be able to develop a fair and ethical policy regarding the use of part scores for ESE eligibility.

- Under what conditions will the district accept the use of part scores?
- What types of written justifications are appropriate for the use of part scores?
- How will the use of the part score be documented in the psychological report?
- Is the school psychologist required to participate in staffing meetings where part scores are being considered?