

## From Conceptual to Practical: Making Gifted Testing Relevant

Barbara Gilman, M.S.  
Kathi Kearney, M.A. Ed.  
Gifted Development Center  
Denver, Colorado

Conceptual Foundations Division  
National Association for Gifted Children  
51<sup>st</sup> Annual Convention

Salt Lake City, Utah  
November 7, 2004

Not since Terman's 1916 Stanford Binet, the Terman Concept Mastery Test, and the original SAT, have major tests actually been designed to assess giftedness. Finding gifted children today requires instruments that both assess the child's strengths and remain sensitive to gifted characteristics. A new generation of IQ tests has recently been released. We are in the transition phase now, using the new tests and finding what works and doesn't work, or which portions to utilize to satisfy various needs. New tests are based on stronger theoretical foundations; however, they still represent both improvements and losses for the gifted. We have seen some extension of test ceilings, and enhanced assessment of nonverbal and visual-spatial abilities, which benefits the gifted, as well as increased emphasis on processing skills and reduced emphasis on verbal reasoning skills, which does not. As tests change, different groups of children are identified by gifted-level scores. Clinicians need a variety of test instruments, and an in-depth knowledge of their use, to document gifted children's strengths and their need for special programming. The tests we find useful include the following:

**Wechsler Tests.** Wechsler tests continue to be good initial IQ tests for the gifted. Although they rarely yield scores above the 140s, they offer useful diagnostic information, and indications of ability beyond their limits.

**WISC-IV** (Wechsler Intelligence Scale for Children-Fourth Edition)- replaces the WISC-III. The WISC-IV's increased ceiling offers additional harder questions at the upper ends of a number of subtests. The Verbal IQ and Performance IQ scores of its predecessors have been eliminated. The 10 required subtests (5 are supplementary) yield a Full Scale IQ score and four Composite scores: Verbal Comprehension, Perceptual Reasoning, Working Memory, and Processing Speed. The Verbal Comprehension and Perceptual Reasoning Composites are very good indicators of giftedness, that do an admirable job of assessing verbal abstract reasoning and provide very useful tests of visual reasoning with less timing emphasis. Working Memory and Processing Speed are less correlated with giftedness. The Dumont-Willis Indices (check out John Willis and Ron Dumont's website at <http://alpha.fdu.edu/psychology/>) offer another approach to evaluating WISC-

IV scores besides the Full Scale IQ when Verbal Comprehension and Perceptual Reasoning Composite scores are higher than Working Memory and Processing Speed (the WISC-IV technical manual suggests this will usually be the case). A DWI-1 score can be computed for the combination of Verbal Comprehension and Perceptual Reasoning, while a DWI-2 score can be computed for the combination of Working Memory and Processing Speed. These computations, based on the Tellegen and Briggs formula, could be helpful to schools. The DWI-1 score would be an excellent identifier of gifted children for school programs, and only six subtests of the WISC-IV are needed to produce it.

**Recently, The Psychological Corporation, publisher of the WISC-IV, has trained testers in the calculation of a General Ability Index (GAI). Utilizing the same approach as the DWI-1 score, the GAI combines scores from the Verbal Comprehension and Perceptual Reasoning Composites. Look for tables on the Harcourt Assessment website by the end of 2004. Similar tables in Flanagan and Kaufman's *Essentials of WISC-IV Assessment* (pp. 331-333) are available until then.**

The WISC-IV is yielding many gifted-level scores at the GDC. However, some of the Full Scale IQ scores are excessively lowered by Working Memory and Processing Speed scores. As intelligence is primarily abstract reasoning ability, emphasizing short-term auditory memory and processing speed on paper-and-pencil tests is less helpful. Two Working Memory subtests (only one was required on the WISC-III) and two Processing Speed subtests (only one was required on the WISC-III) place more weight on these processing skills in the Full Scale IQ score. This is unfortunate for gifted children and confounds the FSIQ as a gifted identifier at times. Flanagan and Kaufman advise against the use of Full Scale IQ scores if composite scores vary by 1.5 SDs or more (at least 23 points) (p. 128). If the Full Scale score is inappropriate and the standard score difference between Verbal Comprehension and Perceptual Reasoning is less than 23 points, "then the GAI may be used to describe overall intellectual ability." (p. 128)

In the normative sample for the WISC-IV, a group of 63 gifted children (which had previously scored at least two standard deviations above the mean) earned the following scores:

Full Scale IQ	123.5
Verbal Comprehension	124.7
Perceptual Reasoning	120.4.
Working Memory	112.5
Processing Speed	110.6

(WISC-IV Technical Manual p. 77). It is notable that the gifted group in the normative sample scored lower in Working Memory and Processing Speed.

A study of 103 children tested on the WISC-IV at the Gifted Development Center (GDC) yielded the following scores:

Verbal Comprehension	131.69
Perceptual Reasoning	126.42
Working Memory	*117.65
Processing Speed	104.33.

\*Arithmetic was substituted for Letter-Number Sequencing unless a child showed evidence of mathphobia.

Scores for the Gifted Development Center sample showed even greater variance, despite higher Working Memory scores, likely due to the substitution of Arithmetic for Letter-Number Sequencing. This was done because Arithmetic has a higher *g*-loading (Kaufman, p. 308) and is more engaging for most gifted children, as well as our experience that some children responded unpredictably to L-N Sequencing. Arithmetic might have been substituted for Digit Span instead. The variance of the GDC sample is greater than 23 points, suggesting that all WISC-IV score profiles of gifted children should be scrutinized before calculating and reporting a Full Scale score.

**WPPSI-III** (Wechsler Preschool and Primary Scale of Intelligence-Third Edition) The WPPSI-III offers improvements over its predecessor for gifted children. It appears to be a better diagnostic test and it emphasizes abstract reasoning well. It has two forms: one for children ages 2-6 (2 years 6 months) to 3-11, and one for ages 4-0 to 7-3 (we prefer a WISC-IV at 6-0). The form for very young children is short (4 required subtests) and only Object Assembly is timed (rather generously). The portion for ages 4-0 and up requires 7 subtests and has a good mix of verbal and visual reasoning in a child-friendly test (administration is more comfortable than with the WPPSI-R for little ones). The test does have a Processing Speed Quotient, but only one of the two subtests from which it is calculated is included in the Full Scale IQ score. This is a timed handwriting-like test. Slower processing speed on handwritten paper-and-pencil tests is common in the gifted and should not rule out admission to gifted programs.

**SB5** (Stanford-Binet Intelligence Scale-Fifth Edition) The SB5 has highly advanced mathematical reasoning and spatial reasoning items. Gale Roid, the test's author, has shown flexibility in considering alternate ways to score this test when young children perform at a level far above age expectations, e.g., using Rasch-ratio scores. Using this different metric, testers can explore and differentiate higher levels of giftedness by comparing test age to chronological age. Roid has also condoned omitting a portion of the SB5 if it is deemed less relevant for gifted identification (e.g., the Working Memory portion) and prorating the score. However, the reporting of more than one score can be confusing for schools, especially if one score meets gifted program admission standards and another does not. Efforts are under way to clarify the appropriate uses of these

scoring approaches. Because the SB5 is limited in its assessment of verbal abstract reasoning, it is useful with gifted children with verbal expression issues. The SB5 has continued the Binet tradition of being largely untimed, and potentially offers a very high ceiling when Wechsler tests prove too limited.

**SBL-M** (Stanford-Binet Intelligence Scale—Form L-M)

The SBL-M's use is still marginally supported by Riverside Publishing as it remains probably our strongest test of verbal abstract reasoning (with excellent mathematical and spatial items) and has established its usefulness over many years. Despite older norms, the test continues to successfully discriminate between higher levels of giftedness, beyond the limits of Wechsler tests, allowing professionals to fine tune educational recommendations. It has been the test of choice in research with exceptionally and profoundly gifted children, such as Miraca Gross's study of exceptionally and profoundly gifted children in Australia. Revising and renorming this instrument as a test of giftedness would be appropriate. Without it, a population of extremely high-risk children becomes invisible and their needs become virtually impossible to address. The SBL-M and SB5 show real strengths used together because their content varies significantly.

**What are the differences between the SBL-M and SB5?** The SBL-M is an excellent indicator of verbal abstract reasoning—required in gifted classrooms where students discuss concepts at advanced levels, compare and contrast ideas, make inferences, debate and write at depth about issues. This kind of reasoning is primarily assessed in only 1 of the 10 subtests of the SB5: Verbal Fluid Reasoning. Although the SB5 is composed of 5 "Verbal" and 5 "Nonverbal" subtests, this designation refers only to the way the items are administered. Other Verbal subtests include "Verbal Quantitative Reasoning," "Verbal Visual-Spatial Reasoning," "Verbal Working Memory," and "Verbal Knowledge" (vocabulary). So, of 10 subtests, only 2--Verbal Fluid Reasoning and Verbal Knowledge (vocabulary), assess the most important elements of verbal comprehension in the typical sense.

However, because the Nonverbal portion of the SB5 includes Nonverbal Fluid Reasoning (matrices), Nonverbal Knowledge (visual reasoning), Nonverbal Quantitative Reasoning, Nonverbal Visual-Spatial Processing, and Nonverbal Working Memory, the SB5 is a strong assessment of visual reasoning, visual-spatial reasoning (presented both verbally and visually) and mathematical reasoning (both verbal and visual). It expands areas the SBL-M taps. The SBL-M is a wonderful test of verbal abstract reasoning, which also has some excellent mathematical and visual-spatial items, while the SB5 emphasizes mathematical and spatial reasoning. The tests complement each other.

It is unlikely the SBL-M will be significantly redesigned to include more visual-spatial items because producing an entirely new IQ test like the SB5 is very costly. However, Riverside would be wise to continue supporting the SBL-M because its content exists nowhere else and it is a proven test of giftedness. Just slightly revising it and renorming it as solely a test for the gifted is greatly needed.

*“You’re Asking Me to Do WHAT With Those  
Letters/Numbers/Words/Pictures That You Just Told/Showed Me?????????”*

### **The Increased Emphasis on Processing Skills Measures as Part of the FSIQ**

\*Perhaps\* the inclusion of more processing skills measures is appropriate for lower functioning children. If the child's processing speed on paper-and-pencil tasks is so slow that he or she cannot complete work in a reasonable amount of time in the classroom, processing speed may be such a limiting factor that it should be included in IQ scores. Likewise, if short-term auditory memory is so poor that the teacher's instructions can't be retained at all, this is a significant problem. However, gifted children rarely perform extremely poorly in these areas on an absolute scale. It makes much more sense to identify them as gifted based on assessments emphasizing reasoning, provide them gifted learning experiences, and then add any accommodations based on relative weaknesses to the gifted accommodations. A Full Scale IQ score that averages *gifted* reasoning and *average* processing skills fails to identify either the giftedness or the relative weaknesses.

Test authors have wrongly assumed gifted children are fast processors. Some are very quick; others are reflective or perfectionistic, slowing their speed. Gifted children also show a preference for meaningful test materials, and may not perform well on short-term memory tests or other tasks that utilize non-meaningful material. They usually perform so much better with meaningful material that their scores with non-meaningful material are difficult to interpret.

If a strand is added to an IQ test that identifies a different group as scoring the highest than was identified by the other strands, there will be a *confounding* of the Full Scale IQ score. The newly revised and renormed tests do exhibit confounding in the Full Scale scores (note the fact that the gifted group in the WISC-IV normative sample scored a 124.7 on Verbal Comprehension and a 120.4 on Perceptual Reasoning, but only earned a 112.5 in Working Memory and a 110.6 in Processing Speed, according to the WISC-IV Technical Manual, p. 77). Given these issues, it will be a challenge for testers of the gifted to choose tests appropriate to document gifted strengths and diagnose weaknesses, without eliminating children from gifted program entrance requirements.

### **Nonverbal Tests**

New tests are being marketed to assess intelligence on visually-presented items, with either verbal or nonverbal instructions. We find these tests to be excellent additional measures for gifted assessment—to document visual reasoning abilities or identify capable children with language limitations—providing there are other options, as well. The visual reasoning abilities of children can be obscured by visual processing deficits on such tests when complex visual patterns are involved. Likewise, we find that some children simply demonstrate stronger verbal than visual reasoning. Furthermore, it is important to use a test that is compatible with the program offered. If the program is a gifted classroom, emphasizing conceptual complexity in discussion, debate, and in-depth

writing, a nonverbal test may not identify the children who fit best. One child at the GDC, who had previously scored at the 99.9<sup>th</sup> percentile on a Ravens Progressive Matrices Test, and was offered admission to a highly gifted program in a California school, subsequently scored no higher than the 75<sup>th</sup> percentile on any WISC-III subtests. An autistic boy, his strength was in pattern recognition, and gifted programming for him would need to revolve around such ability. In general, classrooms for children with gifted visual-spatial strengths need to employ related activities—visual/hands-on projects, constructions, teaching approaches that utilize visual aids (maps, models, etc.). Typically, identification procedures that employ varied measures are more successful identifying a variety of gifted children.

Dumont, R. & Willis, J. *Dumont Willis*. <http://alpha.fdu.edu/psychology/> (7 Nov. 2004).

Falk, R.F., Silverman, L.K., & Moran, D.M. (2004, November). *Using two WISC-IV indices to identify the gifted*. Paper presented at the 51<sup>st</sup> Annual Convention of the National Association for Gifted Children, Salt Lake City, UT.

Flanagan, D.P. & Kaufman, A.S. (2004). *Essentials of WISC-IV assessment*. Hoboken, NJ: John Wiley & Sons.

Gross, M.U.M. (2004). *Exceptionally gifted children*. London: RoutledgeFalmer.