Breakthroughs in Assessment of the Gifted

Linda Kreger Silverman, Ph.D.
Licensed Psychologist

Nebraska Association for the Gifted
Omaha, Nebraska
February 28, 2013
The new IQ tests have rendered all of these ideas obsolete. Different IQ tests no longer measure the same basic construct of intelligence. Each test identifies a different population as gifted. Gifted students may obtain widely varying scores on different instruments. Full Scale IQ scores are not the unitary constructs they once were; therefore, they may not be the best representation of the child’s intellectual capacities. The familiar Verbal and Performance IQ scores are gone. There are multiple ways of administering and scoring the new instruments. The designation of giftedness is now unclear. Under these circumstances, how should decisions be made about the use of IQ tests in the selection of students for gifted schools, programs and services?

The following recommendations have been drawn from the Symposium on Assessment Techniques in the Identification of Gifted Learners hosted by the World Council for Gifted Children 16th Biennial Conference in New Orleans, Louisiana, August 7, 2005; the Second National Symposium on Assessing Gifted Learners, Van Nuys, California, March 27, 2009; the U.S. National Association for Gifted Children (NAGC) Position Statement on “Use of the WISC-IV for Gifted Identification”; recommendations of the NAGC Task Force on IQ Interpretation; chapters published in Alternative Assessment of Gifted Learners, High IQ Kids and The International Handbook on Giftedness; and several studies presented at NAGC and World Council for Gifted Children Conferences in the last several years. Please review the references for additional information on each recommendation.
Recommendations

1. Individual IQ tests provide better information for high-stakes decision making for gifted students than group tests. Group IQ tests are best used for general screening purposes (Rimm, Gilman & Silverman, 2008).

2. Gifted students are a special needs population; therefore, comprehensive assessment is needed to determine strengths and weaknesses, information which should be used in their programming (VanTassel-Baska & Baska, 1993).

3. When students are selected for gifted programs on the basis of achievement tests, grades and teacher recommendations, programs are likely to miss children who are economically disadvantaged, who have few books in their homes, and fewer role models for achievement. Individual ability tests provide greater access to programs for diverse cultural groups and children of low socio-economic status (Silverman & Miller, 2009).

4. Raven’s Progressive Matrices appears to be the most popular intelligence scale worldwide (Silverman, 2009b). The Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV) is the most popular individual IQ test used in the United States for the selection of students to gifted programs, followed by the Stanford-Binet Intelligence Scale, Fifth Edition (SB5). The WISC-IV and the SB5 offer several methods of scoring. The Full Scale IQ scores generated on these tests do not measure general intelligence as cohesively as in prior editions. When there are significant discrepancies between composite scores on the WISC-IV (23 points or more), the Full Scale IQ score should not be derived (Flanagan & Kaufman, 2004). Therefore, the Full Scale IQ score should not be the main score used to determine program selection (Rimm, Gilman & Silverman, 2008).

5. It is not necessary to calculate a Full Scale IQ on the SB5; it is permissible to use either the Verbal IQ or the Nonverbal IQ independently to locate gifted children with different strengths. The highest index, composite or factor score is often the best predictor of success in the gifted program, if the program is responsive to the learning strengths of the students (Silverman, 2009b).

6. When using the WISC-IV, either the General Ability Index (GAI), which emphasizes reasoning ability, or the Full Scale IQ Score (FSIQ), are acceptable for selection to gifted programs. The GAI should be derived using the table provided by Pearson Assessment (Technical Report 4). “The Verbal Comprehension Index (VCI) and the Perceptual Reasoning Index (PRI) are also independently appropriate for selection to programs for the gifted, especially for culturally diverse, bilingual, twice-exceptional students or visual-spatial learners. It is important that a good match be made between the strengths of the child and the attributes of the program. Students who have special learning needs should be admitted to gifted programs, provided that there are other indications of giftedness
and instructional modifications are made to fit the needs of the students”
(National Association for Gifted Children, 2008).

7. Processing Speed and Working Memory often lower composite and Full Scale IQ
scores for the gifted. The 63 gifted children in the norm sample reported in the
WISC-IV Technical Manual had a mean Working Memory score of 112.5 and a
mean Processing Speed score of 110.6 (Wechsler, 2003). Only 4 of 103 gifted
children tested at the Gifted Development Center scored 130 or above on
Processing Speed (Silverman, Gilman & Falk, 2004). The Family Achievement
Center found Processing Speed to be in the high average range for 42 children
reported that “over 70% of the students applying for gifted placement have
Processing Speed Index scores in the average range or below” (p. 2). On the
WISC-IV, the GAI omits both Processing Speed and Working Memory in the
scoring. In using the SB5, a Gifted Composite Score and a Gifted Nonverbal
Composite score can be generated in which Working Memory is eliminated
(Silverman, 2009b).

8. Gifted children, particularly those with learning disabilities, succeed on difficult
items while failing easier items. They are often capable of answering items
beyond the discontinue criteria. The Stanford-Binet Intelligence Scale (Form L-
M) allowed credit for all correct items, and examinees were offered the
opportunity to define all the vocabulary items on the exam (Silverman, 2009a). In
modern testing, discontinue criteria have been reduced to as few as 2 or 3 items,
which penalizes twice exceptional children (Silverman, 2009c). It is
recommended that more flexible use of discontinue criteria be employed to test
the limits of children’s abilities and that both standardized and non-standardized
scores be reported. It is also wise to report the number of subtests where no
discontinue criterion was reached. This indicates an underestimate of ability.

9. As giftedness and disabilities mask each other, individual assessment of twice
exceptional children is strongly recommended. To locate twice exceptional
children, it is necessary to examine intrapersonal variables rather than comparing
the child to the norms for average children: “To what extent does the discrepancy
between this child’s strengths and weaknesses cause frustration and interfere with
the full development of the child’s abilities?” (Silverman, 2009c).

10. Funds for assessing gifted students are usually limited. For this reason, many
districts employ short forms of individual intelligence scales, such as the
Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999). A
reasonable alternative would be to administer only the six core subtests of the
WISC-IV from which the General Ability Index (GAI) can be derived:
Vocabulary, Similarities, Comprehension, Block Design, Matrix Reasoning and
Picture Concepts (Silverman, Gilman & Falk, 2004). As most of these subtests
are richly loaded in general intelligence (g), they are likely to locate the students
who would be most successful in a gifted program.
11. The gifted validation sample reported in the *Technical Manual* of the WISC-IV achieved a mean Full Scale IQ score of 123.5 (Wechsler, 2003). The mean IQ score of 202 children in the gifted validation sample of the SB5 was 124. Therefore, cut-off scores for gifted programs should be lowered to 120, rather than 130 (Rimm, Gilman & Silverman, 2008; Silverman, 2009b).

12. In selecting an instrument to use for assessing the gifted, it is necessary to keep in mind that the WISC-IV and the SB5 identify different students as gifted. As 30% of the WISC-IV measures abstract verbal reasoning, compared to 10% of the SB5, the WISC-IV is likely to find more highly verbal children. As 20% of the SB5 measures mathematically gifted children, compared to 0 – 10% of the WISC-IV (depending if Arithmetic, an optional test, is administered), the SB5 is likely to find more mathematically gifted children. While 20% of each test is devoted to the measurement of visual-spatial abilities, there may be more visual-spatial content in the SB5, so it may be preferable for locating spatially gifted children.

13. Those instruments or portions of instruments with the richest loadings on general intelligence ($g$) are the most useful for locating gifted children. *Raven’s Progressive Matrices*, the Stanford-Binet scales and the Wechsler scales were all founded on the conception of intelligence as abstract reasoning ($g$). Abstract reasoning and general intelligence ($g$) are synonymous. Giftedness is high abstract reasoning. Therefore, $g$ could as easily stand for giftedness as for general intelligence (Silverman, 2009b).

**$g$-loadings on the WISC-IV**

<table>
<thead>
<tr>
<th>Good Measures of $g$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary</td>
<td>.82</td>
</tr>
<tr>
<td>(Information)</td>
<td>.79</td>
</tr>
<tr>
<td>Similarities</td>
<td>.79</td>
</tr>
<tr>
<td>(Arithmetic)</td>
<td>.74</td>
</tr>
<tr>
<td>(Word Reasoning)</td>
<td>.70</td>
</tr>
<tr>
<td>Comprehension</td>
<td>.70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fair Measures of $g$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Matrix Reasoning</td>
<td>.68</td>
</tr>
<tr>
<td>Block Design</td>
<td>.67</td>
</tr>
<tr>
<td>(Picture Completion)</td>
<td>.63</td>
</tr>
<tr>
<td>Letter-Number Sequencing</td>
<td>.60</td>
</tr>
<tr>
<td>Symbol Search</td>
<td>.58</td>
</tr>
<tr>
<td>Picture Concepts</td>
<td>.57</td>
</tr>
<tr>
<td>Digit Span</td>
<td>.51</td>
</tr>
</tbody>
</table>
Poor Measures of g

Coding (.48)
(Cancellation) (.25)

*Items in parentheses are optional tests.
(Flanagan & Kaufman, 2004, p. 309)

Recommendations for Differentiating Children at Higher Levels of Intelligence

14. Exceptionally gifted children are among the highest risk gifted populations (Rimm, Gilman & Silverman, 2008). An international group determined nomenclature for the higher IQ levels. Their results were published in a chapter on “Assessment of Intellectual Functioning” by John Wasserman (2003).

Table 5
Levels of Giftedness

<table>
<thead>
<tr>
<th>Level</th>
<th>IQ Range</th>
<th>Standard Deviations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profoundly Gifted</td>
<td>above 175</td>
<td>+5 SD</td>
</tr>
<tr>
<td>Exceptionally Gifted</td>
<td>160 – 174</td>
<td>+4 SD</td>
</tr>
<tr>
<td>Highly Gifted</td>
<td>145-159</td>
<td>+3 SD</td>
</tr>
<tr>
<td>Gifted</td>
<td>130-144</td>
<td>+2 SD</td>
</tr>
</tbody>
</table>

(adapted from Wasserman, 2003, p. 435)

15. The low ceilings on the Raven’s, WISC-IV and SB5 make it difficult to locate highly, exceptionally and profoundly gifted children. To document that the child’s abilities exceed the measuring tool, Betty Meckstroth created a method of tracking the number of raw score points earned beyond the minimum score required to attain a 19 (the highest possible subtest score on the WISC-IV and SB5). One child scored 13 raw score points beyond the ceiling on Vocabulary (19 + 13) and 8 extra points on Similarities (19 + 8) on the WISC-IV (Rimm, Gilman & Silverman, 2008). Another child scored 12 extra raw score points on Vocabulary and 8 on Similarities (Zhu, Cayton, Weiss, & Gabel, 2008).

16. As of February 7, 2008, a new set of extended norms is available for the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV) for assessing exceptionally and profoundly gifted children. The WISC-IV extended norms were developed in response to a request from the National Association for Gifted Children (NAGC) Task Force on IQ Interpretation. NAGC sponsored a study of WISC-IV scores of 334 gifted children from 8 sites and the data were sent to the publisher. Based on these data, the maximum subtest scaled score was raised from 19 to 28. The maximum index (composite) and full scale scores were raised from 160 to 210. It is necessary to achieve subtest scaled scores of 18 or 19 for
the extended norms to apply. The new norms were posted on the Pearson Assessments website in *WISC-IV Technical Report #7:*


17. The *Interpretive Manual* of the SB5 (Roid, 2003) offers a table of Extended IQ scores for children who score above 150 IQ or below 40 IQ. Scores range as low as 10 and as high as 225 IQ. Based on Rasch scoring, the examinee is credited with all raw score points beyond the requirement to obtain the ceiling score of 19. As few score above 150, Rasch-Ratio scores may also be derived by hand for students who score 130 and above on the SB5 (Carson & Roid, 2004). It is recommended that students be selected for gifted programs who attain a score of 120 or above on the SB5, and that Rasch-Ratio scores be derived to qualify students for services for the highly and exceptionally gifted or to determine the degree of acceleration needed (Silverman, 2009b).

18. Sylvia Rimm created another scoring method for estimating scores beyond the norms on the WISC-IV. Rimm Ratios can be derived by utilizing test-age equivalents for subtests provided in the WISC-IV manual (p. 253), converted into months, to determine a child’s mental age. Age-equivalent scores reflect all correct items for each subtest. The child’s mental age is divided by the chronological age and multiplied by 100 to derive a Rimm Ratio (Rimm, Gilman & Silverman, 2008). This method is most appropriate for children 10 and under, due to the low ceiling of the test-age equivalents.

19. The two-step process employed in the Talent Searches for differentiating the most able students at 12 or 13 years of age is the best model for locating exceptionally gifted children. Talent Searches provide out-of-level testing to children who score at or above the 95th or 97th percentile in reading and mathematics. Since 1989, younger children in these ranges have been found by using a combination of two different measures: one comparing them with others their own age and one with a higher ceiling, like the SAT, that compares their abilities to those of older children. Because it is organized by age levels, with increasing levels of difficulty, all the way to Superior Adult III, the “Binet-type age scale might be considered the original examination suitable for extensive out-of-level testing” (Stanley, 1990, p. 167).

20. Examiners who assess the exceptionally gifted offer out-of-level testing to children who achieve on a standard IQ test (e.g., the WISC-IV, WPPSI-III, SB5, DAS-2, KABC-2, etc.) at or above the 99th percentile on at least two subtests (particularly those subtests that are good measures of g). In the case of the *Raven’s Progressive Matrices*, a score at the 97th percentile would probably warrant an out-of-level supplementary test. The *Stanford-Binet Intelligence Scale (Form L-M)* is given purely as a *supplemental* measure to test the limits of children’s abilities when they achieve ceiling-level scores on tests with lower
ceilings (Silverman & Kearney, 1989; see also, Silverman & Kearney, 1992; Wasserman, 2007a). When the child exceeds the scores in the norm table of the SBL-M, a formula IQ is derived according to the instructions on page 339 in the manual (Terman & Merrill, 1973). The formula IQ is a ratio metric.

The SBL-M remains unmatched in its breadth of procedures and is probably truer to the changing nature of cognitive-intellectual abilities over development than any test subsequently published. Its unique age-scale format and liberal discontinue rules enable testing to continue far beyond one’s chronological age, thereby providing examinees with an opportunity to demonstrate considerably advanced competencies.

I consider this test appropriate only after an examinee has approached the ceiling of a more recently normed test (such as the WISC-IV or SB5…), as a method of resolving just how far above the ceiling the examinee’s true abilities may lie. When reported in an appropriately conservative manner (because of its limitations), the ratio IQ approach provides the only available means of estimating intelligence in exceptionally and profoundly gifted ranges that has any prior foundations in research (e.g., the work of Terman and Hollingworth). (Wasserman, 2007, p. 51)

21. It is permissible to use the SBL-M as a supplemental test, as long as examiners acknowledge that the scores are on a different metric and, therefore, not comparable to deviation IQs (Carson & Roid, 2004).

22. The Flynn Effect is the most frequently cited reason for not using a test with older norms. Newer studies suggest that the Flynn effect may have tapered off at the beginning of the 1990s (Teasdale & Owen, 2005). John Wasserman (2007b) recently studied the Flynn Effect and writes:

My January, 2007 examination of psychological research databases suggests that the Flynn effect has not yet been adequately demonstrated for all levels of ability … there is no substantive evidence for its validity with high ability individuals (particularly those who are intellectually gifted). … I have yet to see any sound empirical studies of the Flynn effect in gifted samples. (p. 1)

For more information on the recommendations listed above, please see (Rimm, Gilman & Silverman, 2008; Silverman, 2009b; Silverman, 2007; Silverman & Miller, 2009; Wasserman, 2007a).
References


**Note:** This article may be reprinted on other websites and in newsletters, provided that the author is given full credit, receives a copy of the publication, and is told when and where the article is posted on a website.

October 31, 2011 © Linda Kreger Silverman

Gifted Development Center
1820 Sheridan Boulevard, Suite C-111
Westminster, CO 80003
[www.gifteddevelopment.com](http://www.gifteddevelopment.com)
(303) 837-8378
gifted@gifteddevelopment.com

Gifted Development Center (GDC) is a nonprofit 501(c)3 agency that has tested over 6,000 children over the last 33 years. GDC was instrumental in the development of extended norms on the WISC-IV.
Using Test Results to Support Clinical Judgment

Linda Kreger Silverman, Ph.D.
Gifted Development Center

Medical doctors order tests, but they don’t base their diagnoses solely on results of those tests. They require a complete medical history of the patient’s family, as well as the patient, and analyze test results in conjunction with other information obtained, such as presenting symptoms, medical history, family history, and patient interview. Test results are of limited value unless interpreted by a skilled clinician who has had experience with the presenting problem.

Yet, in the diagnosis of giftedness, high stakes decisions frequently are made on the basis of test scores alone (sometimes even from group-administered tests). Scores become the definitive arbiters of children’s qualification for placement, often determining their futures. Clinical judgment, if employed at all, tends to be subservient to the numbers. Similar to the medical field, accurate assessment of giftedness is dependent upon the skill and experience of the examiner in interpreting protocols of gifted children within the context of all the other information obtained.

The best evaluators of gifted children utilize some aspects of qualitative assessment. With sufficient experience, they can accurately estimate a child’s level of intelligence through clinical observation, a brief discussion with the child, an interview with the parents, developmental milestones, family history, or some combination of these sources of information. Test results are interpreted within this broader framework and judged to be valid only if they conform to the clinical picture that has emerged from a more comprehensive appraisal of the child. If the test results fail to support the examiner’s clinical judgment, further evaluation is sought to determine the cause of the discrepancy. The more experience an examiner has with gifted children, the more effective his or her clinical judgment will be. Obviously, this type of assessment is more time consuming, and, therefore, more costly than simply determining whether or not a child qualifies for a gifted program.

Examiners who assess gifted children need grounding in the developmental and psychometric patterns of the gifted, as well as knowledge of assessment; otherwise, boilerplate analyses are likely to ensue in which numbers take precedence over clinical judgment. Such interpretations are often inaccurate.

Traditional test interpretation involves averaging of subtests measuring verbal ability, perceptual reasoning, working memory and processing speed, from which a Full Scale IQ is derived. Relative strengths and relative weaknesses are reported based on the degree of discrepancy between specific subtest scores and the subject’s average scores in these
areas. The child’s scores are compared to standardized norms to determine if they are above or below the average for their age group. This is standard practice.

While normative interpretation may be suitable for assessments with 95% of the population, it is likely to underestimate the abilities of gifted children—and render acute underestimates for profoundly and twice exceptional children. The assessment issues unique to the gifted population are not commonly understood. Variables such as test ceilings, discontinue criteria, and environmental factors differentially affect the performance of the gifted due to the small number of items that distinguish the gifted from the average.

**Selection of IQ Tests**

While children in the average range score similarly on different measures of intelligence, gifted children exhibit more variability on IQ tests than any other group (Silverman, 1995a). Most IQ tests suffer from ceiling effects that diminish scores in the gifted range. What appears as a “relative strength” on one test may turn out to be an astronomical strength on a test with a higher ceiling. The talent search model serves as a clear example of this principle. Two 7th graders who score at the 97th percentile in mathematics on a 7th grade achievement test may attain radically different scores on the Mathematics section of the SAT, when taken as an above-level test in one of the talent searches: one may score 300 and the other 700. The grade-based achievement test indicates that the two students are in the top 3 percent of students their age, which may qualify them for an advanced mathematics program. But the SAT reveals that one of the two has mastered the math curriculum of the next several years and needs considerably more acceleration than the other.

IQ scores for children in the exceptionally and profoundly gifted ranges vary dramatically depending on the ceiling of the test, item difficulty, discontinue criteria, and whether or not the child is allowed credit for all raw score points earned (as in the extended norms of the WISC-IV). There are several types of ceiling issues in testing the gifted. The highest possible score may not differentiate children in the higher ranges (e.g., 135 on Raven’s Progressive Matrices). The child might answer the hardest problems in a subtest, never reaching discontinue criteria (subtest ceiling). When no ceiling is reached, the score is probably an underestimate; there is no way of knowing how much better he or she might have performed had there been harder items. Maximum test ages need to be examined: a test that tops out at 17 renders different results from a test with a maximum test age of 23 (i.e., the SBL-M, administered as a supplemental test for exceptionally and profoundly gifted children).

Untimed tests are better for locating gifted children than timed tests. Test constructors are often surprised to learn that gifted children are not always faster than children of lesser ability (Reams, Chamrad & Robinson, 1990). Processing speed is usually considerably lower than other composite scores in gifted samples (Rimm, Gilman & Silverman, 2008). Bonus points for speed in IQ tests depress IQ scores for children who are reflective or who suffer from slow processing speed or poor motor coordination (Kaufman, 1992).
Timed tests exact a greater toll on the gifted because of the extent of the discrepancy between their competence and performance under timed conditions. To improve diagnostic accuracy, gifted children should be allowed to continue after the time limits, and both timed and untimed performance should be reported.

IQ tests may not tell the whole story. Achievement tests are needed as well. Most children attain lower scores on achievement measures than on measures of ability (Richard Woodcock, personal communication, July, 2000). However, as the Woodcock-Johnson III Tests of Achievement (WJ-III) have higher ceilings, some gifted children obtain higher scores on the WJ-III than on their IQ test. When achievement scores surpass ability measures, IQ scores are depressed and the achievement scores are better estimates of the child’s capabilities.

Test content also must be scrutinized. An appropriate IQ test for the gifted should be an excellent measure of abstract reasoning (Silverman, 2009b). Tests that emphasize working memory, processing speed, and nonmeaningful material are likely to produce less relevant results for this population than instruments designed to measure abstract reasoning or general intelligence (g).

**Interpreting Subtest Scores**

Asynchronous development is the norm for gifted populations (Silverman, 2012). Discrepancies among subtest scores are greater for the gifted than other groups (Rimm, Gilman & Silverman, 2008). The *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)* (American Psychiatric Association, 1994), which establishes the criteria used by mental health professionals for various diagnoses, provides clear admonition against averaging significantly discrepant subtest scores.

> When there is significant scatter in the subtest scores, the profile of strengths and weaknesses, rather than the mathematically derived full-scale IQ, will more accurately reflect the person’s learning abilities. (p. 40)

This advice appears in the DSM-IV section on mental retardation. The same caveat should be used with the gifted. When discrepancies among subtest scores exceed 8 points, or when Composite scores vary by 23 points, the child’s strengths and weaknesses should be discussed separately rather than averaged. The strengths should be used as the best indication of the child’s giftedness.

A different problem occurs when discrepancies typical for the gifted are misinterpreted as signs of abnormal brain functioning. Gifted children usually have higher Verbal Comprehension scores than Perceptual Reasoning scores because the verbal tests are better measures of cognitive ability and perceptual reasoning tests are more dependent on the child’s physical coordination and speed. Large discrepancies between Verbal Comprehension and Perceptual Reasoning have been misdiagnosed as evidence of a right hemispheric disorder (e.g., Nonverbal Learning Disorder).
Visual perception weaknesses need to be ruled out before diagnosing a child with Nonverbal Learning Disorder. Six months of vision therapy, faithfully practiced every day, have been found to increase scores in Perceptual Reasoning by as much as one or two standard deviations (Silverman, 2001, 2002). In the same vein, a Central Auditory Processing Battery is recommended to rule out auditory processing weaknesses before labeling a child AD/HD Inattentive Type (Silverman, 2002).

The General Ability Index (GAI) is usually the best estimate of a gifted child’s intelligence (Rimm, Gilman & Silverman, 2008). The GAI is derived from six core subtests of the WISC-IV, which compose the Verbal Comprehension Index (VCI) and the Perceptual Reasoning Index (PRI). It eliminates Working Memory and Processing Speed. However, when the discrepancy between VCI and PRI exceeds 22 points (1 ½ s.d.), either VCI or PRI is independently appropriate for selection to gifted programs (National Association for Gifted Children, 2008).

Some subtests are more relevant for the assessment of giftedness than others, and certain combinations of subtests indicate mathematical talent (e.g., Arithmetic and Block Design) or visual-spatial abilities (e.g., Block Design and Matrix Reasoning). Two optional WISC-IV subtests strongly correlated with general intelligence serve as excellent measures of giftedness: Information and Arithmetic. When the primary purpose of the assessment is to document giftedness rather than disabilities, Arithmetic can be substituted for Digit Span, and Information might be considered as a suitable substitution for Comprehension. (Two a priori substitutions are allowable.) Knowledge of factor loadings of the subtests enhances the examiner’s ability to pinpoint specific strengths and diagnose subtle weaknesses (Silverman, 2009b).

Extended Norms

Extended norms enable children to qualify for programs for the highly gifted and support acceleration. They are available for the WISC-IV (Zhu, Cayton, Weiss & Gabel, 2008) on the Pearson website: [http://pearsonassess.com/NR/rdonlyres/C1C19227-BC79-46D9-B43C-8E4A114F7E1F/0/WISCIV_TechReport_7.pdf]. Based on the same measurement metrics used in all Wechsler scales, they allow meaningful comparisons of strengths and weaknesses and illuminate the true extent of discrepancies. Extended norms take into account most of the raw score points earned by extremely gifted children, raising the possible scaled score from 19 to 28. Information, an optional test, appears to generate higher scores on the Extended Norms than some core subtests.

The Extended Norms table should be consulted when a child attains ceiling-range scaled scores (18 or 19) on two or more subtests. The norms can be calculated on prior WISC-IV test results: it is not necessary for the child to be retested.

Twice Exceptional Children

A remarkable number of gifted individuals suffer from hidden disabilities (Kennedy & Banks, with Grandin, 2011). These include stealth dyslexia (Eide & Eide, 2006), sensory
processing disorder, central auditory processing disorder, visual processing weaknesses, AD/HD and Asperger Syndrome (Lovecky, 2004). Hidden learning disabilities can be difficult to diagnose in children whose extraordinary abstract reasoning enables them to find other ways to solve problems. And disabilities can depress IQ scores so that a gifted child does not score in the gifted range (Rimm, Gilman & Silverman, 2008). A history of chronic ear infections, for example, exerts a greater impact on IQ scores in the gifted range than in the average range (Silverman, 1995b). An average child who misses many of the auditorally-presented items will still test in the average range. A gifted child who misses the same number of auditory items will also score in the average range—a decided loss.

Gifted children with learning disabilities demonstrate much more erratic IQ scores over time than other groups. They tend to do poorly on group tests, as well as on tests that are timed, require handwriting, or are administered later in the day when they are fatigued. Twice exceptional children often pass the harder items and miss the easy ones. This population is clearly at risk with short discontinue criteria. They should be offered items after they have met the discontinue criteria, and two sets of scores should be reported (Silverman, 2009a).

It takes a good detective to be able to ferret out disabilities in gifted children and recognize giftedness in disabled children. Family histories should be taken routinely to determine the degree of giftedness in the family and the presence of disabilities as well, since both have a strong hereditary component.

**Environmental Factors**

Environmental factors can affect test results for all children, but they influence the gifted to a greater degree. Among the factors that can prevent gifted children from demonstrating all that they know are:

- choosing to hide their abilities out of fear of the consequences of being labeled gifted (e.g., being removed from a current placement and being placed in a new environment; greater expectations of parents and teachers; losing friends; effect on siblings; etc.);
- unwillingness to guess for fear of making a mistake and appearing foolish;
- anxiety at being evaluated;
- feeling uncomfortable with the examiner;
- discomfort with aspects of the physical surroundings.

One child scored no higher than 3 on 10 subtests and 19 on the eleventh! Later, with another examiner, he obtained a score of 151, in the highly gifted range. His mother reported that he was uncomfortable in the first setting and refused to perform.
Rapport building is essential for all children and takes time. They can be asked to bring a favorite toy or a photograph album to share with the examiner (Meckstroth, 1989). If the child becomes afraid of making mistakes, a toy or a hand puppet can help answer the questions. The testing room should be free of distractions (e.g., noise, flickering bulbs, etc.). Frequent breaks should be allowed as needed and the child should be shown how to find the bathroom and his or her parent. If anxiety causes a child to freeze up, the examiner may move to a different section of the test and return to the anxiety-producing items when the child is more at ease, or postpone the rest of the exam for another day.

Some highly gifted children will not respond if a test item is too easy. They think it is a “trick question” and read many deeper meanings into it (Lovecky, 1994). Their IQ scores may be depressed because they know too much about a subject rather than too little. For example, examiner Melody Wood asked one child, “Who discovered America?” The girl thought for a long time and then said she didn’t know. When the test was over, Melody asked her the question again and she replied that she knew it wasn’t Christopher Columbus, because that theory had been disproven, but she couldn’t remember who it was!

Examiners should encourage guessing, as there are many gifted children (particularly girls) who will not offer an answer unless they are certain. Sometimes practicing simple guessing games like, “Guess what I ate for breakfast?” helps the child relax enough to speculate on the more difficult questions. Positive feedback for good guesses increases risk-taking and supports effort over performance.

How the evaluator feels about the child exerts a powerful effect on test scores. Some gifted children are extremely intuitive and pick up on facial expressions, body language, and other signals that the tester is unaware that he or she is emitting. If the examiner is hungry and is annoyed that the child is answering so many items correctly that the evaluation is taking longer than expected, the gifted child is likely to oblige by missing sufficient items so that the tester can go to lunch. On the other hand, if the evaluator thoroughly enjoys the workings of a gifted child’s mind and delights in every correct answer, the child responds to the twinkle in the examiner’s eye and tries his or her best.

**Conclusion**

With gifted children, there are many nuances in both testing and test interpretation. False positives are very unlikely. Scores in the gifted range do not occur “accidentally,” as one cannot fake abstract reasoning (Silverman, 1986). However, false negatives are abundant. Many more children are gifted than test in the gifted range. Underestimation of gifted children’s abilities, unfortunately, is much more common than accurate appraisal.

When the examiner knows enough about giftedness to recognize the pitfalls that may cause underestimates, he or she confirms test results with other data. If, for example, a child’s Broad Reading score is 160, but the IQ score is 125, the IQ score must be an underestimate. It is impossible for a child to achieve beyond his or her capabilities. (This is why the term “overachiever” is an oxymoron.) The highest indicator of a child’s
abilities at any age should be seen as the best estimate of the child’s giftedness (Silverman, 2009b). When other measures fall short of this indicator, the evaluator needs to explore carefully to determine possible causes of the underestimate.

The IQ scores of parents or siblings, early achievement of developmental milestones, profound curiosity, deep moral concern, remarkable associations or generalizations, perfectionism, advanced vocabulary, keen attention to detail, unusual empathy, vivid imagination, superb memory, early reading or fascination with Legos, school achievement, reading interests, and parental anecdotes of unusually precocious reasoning should all be taken seriously in determining the abilities of a child. A seasoned tester of the gifted uses this information to create a composite picture of the level of the child’s abilities. IQ test results, then, are nested into this schema to add further information.

Accurate diagnosis of the degree of a child’s developmental advancement is worth the investment. It should be based upon clinical judgment, rather than psychometric data. IQ scores are never an end in themselves; they are simply tools to be used wisely in the hands of professionals who understand giftedness.

REFERENCES


---

**Linda Kreger Silverman, Ph.D.,** a licensed psychologist, directs the Institute for the Study of Advanced Development (I.S.A.D.) and its subsidiary, the Gifted Development Center, in Denver, Colorado. She has studied the gifted since 1958, and authored 300 publications in the field, including *Upside-Down Brilliance: The Visual-Spatial Learner* and *Counseling the Gifted & Talented*. She works closely with test publishers on the development of intelligence scales that serve gifted children. Over 6,000 children from around the world have been assessed at the Gifted Development Center in the last 33 years.


---

**Gifted Development Center**  
* A service of I.S.A.D.*  
1820 Sheridan Boulevard  
Suite C-111  
Westminster, CO 80003  
www.gifteddevelopment.com  
1-888-GIFTED1