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

American Psychiatric Association. (1994*). Diagnostic and statistical manual of mental disorders* (Fourth edition). Washington, DC: Author.

Eide, B., & Eide, F. (2006). *The mislabeled child.* NY: Hyperion.

Kaufman, A. S. (1992). Evaluation of the WISC-III and WPPSI-R for gifted children. *Roeper Review, 14*, 154-158.

Kennedy, D. M., & Banks, R. S., with Grandin, T. (2011). *Bright not broken: Gifted kids, ADHD and autism.* San Francisco, CA: Jossey-Bass.

Lovecky, D. V. (1994). Exceptionally gifted children: Different minds. *Roeper Review, 17*, 116-120.

Lovecky, D. V. (2004). *Different minds: Gifted children with AD/HD, Asperger Syndrome, and other learning deficits.* London: Jessica Kingsley.

Meckstroth, E. (1989). On testing. *Understanding Our Gifted, 1*(5), 4.

National Association for Gifted Children (Jan. 2008). *Use of the WISC-IV for gifted identification.* Retrieved January 23, 2008, from http://www.nagc.org/index.aspx?id=375

Reams, R., Chamrad, D., & Robinson, N. (1990). The race is not necessarily to the swift: Validity of WISC-R bonus points for speed. *Gifted Child Quarterly, 34,* 108-110.

Rimm, S., Gilman, B. J., & Silverman, L. K. (2008). Non-traditional applications of traditional testing. In J. VanTassel-Baska (Ed.), *Alternative assessments with gifted and talented students* (pp. 175-202). Waco, TX: Prufrock Press.

Silverman, L. K. (1986). The IQ controversy—conceptions and misconceptions. *Roeper Review, 8,* 136-140.

Silverman, L. K. (1995a). Highly gifted children. In J. Genshaft, M. Bireley, & C. L. Hollinger (Eds.), *Serving gifted and talented students: A resource for school personnel* (pp. 217-240). Austin, TX: Pro-Ed.

Silverman, L. K. (1995b). Lost IQ points: The brighter the child, the greater the loss. In D.J. Lim, C.D. Bluestone, M. Casselbrant, J.O. Klein, & P.L. Ogra (Eds.), *Proceedings of the Sixth International Symposium on Recent Advances in Otitis Media* (pp. 342-346). Hamilton, Ontario: B.C. Decker.

Silverman, L. K. (2001). Diagnosing and treating visual perceptual issues in gifted children. *Journal of Optometric Vision Development, 32,* 153-176.

Silverman, L. K. (2002). *Upside-down brilliance: The Visual-spatial learner.* Denver: DeLeon.

Silverman, L. K. (2009a). Searching for asynchrony. A new perspective on twice-exceptional children. In B. MacFarlane & T. Stambaugh (Eds.). *Leading change in gifted education: The Festschrift of Dr. Joyce VanTassel-Baska* (pp. 169-181). Waco, TX: Prufrock Press.

Silverman, L. K. (2009b). The measurement of giftedness. In L. Shavinina (Ed.). *The international handbook on giftedness* (pp. 947-970). Amsterdam: Springer Science.

Silverman, L. K. (2012). Asynchronous development: A key to counseling the gifted. In T. L. Cross & J. R. Cross (Eds.), *Handbook for counselors serving students with gifs & talents* (pp. 261-279). Waco, TX: Prufrock Press.

Zhu, J., Cayton, T., Weiss, L., & Gabel, A. (2008). *Wechsler Intelligence Scale for Children–Fourth Edition: Technical report #7.* Upper Saddle River, NJ: Pearson Education.

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 32 years.

[An earlier version of this article appeared in *Gifted Education Press*, 1997, *12*(1), 2-5. © Linda Silverman.]

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