Gifted Development Center
a service of **The Institute for the Study of Advanced Development**
1452 Marion Street Denver, CO 80218 (303) 837-8378

Lost: One IQ Point Per Year for the Gifted

Linda Silverman, Ph.D.

Several factors combine to undermine our ability to detect the full incidence of highly gifted children in the population: (1) gifted children are not systematically assessed as are mentally retarded children; (2) it is common for gifted children to hide their abilities; (3) IQ tests are not constructed with sufficient difficulty to capture the full range of gifted children's abilities; (4) the tests provide no norms beyond the third or fourth standard deviation; (5) mental ages are no longer provided on current tests, which prevents extrapolating ratio IQs for the highly gifted; (6) scores in the highly gifted range are significantly depressed with each new set of test norms.

One of the authors of the Stanford-Binet Revision IV, Elizabeth Hagen, explained in an interview (Silverman, 1986) the constraints faced by test constructors. Their objective is to produce instruments that can adequately appraise the full range of individual differences in a chronological age group from the slowest level of development to the most rapid, but the instruments must be fairly easy to administer within a reasonable amount of time. The compromise is to produce instruments that are most effective for the **majority** of students. Tasks are eliminated if they are successfully completed by 99.99 percent of an age group or are failed by 99.99 percent of an age group. Items that can only be solved by children in classes for the gifted are purposely excluded because they would not be functional for a wide enough group. This practice produces a ceiling effect artificially depressing IQs in the upper ranges.

A more difficult problem to solve comes from the fact that methods of assessing giftedness are dependent upon the supposition that there is a normal distribution of intelligence in the population. If this supposition proves to be in error, the problem of assessing the gifted becomes exceedingly more complex.

The Theoretical Distribution of Giftedness

Most intelligence tests and other assessments provide standard scores, stanines, scaled scores and deviation scores - all derived from the Gaussian distribution. This is the theoretical basis for standardized testing as we know it today. The bell curve (using a 16 point standard deviation) predicts that IQ scores at the upper end of the scale are distributed among the general population in the following manner:

16 out of 100	above 116 IQ
11 out of 100	above 120 IQ
5 out of 100	above 125 IQ
2 out of 100	above 132 IQ
1 out of 100	above 136 IQ
6 out of 1,000	above 140 IQ
1 out of 1,000	above 148 IQ
1 out of 10,000	above 160 IQ
1 out of 100,000	above 168 IQ
1 out of 1,000,000	above 180 IQ

However, is this a true picture of the distribution of giftedness?

The Gap Between Theoretical and Real Distributions

Research is accumulating to suggest that the **actual** distribution of high capabilities may vary considerably from the theoretical distribution. There appear to be many more gifted children, especially highly gifted, in the population than anyone realizes. Several investigators have discovered unexpectedly frequent scores at the upper end of the IQ distribution that parallel the higher frequencies found at the lower end of the continuum (Dunlap, 1967; Gallagher & Moss, 1963; Jensen, 1980; McGuffog, Feiring, & Lewis, 1987; Robinson, 1981; Stott & Ball, 1965; Terman, 1925).

If intelligence were normally distributed at all levels, there would probably be less than six children who score above 168 IQ (1 in 100,000) in the entire school population of the state of Colorado (approximately 575,000). Yet, in the past ten years, we have found **over 115** children in this IQ range at the Gifted Child Development Center in Denver, and 23 of these score beyond 180 on the Stanford-Binet (Form L-M). We noticed an unusually large number of highly gifted children in our sample right from the start. Among the first 40 children brought to our center for assessment, one-third had IQs above 150 and one-sixth above 160. When we had 100 cases, 20 were above 150, 12 above 160 and 4 went beyond the limits of the test manual. Using the formula in the manual (p. 341), we computed those four to be in excess of 180. We have found similar distributions in subsequent years.

Puzzled by these figures, I inquired into the incidence of highly gifted children in programs around the country, and found that there were disproportionate numbers of highly gifted children in programs in New York, California, Washington, Illinois, Indiana and Alaska. The Talent Searches, which began at Johns Hopkins University under the direction of Dr. Julian Stanley, and have now involved over one million youth, have unearthed many times the number of gifted and highly gifted students that would be predicted (Stanley, personal communication, November 10, 1983).

Unusually high frequencies in the upper ranges were actually noted from the very beginning of IQ testing (Terman, 1925). Terman's first sample of 643 gifted children contained 98 above 150 IQ (15%) and 32 above 160 IQ (5%). Terman was a major proponent of the proposition that intelligence is distributed along a continuum resembling the normal curve. Yet, from the results of his study, he was forced to conclude the following:

> *The group contains an unexpectedly large proportion of cases in the upper IQ ranges. ...there is an appreciable excess of 150 IQ cases, or better, over and above the theoretical expectation. Above 160 IQ the number of cases found increases out of all proportion to the theoretically expected number and by IQ 170 exceeds it several times. Unless this discrepancy can be explained as due to the imperfection of the IQ technique it would appear that the distribution of intelligence in the child population departs considerably from that described by the normal probability curve. (Terman, 1925, p. 633)*

Other test constructors have also raised serious doubts about the assumption of the normal distribution of intelligence. In 1944, Wechsler wrote:

> *Some authors also believe that the resulting frequency curve (of scores on intelligence measures) ought to be Gaussian or as nearly Gaussian as possible. The last requirement seems to be a result of the wide-spread but mistaken belief that measures distribute themselves according to the normal curve of error. (p. 126)*

In large-scale studies in several countries, J. C. Raven (1959) and J. Raven (1983) found distributions with the Progressive Matrices Intelligence Test which sharply depart from the Gaussian curve.

> *Finally, in relation to Poulsen's observations about the 'peculiar' shape of the distribution of test scores, it is important to note that J. C. Raven (1959) used the data he obtained with his own (and others') tests to question the notion that test scores 'should' be 'normally' distributed. He again argued that it might be the assumption about what reality should be like, and not the obtained distributions of test scores, which was wrong. For this reason he resisted the use of deviation IQs as the means of expressing test scores. (J. Raven, 1983, p. 174)*

The frequency of scores at the lower end of the IQ continuum clearly departs from the normal curve. The excess of frequencies in one tail of the curve is observed in the distribution of height as well as intelligence, and has shown up in practically every large study of either trait (Jensen, 1980). According to Jensen, the normal curve theoretically and empirically fits the distribution of scores for 98 percent of the population, from 60 to 150 IQ. However, it does not apply at the extremes.

> *Within the IQ range from about 60 to 150, the IQ distribution does not depart significantly from the normal curve. But there is an excess of high IQs beyond about 150. Like the excess of very low IQs, it is a genuine phenomenon, though less pronounced. It is generally not statistically detectable in samples numbering fewer than two or three thousand. The causes of this excess of IQs above 150 are not certain, but two factors most likely account for it: (1) a high degree of assortative mating (like marrying like) among highly intelligent parents and (2) covariance of variable environmental and genetic influences. (p. 84)*

How IQ points Get Lost

The lack of symmetry at the extremes of the normal curve has significant implications for our interpretation of IQ scores of the gifted. Since the tests are constructed with the assumption of the normal distribution of scores, with few items of sufficient difficulty for the gifted, a problem results which has not been adequately understood or resolved. The larger portion of gifted students actually found in the population still has to fit within the normal curve of distribution in order for standard scores (deviation IQs) to apply. This compresses the range of abilities that can be expressed, lowering IQ scores in the gifted range. The more unusual the child's ability, the greater the impact of this compression. The problem appears to intensify with each new set of published norms.

According to a large-scale analysis of studies of intelligence, Americans are gaining approximately .3 IQ points in intelligence each year (Flynn, 1984). For this reason, newer, more stringent test norms are usually considered more accurate than old norms. The way this translates into test interpretation is as follows. In 1960, when a child of 8 years 0 months achieved a mental age of 8.0 on the Stanford-Binet, the resulting IQ score was 100. In 1972, the same raw score yielded an IQ score of around 96 (95 at this particular age). In 1986, if the same test could have been used with a new set of norms, the raw score would have resulted in an IQ of around 92.

But a rather curious situation occurs when we examine the scores of gifted students on these various sets of norms. In 1960, a five-year-old achieving a mental age of 8.0 would have had an IQ score of 165. In 1972, that same raw score only yielded an IQ of 153, a difference of 12 points. Differences between the Stanford-Binet Revision IV, published in 1986, and the 1972 norms appear to be at least 13.5 points in the moderately gifted range (Thorndike, Hagen & Sattler, 1986), which would bring the same child's score down below 140. This is a loss of one IQ point per year from 1960 to 1986 for children in the gifted range. In this 26 year period, average students needed to obtain only **8** more points to make up for the average gains in intelligence of the general population, whereas gifted children needed to obtain **over 25** more points to match previous scores **- 1 1/2 standard deviations of IQ**. This seems like an unreasonable demand.

In the latest edition of the Stanford-Binet, two samples of gifted students are presented in the technical manual: one sample (N = 82) received the L-M edition of the test (1972) and the 4th edition (1986); the other (N = 19) received the 4th edition of the Stanford-Binet and the Wechsler Intelligence Scale for Children-Revised (WISC-R). In the first sample, the mean score on Form L-M was 135.3, whereas the composite score for the same group on the 4th edition was **121.8** (Thorndike, Hagen & Sattler, 1986, p. 70). In the second sample, the 4th edition yielded a mean composite score of **116.3** compared to a mean WISC-R Full Scale score of 117.7 (Thorndike, Hagen & Sattler, 1986, p. 72). There is no appreciable difference between WISC-R and 4th edition S-B scores at 1 standard deviation above the mean, but obviously children who scored 2 standard deviations above the mean (132+) on previous editions of the Stanford-Binet are now scoring in the 120 range.

Implications

Which set of norms is the most "accurate" for the gifted? Most psychologists will answer, "the most recent set." However, for the highly gifted child, this leads to a rather absurd situation. When too many children are found in the upper ranges, the scores are adjusted to fit the theoretical curve. This swells the number of scores in the 120-130 range and depresses the IQ scores of the entire gifted population. The attempt to artificially force the distribution of giftedness into the normal curve results in the disappearance of 1 1/2 standard deviations of intelligence. With today's measuring devices, all IQ scores in the gifted range are most likely **underestimates** of ability.

Broader Selection Criteria

This information is not commonly known by psychologists, educators, or even test constructors, and the newer norms can be used punitively against children who truly qualify as gifted but are dealing with a stacked deck. Fewer children qualify for placement in gifted programs using the Stanford-Binet IV than using the Stanford-Binet (Form L-M) (Kitano & DeLeon, 1988).

The situation differentially affects children in public and private schools. In public schools, where funds are limited and there are fewer places available than qualified students, only the most recent test scores are perceived as credible. The objective is to limit the number of students served. If a 130 or 132 cut-off score on 1986 norms is used as a basis for selection, the program may actually serve only the highly gifted - children who might have scored 155 or higher on previous sets of norms.

In private schools for the gifted, the child's ability to “fit in with the others” is often the major criterion for acceptance, and the norms are not seen as particularly relevant. Many schools accept children on the basis of 1960 norms (albeit unofficially) because they feel that these norms do identify children who perform like other gifted children (Silverman, Chitwood & Waters, 1986). Even highly selective schools that have many more applicants than available places usually do not rely on recent norms as a selection criterion.

In light of the problems in assessment, broader conceptualizations of giftedness, which include children in the 120 to 130 IQ range, seem more appropriate than strict adherence to the 130 or 132 cut-off score traditionally employed for gifted programs. Young children who score in the 120s, the superior range of the IQ test, probably would have attained scores in the gifted range on prior sets of norms, as indicated above. It is worth noting that the two “gifted” samples described in the Stanford-Binet 4th edition manual had mean composite scores of **116** of 121 (Thorndike, Hagen & Sattler, 1986, p. 72). Therefore, if the test constructors consider 116 high enough to be called a “gifted” sample, the schools would be hard-pressed to only consider 130 the official gifted range.

Recommendations for the Highly Gifted

Until intelligence tests are developed specifically for the gifted, we may never know the true incidence of the highly gifted. Scores near the top of the range may actually be considerably higher. This has been demonstrated through the Talent Searches, in which seventh and eighth graders who achieve above the 95th percentile in their achievement tests are given the opportunity to take the Scholastic Aptitude Test (SAT), a much more difficult test designed for high school seniors. Students with similar scores on grade level achievement tests may score anywhere from 200 to 800 on the SAT. Scores of the top five percent of junior high school students form a normal curve of distribution on the SAT (Cohn, 1983).

Talent Searches provide a good indication of the abilities of many students, but the SAT is as much a measure of achievement as intelligence, so it still misses a substantial number of gifted students. However, it does show the importance of developing instruments with sufficient levels of difficulty to disclose the full extent of students' abilities. Without these tools, the cluster of highly gifted children at the upper end of the range may never be detected.

Since none of the current instruments properly assess the full strength of the abilities of exceptionally gifted children, alternative methods of evaluation must be sought. In the *National Report on Identification* (Richert, 1982), the consensus of the national task force was as follows:

> *it was recognized that new instruments and methods need to be developed for identification of gifted students in specific populations, such as disadvantaged, ethnic minorities, students with limited English-speaking ability, exceptionally gifted students and handicapped students. (pp. 77-78, italics added)*

This recommendation is already being implemented with all of the specific populations listed except one - the exceptionally gifted. In this case, an **old** method of identification is more appropriate than newer methods. It is recommended that when a child obtains **three** subtest scores at or near the ceiling of any current instrument, he or she should be retested on the Stanford-Binet Form L-M (Silverman & Kearney, 1989). Ratio IQ scores should be computed for all children who score beyond the norms in the manual (using the formula on p. 341 of the Stanford-Binet Form L-M manual), and perhaps for those who score above 155 on the 1972 norms, because at some age levels the highest score in the norm table is 158 (Terman & Merrill, 1973).

The Stanford-Binet (Form L-M) is of greatest value for children ages 3 through 11. It is advised that older students take the Scholastic Aptitude Test for a better estimate of their advanced abilities. The SAT is available through the national Talent Searches to highly capable students as early as the 5th or 6th grade. This type of above-level testing is the only method we currently have of assessing the full range of highly gifted children's abilities.

Conclusion

We now know that there are more gifted children than we originally thought there were, and that many of these children would test much further from the norm than is now possible if we had appropriate assessments of high levels of ability. In other words, many gifted children are a great deal brighter than their test scores would indicate. For the highly gifted, where discrepancies may be as large as 50 points from one test to another (Silverman & Kearney, 1989), it is imperative that supplemental testing (i.e., the Stanford-Binet, Form L-M) or above-level testing (e.g., the SAT for ages 10 through 13) be employed for more accurate estimates of their abilities.

***Note: Presented at the 36th annual convention of the National Association for Gifted Children, Cincinnati, Ohio, November 2, 1989.**

REFERENCES

Cohn, S. (1983, Feb.). Two-step identification process. *Second annual training conference: Providing appropriate educational services for intellectually gifted students*, (p. A93). Tempe, AZ: Arizona State University.

Dunlap, J. M. (1967). The education of children with high mental ability. In W. M. Cruickshank & G. O. Johnson (Eds.), *Education of exceptional children and youth* (2nd ed.) (pp. 143-193). Englewood Cliffs, NJ: Prentice-Hall.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95, 29-51*.

Gallagher, J., & Moss, J. (1963). New concepts of intelligence and their effect on exceptional children. *Exceptional Children*, 30(1), 1-5.

Jensen, A. R. (1980). *Bias in mental testing.* New York: The Free Press.

Kitano, M., & DeLeon, J. (1988). Use of the Stanford-Binet Fourth Edition in identifying young gifted children. *Roeper Review, 10*(3), 156-159.

McGuffog, C., Feiring, C., & Lewis, M. (1987). The diverse profile of the extremely gifted child. *Roeper Review, 10*(2), 82-88.

Raven, J. C. (1959). A note on Burt's “The distribution of intelligence.” *British Journal of Psychology, 10,* 95-100.

Raven, J. (1983). Raven Coloured Progressive Matrices Intelligence Test in Thailand and in Denmark: A response. *School Psychology International, 4,* 173-176.

Richert, E. S., with Alvino, J. J., & McDonnel, R. C. (1982). *National report on identification: Assessment and recommendations for comprehensive identification of gifted and talented youth.* Sewell. NJ: Educational Improvement Center-South.

Robinson, H. B. (1981). The uncommonly bright child. In M. Lewis & L. A. Rosenblum (Eds.), *The uncommon child* (pp. 57-81). New York: Plenum Press.

Silverman, L. K. (1986). An interview with Elizabeth Hagen: Giftedness, intelligence and the new Stanford-Binet. *Roeper Review, 8*, 168-171.

Silverman, L. K., Chitwood, D. G., & Waters, J. L. (1986). Young gifted children: Can parents identify giftedness? *Topics in Early Childhood Special Education, 6*(1), 23-38.

Silverman, L. K., & Kearney, K. Parents of the extraordinarily gifted. *Advanced Development, 1*, 41-56.

Stanley, J. (1983, Nov. 10). Personal communication.

Stott, L., & Ball, R. (1965). Infant and preschool mental tests: Review and evaluation. *Monograph of the Society for Research in Child Development, 30*, 4-42.

Terman, L. M. (1925). *Genetic studies of genius: Vol. 1. Mental and physical traits of a thousand gifted children.* Stanford, CA: Stanford University Press.

Terman, L. M., & Merrill, M. A. (1973). *Stanford-Binet Intelligence Scale: 1972 norms edition*. Boston: Houghton Mifflin.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *The Stanford-Binet Intelligence Scale: Fourth edition*. Technical manual. New York: The Riverside Publishing Co.

Wechsler, D. (1944). *Measurement of adult intelligence* (3rd ed.). Baltimore: Williams & Wilkins.

Linda Silverman, Ph.D., is a licensed psychologist
and Director of the *Gifted Development Center* in Denver, Colorado.