HOW MISLEADING THE FIFTY YEARS? A REPLY TO JACOBS

William H. Graves

After reading "Are We Being Misled by Fifty Years of Research on Our Gifted Children?" (Jacobs, 1970), it became apparent that one of the more misleading events of the past fifty years was the research design of this study. Jacobs reported administering individual IQ tests to incoming kindergarteners in a school district, selecting as gifted those nineteen who scored at or over 125, and randomly selecting nineteen others to serve as a control group. After one year, the gifted and control groups were retested and the gifted group was found to have suffered a statistically significant 7.7 point drop in IQ, a change not mirrored in the control group.[1] One conclusion drawn was that there was a negative growth trend for these gifted children.

When a group is identified on the basis of directionally deviant scores on a test of less than perfect reliability (for simplicity, high IQ scores will be referred to hereafter), subsequent administrations of the test to that group yields a group mean closer to the overall mean. This is known as statistical regression, and it accounts for Jacobs' findings more parsimoniously than do his conclusions. This phenomenon is discussed at length by Goodenough and Maurer (1940), Thorndike (1942), and Campbell and Stanley (1963), but a brief explanation is offered.

The score a person obtains on an IQ test, group or individual, is not his "true" IQ, but only an approximation thereof. An important consideration is the likely scoring range of an individual. An index of this variability is the standard error of measurement, which is about five for the WPPSI. If a child were to take the test numerous times, he would tend to score within five points (high or low) of his true score about two-thirds of the time. A true score shall be (imprecisely) defined as the average of multiple trials.

We may accept, for the moment, the criterion of giftedness cited by Jacobs of an IQ at or over 125. Administration of the WPPSI to a sizeable number of children would yield some who, on that trial, scored lower than 125, but whose true score was at or over 125. These children would be unjustly denied a gifted classification, and shall be called excluded gifted (EG). In the same administration would be some children who scored 125 or over, but whose true score was lower than 125. These children would be falsely classified as gifted, and shall be called false gifted (FG). Upon

[1] It was reported that, in fact, *no* change occurred in the control group on any test or subtest. This was interpreted to mean that no *significant* change occurred.

Dr. Graves is Assistant Professor of Education, The University of Oklahoma, Norman, Oklahoma 73069.

Downloaded from gcq.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/25601

reexamination at the end of the school year (or the next day, for that matter), these children would tend to score around their true score rather than around their score on the first test. If, in a massive testing program, we find 100 children with a true IQ of 120, we would expect to find about thirty with a true IQ of 130, since the WPPSI follows a normal (bell-shaped) distribution with a standard deviation of 15. Of the 100 children of true IQ 120, the standard error would lead us to predict that about sixteen would score at or over 125 on any one administration, and will be false gifteds (FGs). At the same time, about five of the thirty children of true IQ 130 would probably score below 125, the excluded gifteds (EG).[2] The FG group will average about 120 on a second administration and the EG group will not have the opportunity to average their 130. The FGs will therefore lower the mean on the readministration directly by their lower scores, and the EG group's absence will keep the mean of the second administration reduced.

As we get closer to the IQ 125 cutting line, the percentage of erroneous classification increases (of the true IQ 124 children, more than 40% will be FGs, and at true IQ 125, more than 40% will be EGs). As we get closer to the cutting score, the differential between the group means on the first and second tests will become smaller, but there will be more misclassified people, and the effect will still be towards reducing the overall mean on the second test, and it will be a cumulative effect.

Goodenough and Maurer (1940), in their critique of the Iowa studies, disclosed the results of studies at the University of Minnesota on children of nursery school age. Thirteen children were identified who scored 130 or over on the first administration of the Kuhlmann-Binet. A year later, readministration indicated that this group had a mean score 7.98 points lower (and low IQ groups had substantial increases), which was attributed to statistical regression. It should be emphasized that Jacobs' "gifted" group was a high intelligence group, even considering the second administration. Research employing independent measures like the Rorschach (Jacobs, 1971) is not contraindicated by the considerations which invalidate Jacobs' 1970 study.

Even if Jacobs' results could not be attributed to statistical artifact, he mentioned a negative growth trend for the gifted children in his study which was in conflict with the preponderance of studies which indicated positive growth trends (Jacobs cites Terman, 1925-52, and alludes to others). It should be emphasized that in many of the conflicting studies, these positive growth trends were in logically independent areas, such as adjustment social effectiveness, whereas in Jacobs' study the negative growth was in the dependent area of IQ.

Jacobs referred to the potential bias in Terman's procedure of pre-screening children through group IQ tests and teacher nominations. While it is agreed that this almost certainly introduces bias in several

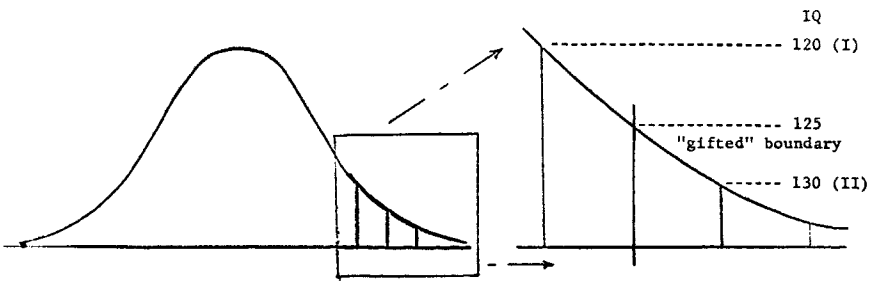
[2] It is acknowledged that there will be sixteen children of true IQ 120 who will score below 115 and five children of true IQ 130 who will score above 135. The effects of these will be analyzed in the figure.

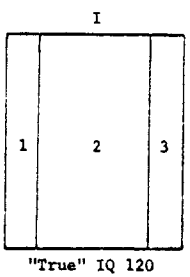
Downloaded from gcq.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

dimensions, it is worthwhile to explore the likely sources and consequences of such bias. When considering which children "possessed qualities which allowed teacher to recognize them as gifted" (Jacobs, p. 122), the aspects of interest and motivation as perceived by the teacher would seem to weigh heavily, so those nominated would likely be high achievers. The biasing influence of group administered IQ tests is more difficult to isolate, but would likely tend to exclude those who are less socialized, and might be construed to be detrimental to the creative child, due to considerations raised by Torrance (1963) and Wallach and Kogan (1965).

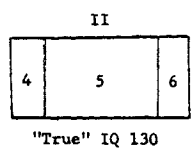
FIGURE 1

REPRESENTATIVE RESCORING PATTERNS WITHIN A GROUP SELECTED FOR INITIAL HIGH IQ SCORES





Figures I & II represent IQ scores of children on any given administration of an IQ test when the true IQ of those children is 120 (I) and 130 (II).



Group No.	A Group True Score	B Approximate Group Mean, Test 1	C Group Test 1	D Approximate Group Mean, Test 2	E True Gifted?	F Classified as Gifted?	G Relative No. of People	H Effect on Second Test Mean, Gx(D-B)
1	120	113	114-down	(120)	no	no	16	(112)
2	120	120	115-124	(120)	no	no	68	(0)
3	120	127	125-up	120	no	yes	16	-112
4	130	123	124-down	(130)	yes	no	5	(35)
5	130	130	125-134	130	yes	yes	20	0
6	130	137	134-up	130	yes	yes	5	-35
							Net loss:	-147

Downloaded from gcq.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

The later consideration is indeed serious, but the case for teacher nomination bears further investigation. From the standpoint of research exploring the correlates of high intelligence, any biasing factors are lamentable. From the practitioner's viewpoint, however, and for those who prefer to define "gifted" in a broader sense than a cutting score on a single test (Torrance, 1970), the limitations of single measures outweigh their administrative ease and simplicity. Terman (1916, p. 67) alluded to the unreliability of single measures as he quoted Binet: "let the tests be rough, if there only be enough of them." Increased reliability and validity in assignment means more fairness and more efficient use and development of resources. Lord (1963) has reported a procedure for determining cutting boundaries of multiple correlated measures. Those children who pass the triple test of interest (teacher nomination), potential (IQ) and desire (pupil and parental volunteering) would seem to be better candidates for intensive enrichment programs which might "turn off" those of similar potential, but who lack interest. The problem of "turning on" underachievers is recognized as an important area, somewhat more complex due to the moral question of unsolicited intervention and to the multiplicity of reasons for non-interest. It would seem that programs for high intelligence underachievers should be separate from (but probably leading to) programs for the gifted.

REFERENCES

Campbell, D. T., and Stanley, J. C. *Experimental and quasi-experimental designs for research.* Chicago: Rand-McNally, 1963.

Goodenough, F. L., and Maurer, K. M. The relative potency of the nursery school and the statistical laboratory in boosting the IQ. *Journal of Educational Psychology,* 1940, *31*, 541-549.

Jacobs, J. C. Are we being misled by fifty years of research on our gifted children? *Gifted Child Quarterly,* 1970, *14,* 120-123.

Jacobs, J. C. Possible misinterpretation of personality traits of the gifted. *Gifted Child Quarterly,* 1971, *15,* 195-200.

Lord, F. M. Cutting scores and errors of measurement. *Psychometrika,* 1962, *27,* 19-30.

Terman, L. M. *The measurement of intelligence.* Boston: Houghton-Mifflin, 1916.

Terman, L. M. *Genetic studies of genius.* 5 vols. Stanford: Stanford University Press, 1925-1952.

Thorndike, R. L. Regression fallacies in the matched groups experiment. *Psychometrika,* 1942, *7,* 85-102.

Torrance, E. P. *Education and the creative potential.* Minneapolis: University of Minnesota Press, 1963.

Torrance, E. P. Broadening concepts of giftedness in the 70's. *Gifted Child Quarterly,* 1970, *14,* 199-208.

Wallach, M. A., and Kogan, N. *Modes of thinking in young children.* New York: Holt, Rinehart and Winston, 1965.

Downloaded from gcq.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016