

Interpreting IQ Scores

John R. Hills
Florida State University

Notice This Material May Be Protected
By Copyright Law (Title 17 U.S. Code)

Tests that provide IQ scores are widely used in schools. A score that tells how "bright" a student is, or how readily he or she will learn in school, would be a handy bit of information, and it might be essential as a trustworthy and reliable guide for diagnosing learning difficulties, for providing students with special educational opportunities, or for predicting future educational performance. But can IQ scores do all these things for us? Try the following questions about scores from intelligence tests. Indicate whether the stated interpretation of the IQ score is sound, true or false. The correct answers and brief explanations appear on page 38. This "test" may be reproduced without written permission. Please cite *Educational Measurement: Issues and Practice* as the source.

Hills' Handy Hints: IQs

- T F 1. Manny got an IQ score of 115. This score means that Manny's innate ability to do things that require brains or intelligence is about one standard deviation above the mean or at about the 84th percentile.
- T F 2. Manny's IQ score came from a recent revision of a widely used intelligence test with a long and honored history. One can be confident that the score was derived by finding Manny's Mental Age on the basis of his answers to the test questions, dividing that by his Chronological Age, and multiplying by 100, as in $(MA/CA) \times 100 = IQ$.
- T F 3. If Manny had been given a different IQ test, one that was used as widely and with an equally long and illustrious history, the resulting IQ score would be about the same level (within 2 or 3 points either way).
- T F 4. If Manny received the IQ score of 115 when he was in the sixth grade, we know that he was old enough that the score will remain stable, neither increase nor decrease appreciably, for many years.
- T F 5. Manny happens to be black. Knowing this, we can estimate that Manny's score is lower than it should be because IQ tests are biased against minority groups.
- T F 6. The test that Manny took had two scores, one for Verbal skills and one for Performance or Nonverbal skills. The difference between the Verbal and Nonverbal IQs was 15 points. A skilled clinician can use that information by itself to make a useful diagnosis of emotional disturbance or organic brain injury.
- T F 7. Manny's brother, Sherman, took the same IQ test and got a score of 102. Because his score is below 110, there is little chance that he could be admitted to a college.
- T F 8. The IQ test that Manny and Sherman took was an established and highly respected test, but the publisher still may not have checked to see whether the IQ scores are useful for predicting performance in school and may not have published relevant results in the test manual.
- T F 9. Although Manny's teacher needs to know Manny's score, and it should be posted in the school records, Manny should not be allowed to know what his IQ score is.
- T F 10. If someone had taken Manny aside and showed him how to do some of the kinds of items on the IQ test and let him practice on those kinds of items a bit, his score might have been noticeably higher.

John Hills is Professor of Educational Measurement in the Department of Educational Research, Development and Foundations of the College of Education at Florida State University. He has held that position since 1966. Prior to coming to Florida State University, he was in the Research Division of Educational Testing Service in Princeton, New Jersey, and later Director of Testing and Guidance for the Board of Regents of the University System of Georgia. He is a member of the Board of Directors of NCME, has been an Advisory Editor for the Journal of Educational Measurement, the Personnel and Guidance Journal, and Research in Higher Education. He was an Examiner in Aptitude Testing for the College Entrance Examination Board for several years, and was Chairman of that group in 1971-72. His introductory textbook on measurement and evaluation for teachers is widely used and has been translated into Spanish.



Hills' Handy Hints: IQs

continued from page 34

Questions 1-5, 7, and 9 are false. Questions 8 and 10 are true. The answer to Question 6 is true or false depending on whose opinion you believe, or how you evaluate empirical data, conjecture, opinion, and hope. Explanations:

1. IQ tests do not measure innate abilities unmodified by environment. It is even questioned whether tests given at birth could be argued to be measures of innate abilities because the fetus is influenced by its environment in many ways.

2. Nearly all IQ scores from current intelligence tests use a deviation IQ based on standard scores with a mean of 100 and a standard deviation of 15 or 16 instead of the ratio IQ score based on MA/CA.

3. Different IQ tests can result in quite different scores for the same person because different IQ tests contain different kinds of items or different relative emphases. Different IQ tests are based on different norms groups, which can also result in different IQs for the same person.

4. IQ scores may change, and sometimes markedly. Cronbach uses as an illustration a person whose IQ changed by 55 points with changes in environment (*Essentials of Psychological Testing*, Third Edition, Harper and Row, p. 232). Other evidence also points to the strong possibility of significant changes in IQ score with changes in the environment or, in the case of older forms of IQ tests, changes in age of the examinee (Ahmann & Glock, *Evaluating Pupil Growth*, Fourth Edition, Allyn and Bacon, pp. 380-390).

5. Although it is true that blacks on the average score 10 to 15 points lower than whites on IQ tests if the content of the items stresses verbal reasoning, there is no clear evidence that the difference is due to bias in the test or its items. Manny's score of 115 is about a standard deviation above average, or at about the 84th percentile. If he had scored 15 points higher, he would have

placed about two standard deviations above average, or at about the 98th percentile. Most likely Manny's performance will correspond to his score of 115, not a higher score. It would be a mistake to predict greater success for Manny than would be associated with his score of 115 or to expect his performance in activities depending on verbal reasoning to be comparable to those of students at the 98th percentile instead of the 84th percentile. The same kind of interpretation would apply if Manny had scored a standard deviation below the mean, or at IQ 85. We would then recognize that he has more difficulty with verbal reasoning items than average, and most likely he would then have more than average difficulty with the verbal materials he encountered in school unless something occurred to remedy this weakness.

6. Research support for using IQ scores or part scores on intelligence tests for diagnosis of emotional disorders or brain injury is sketchy at best. For example, read Cronbach, op. cit., pages 248-251, or Tittle's review of WISC-R in the *Eighth Mental Measurements Yearbook*, page 353. However, the manual for WISC-R indicates that it is useful for psychological diagnosis. So, you could reasonably argue that based on what you have read you deserve credit for either T or F.

7. The variety in colleges and their requirements is surprising to those who have not developed a familiarity with the literature on college admissions. Cronbach (op. cit., p. 219) is a handy reference indicating, for example, that among college entrants studied by Wolfe in the 1950s, 18 percent had IQ scores between 100 and 109. Very few colleges are highly selective; most of them reject only students who are below average, that is, below IQ 100, and most colleges admit some of the applicants who are below average.

8. One cannot count on finding predictive validity data in the test manual for IQ tests. A specific example is the WISC-R.

Tittle (*Eighth Mental Measurements Yearbook*, p. 353) states that the major weakness of the WISC-R manual is the discussion of validity. Sax (*Principles of Educational and Psychological Measurement and Evaluation*, Second Edition, Wadsworth, pp. 389-391) points out that neither the WISC manual or the WISC-R manual even mentions the word "validity," and he goes on to indicate that the manuals for group intelligence tests similarly fail to provide information about validity (p. 392).

9. One of the mystiques about IQ scores is that the individual should not know what his IQ score is. Any other score is all right, but not the IQ score. Prior to the 1960s, the same mystique applied to College Board SAT scores. I doubt that any educational measurement book espouses such secrecy about any test score. Mehrens and Lehmann (*Measurement and Evaluation in Education and Psychology*, pp. 608-609) firmly state that all achievement and aptitude scores should be disseminated to all professional staff members and to the individuals who were tested, even though they note that according to a study by Goslin, half of teachers surveyed had never given a pupil even a general idea of his intelligence. Remember, IQ scores from modern tests are no more than standard scores on a measure of generalized achievement of cognitive skills and knowledge. They have no special, secret, or magical properties except in the minds of the uninformed.

10. Coaching students for tests of college admission has become an important topic in recent years. Debate continues concerning the effectiveness of coaching for those tests. However, it is often erroneously assumed that IQ tests measure something innate or so fundamental that coaching is ineffective for items on these tests. Some IQ tests include unusual kinds of items for which coaching is beneficial. In fact, some investigators have suggested and demonstrated that IQ scores might be more valid if all children were coached on such items (Cronbach, op. cit., page 245). ■