Source: Educational Measurement: Issues and Practice

Interpreting Grade-Equivalent Scores

John R. Hills
The Florida State University

School teachers use grade-equivalent (GE) scores from commercially published standardized tests because they feel that a grade equivalent is meaningful. Unlike scores such as percentiles, stanines, or scaled scores, GE scores tell you directly what they mean—or do they? Try the following examples. Just indicate whether the stated interpretation of the GE score is sound, true or false. The correct answers and brief explanations appear on page 21. This "test" may be reproduced without written permission. Please cite *Educational Measurement: Issues and Practice* as the source.

This is the first of a series of quizzes by John Hills that will be printed in EM. Subsequent quizzes will cover other types of derived scores. Your reactions and suggestions for other topics may be sent to the editor or directly to Professor Hills.

Hills' Handy Hints: Grade Equivalents

Please circle either T or F for each question. Mark a response for each question as there is no penalty for guessing. This is an untimed test.

T F 1. Tim is a sixth grader. He obtained a GE score of 9.2 in reading. This means that Tim scored well above average sixth graders on reading.

T F 2. A GE score for Tim of 9.2 means that he can read as well as ninth graders in the second month of the school year.

T F 3. Tim's GE score of 9.2 on reading means that when a group of ninth graders in their second month were tested on ninth grade reading material, they received scores equivalent to Tim's score.

T F 4. Tim's GE score of 9.2 on reading means that Tim could well be put in a class of ninth graders for material in which reading skills were important.

T F 5. Tim's 9.2 GE in reading means that in a flexible school in which children work on materials at their own level, Tim should be put into a ninth grade class for instruction in reading.

T F 6. Tim obtained a GE score of 7.3 in arithmetic on the same test battery from which his reading GE score was 9.2. This means that in reading Tim is nearly 2 years ahead of his performance in arithmetic.

T F 7. GE scores of 9.2 in reading and 7.3 in arithmetic indicate that Tim is farther ahead of his own class in reading than in arithmetic.

T F 8. Tim's GE of 9.2 in reading was from fall testing in the sixth grade. Tested in the spring, he received a GE score of 8.0. That indicates that his reading skills declined during that school year. Some effort should be expended to find out why and whether such losses can be expected to continue.

T F 9. When tested in September, 30 percent of the students in Mr. Brown's fifth grade class got GE scores below 5.1. Something needs to be done to help his students reach grade level. Also, the third and fourth grade teachers should improve the instruction given to students before they reach Mr. Brown.

T F 10. Jones Elementary School is an inner-city school. The Jones school mean GE score in reading in first grade was .6. The mean increased each year until by the sixth grade it was up to 3.2. Thus, the Jones mean was .4 years behind at the first grade and nearly 3 years behind by the sixth grade. Because the Jones students are falling farther behind the national average each year, the reading program, the teachers, and the administration are inadequate to meet the learning needs of the Jones school students. ■

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: Grade Equivalents

Question 1 is true. The others are false.

Explanations:

1. Because GE scores are developed by obtaining the mean or median performance at each of several grade levels on a test whose content covers the several grades, a student who scores above the grade in which he is enrolled has performed above average for students in his grade.

2. A student can obtain a high GE score without being able to do the work of students at the grade level indicated by his score. Tim may have obtained a 9.2 score by getting *all* the items that were designed for grades 4, 5, 6, and 7 correct and may not have done particularly well on items designed for grades 8 and 9, if there were any.

3. Often the GEs associated with high or low number-correct scores are obtained by extrapolation. It is possible that no ninth grader was ever tested with the test given Tim.

4. Because Tim could have gotten a GE score of 9.2 by doing well on the easier or lower level items of the test, one cannot tell from these scores whether he could participate effectively with ninth graders or not.

5. In most schools, reading is not taught in the ninth grade except perhaps for remediation of ineffective reading skills. So it does not make sense to consider putting Tim into ninth grade instruction in reading. Even if reading (or any subject for which a 9.2 GE was obtained) were taught at the ninth grade, one would not know whether Tim should be put into a higher level of instruction without evaluating whether he had the prerequisite skills. The GE score cannot be relied on to indicate that ninth grade skills have been mastered.

6. The standard deviations of GE scores vary from one subject to another. Tim's score of 9.2 on reading and 7.3 on arithmetic could be equal scores if one used another score scale such as percentiles. The difference between the two GE scores may be due to the fact that students tend to differ less within a grade on arithmetic than on reading. In addition, GE scores above a student's grade do not mean that he has really mastered skills beyond his own grade level.

7. Because the standard deviations for different subjects differ, we cannot tell whether 9.2 in reading is relatively better than 7.3 in arithmetic, and neither necessarily implies that Tim is ahead of his own class.

8. When GE scores have been extrapolated far above or below a student's grade level, it often occurs that even a single additional item correct can change a student's GE score by more than a year. Tim may simply have gotten one or two fewer items correct in spring than fall.

9. The GE score is based on a mean. One dare not expect all students to be at or above grade level on GE scores. In a typical heterogeneous class, about half would obtain GE scores below grade level and half above grade level. Because only 30 percent of Mr. Brown's fifth graders got GE scores below 5.0, his students may be doing a little better than usual instead of worse.

10. Another peculiar characteristic of GE scores is that the standard deviations get larger year by year. Suppose that a person (or a group average) is at and remains at a given percentile score, say the 16th percentile (which is one standard deviation below the mean). This same percentile each year is translated into a lower GE each year because the standard deviation gets larger from year to year. This can leave the impression that a person (or group) is falling farther behind each year. Similarly, if a student (or group average) is above the mean and stays at that same relative position, he appears to get farther ahead every year in terms of GE scores. This is an illusion created by the GE score system. ■