White Paper

Why Doesn't ESTIMATOR Use Grade-based or Grade-equivalent Scores?

Richard D. Baer, Ph.D

The short answer to this question is that ESTIMATOR uses age-based standard scores because (1) they have a number of technical advantages that other scores do not and (2) federal regulation require that age-based scores be used when making severe discrepancy calculations.

For a more detailed explanation of the advantages of age-based standard scores read on.

Before answering this question it is important to point out that achievement tests can yield basically four types of age and grade related scores. Two are age-equivalent and grade-equivalent scores. They compare the performance of an individual child to the average child his/her age or in his/her grade. They are expressed as raw scores. For example, the average seven-year old child might attain a raw score of 15 on a test or the average third grade child might attain a score of 22, etc. Such scores are "ordinal scaled." This means that higher scores represent greater achievement than lower scores. However, the differences between score points are not necessarily equal. For example, the difference in achievement between scores of 5 and 6 is not necessarily equal to the difference in achievement between 6 and 7, which is not necessarily equal to the difference in achievement between 7 and 8, etc.

Tests can also yield two other age and grade related scores, age-based standard scores and grade-based standard scores. They also compare the performance of an individual child with the average child his/her age or in his/her grade. Standard scores are calculated by setting the mean (or average) of all scores for a particular age at some standard, usually 100. The standard deviation is also set at some standard, usually 15. The standard deviation is a measure of how the scores spread out around the mean score. For a distribution of scores scaled with a mean of 100 and a standard deviation of 15, 68 percent of scores will fall within plus/minus one standard deviation of the mean or between 85 – 115. Similarly, 95 percent of scores will fall within plus/minus two standard deviations of the mean or between 70 - 130 and 99 percent of scores will fall within plus/minus three standard deviations of the mean or between 55 – 145. Also, scores in an age-based or grade-based standard score distribution are "interval scaled" so that the difference between scores is equal. For example, the difference between scores of 100 and 101 is equal to the difference between scores of 101 and 102 which is equal to the difference between scores of 102 and 103 and so on.

Intelligence tests are always expressed as age-based standard scores.

Age-based or grade-based standard scores have a number of advantages over age-equivalent and grade-equivalent scores:

1. They allow us to compare the performance of a particular child on a test with the performance of the average child his/her age along an interval scale. Thus, if a child obtains a score of 85, we can say his/her score is one standard deviation below the mean. Also, that his/her score is below the scores attained by 65 percent of other children his/her age. Similarly, if a child attains a score of 130 we can say his/her score is two standard deviations above the mean. Also, that his/her score is above the scores attained by 95 percent of other children his her age.
2. The equal interval property of standard scores also allows us to add and subtract scores. We can say, for example, that the difference between scores of 100 and 110 is equal to the difference between scores of 80 and 90. We can also calculate certain helpful statistics. For example, we can calculate a standard error of measurement, a measure of the variability in a student's scores if he/she were given the same test several times. We cannot, however, multiply and divide scores. This requires a true zero and something called a "ratio scale." Because standard scores are interval scaled we cannot say that a child with an IQ of 120 is twice as smart as a child with an IQ of 60 or that a child with a reading achievement score of 50 can read only half as much as a child with a score of 100.
3. Identical scores on different standardized tests mean the same thing. For example, a score of 100 on the Wechsler Intelligence Test for Children indicates average intellectual performance. In the same way, a score of 100 on the Kaufman Tests of Achievement also indicates average achievement performance. This is not necessarily true for age and grade-equivalent scores. For example, 50 on one reading achievement test might be the mean while the mean on another reading achievement test might be 70. At the same time, 120 on a math achievement test might be the mean.

If you don't understand all this immediately, don't worry. It requires some study and some getting used to. The important thing to remember is that standard scores allow us to make more meaningful comparisons of where particular children stand relative to other children their age.

Age-based standard scores are based on and compare children from a "homogeneous" population. That is, all the subjects at an age level are the same age. They are all six years old, or seven years old, or eight years old, etc. Grade- based standard scores are based on and compare children from "heterogeneous" populations. For example, fourth graders in a district or state with little grade retention are not like fourth graders in a district or state with a great deal of grade retention. This can make comparing an individual child to the average child in

his/her grade problematic. For example, we would not want to compare a fourth grader from a district or state with low retention to a national standardization sample with a higher rate of retention. Age-based standard scores avoid this problem.

Again, if you don't understand all this immediately, don't worry. The important thing to remember is that age-based standard scores allow us to make more meaningful comparisons of where particular children stand relative to other children their age.

Beyond the technical merits of age-based standard scores, federal regulations require that aptitude achievement discrepancy calculations be based on age. "The first element in identifying a child with SLD should be a child's mastery of grade level content appropriate for the child's age …" (Federal Register / Vol. 71, No 156 / Monday, August 14, 2006 / Rules and Regulations, p. 46652).

To summarize, in order to make its discrepancy calculations, Estimator needs intelligence and achievement scores that are scaled in the same way. Because they have a variety of advantages, Estimator uses age-based standard scores and not grade-based standard scores or age or grade- equivalent scores. Also. age-based achievement scores are required by federal regulations when making aptitude-achievement discrepancy calculations.

June 3, 2011

References

Aiken, L. R. (2000). *Psychological testing and assessment*. Boston: Allyn & Bacon.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Anastasi, A. & Urbina, S. (1997) *Psychological testing* (7th ed.). New York: Macmillan.

Kamphaus, R. W., Dresden, J., & Kaufman, A. S. (1993). Clinical and psychometric considerations in the cognitive assessment of preschool children. In J. L. Culbertson & J. D. Willis (Eds.) *Testing young children: A reference guide for developmental, psychoeducational, and psychosocial assessments* (pp. 55-72). Austin, TX: PRO-ED.

Lahey, M. (1990). Who shall be called language disordered? Some reflections and one perspective. *Journal of Speech and Hearing Disorders*, 55(4), 612-620.

Linn, R. L., & Gronlund, N. E. (2000). Measurement and assessment in teaching (8th ed.). Englewood Cliffs, NJ: Prentice Hall.

Riccio, C. A., Imahoff, B., Hasboruck, J. E. & Davis G. N. (2004). *Test of phonological awareness in spanish examiner's manual*. Austin, TX: PRO-ED.

Salvia, J., & Ysseldyke, J. E. (2001). Assessment (8th ed.). Boston: Houghton Mifflin.