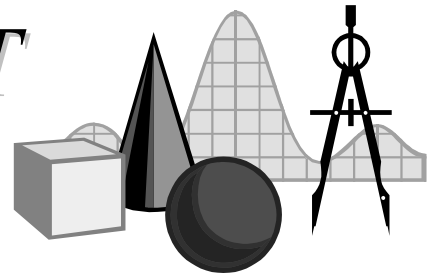


TAKE IT TO THE MAT

A NEWSLETTER ADDRESSING THE FINER POINTS OF MATHEMATICS INSTRUCTION

Math Audit Team
Regional Professional Development Program
March 1, 2000 — Special K–12 Edition



Standardized testing is a major element of public education. Each year, the State of Nevada administers the *TerraNova* by CTB/McGraw Hill. When the results of the test arrive at schools, there are virtually reams of paper with multitudes of different scores requiring interpretation. Digging relevant information from the reports is the most daunting task of administering this exam. In this special edition of *Take It to the MAT*, we will examine the various measures provided by the *TerraNova* and clear some of the fog surrounding them.

Below, terminology from the *TerraNova* is discussed, in addition to some cautions about possible misinterpretations of the various reported scores.

- **Norm-referenced test:** This type of test first collects data from a sample of the expected population called the *norming group*. When individuals or schools later take the test, their results are compared to the norming group's results. In effect, each score is a relative comparison to other examination scores. The *TerraNova* is a norm-referenced test.
- **Criterion-referenced test:** This type of exam compares scores to some pre-determined benchmark, a passing score for example. The *Nevada High School Proficiency Exam* is an example of a criterion-referenced test. The *Objective Performance Indices (OPI)* on the *TerraNova* are criterion-referenced.
- **Raw Score:** The actual number of questions answered correctly.
- **Scale Score:** A score based not only on how many questions were answered correctly, but also on the difficulty and types of these questions. It is analogous to giving different questions on a test different point values. Thus, two students may have the same *Scale Score*, but a different raw score. The *TerraNova* uses *Scale Scores* to determine *National Percentile* and *Normal Curve Equivalent* scores.
- **National Percentile (NP):** When a score is reported as a percentile it means that the individual or group scored as high or higher than that percentage of the population being considered. Hence, a *National Percentile* rank relates how a given scale score compares *nationally*. For example, if an individual's *NP* rank is 65, then that person or group scored as high or higher than 65% of those in the norming group. Percentile scores range from 1 to 99.

Cautions: ♦ A percentile is *not* a raw score or scale score.

- ♦ The median is always the 50th percentile; thus half of all students score below that point.
- ♦ Percentiles are derived from a normal (bell) curve, so differences in *scale scores* will be much greater at very high and very low percentiles than near middle percentiles. Thus, straight comparisons between percentile scores are statistically inappropriate, as is any attempt to average them.

- **Normal Curve Equivalent (NCE):** This score differs from percentile scores in that it is on an equal-interval scale. That is, equal differences in *NCE scores* have equal meaning throughout the scale, unlike percentiles. Differences in *NCE scores* within and between tests can then be compared. *Normal Curve Equivalent* scores also range from 1 to 99.

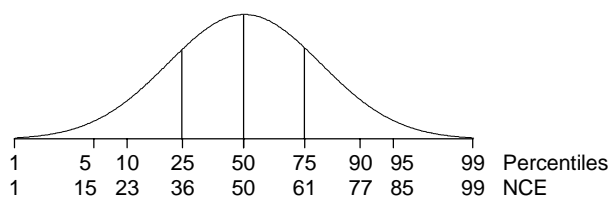


Figure 1

Cautions: ♦ *NCE* scores and percentile scores are only equal for 1, 50, and 99. For percentile scores between 1 and 50, *NCE* scores will be higher than the percentile score; for percentile scores between 50 and 90, *NCE* scores will be lower. See the *Figure 1* for clarification.

- **Stanine:** A stanine is a standardized score based on an equal-interval scale like the *NCE*. Thus, the difference between 8 and 7 stanine students is *approximately* equal to the difference between 7 and 6 stanine students. There are not, however, equal numbers of students in each stanine class. See *Figure 2* for clarification.

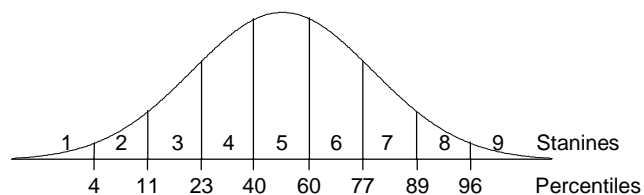


Figure 2

- **Anticipated Scores:** These are estimates of an individual’s score based upon age, grade, and academic aptitude. The aptitude component is determined by the *Test of Cognitive Skills (TCS)*. It yields the *Cognitive Skills Index (CSI)*, a measure of cognitive ability without reference to grade level. The *Anticipated Score* gives a basis for comparison with achievement scores in order to address areas of deficiency and strength.
- **P-Values:** The percent correct on a given test, test item, or group of test items.
- **Item Analysis Report:** This report details *P-Values* for exam objectives, and specific content questions within the objectives. It includes data for comparison from the nation, district, and school. District-nation and school-nation differences are provided, negative numbers indicating performance below national norms.

Cautions: ♦ Goals to improve upon deficiencies should be addressed broadly rather than specifically. For example, a school scores 9 points below the national average on math question number 11, “Convert decimal to fraction.” An inspection of other questions within the same objective reveals no questions asked about converting fractions to decimals. The aim here should not be to *improve conversion of decimals to fractions*, but the broader goal to *improve conversion between decimals, fractions, and percents*.

- ♦ A review of the correlation between *TerraNova* results and the *State Standards* is also warranted. Since meeting State objectives is the goal, not those of the *TerraNova*, deficiencies reported by the test that do not align with the state curriculum should be given decreased emphasis.