Understanding Standardized Testing Statistics

By: Debbie Thompson
Triangle Education Assessments, LLC

THOUSANDS of homeschooled students take standardized achievement tests every year. While homeschoolers generally score in the 80th percentile, well above the national averages, it can be difficult to understand the true meaning and practical applicability of test statistics. Standardized test results can help us analyze our teaching, curriculum and children's academic strengths and weaknesses.

Norm Group

Most nationally standardized tests are based upon a representative norm group, enabling the comparison of an individual's performance to that of a predefined population. Normalized results are representative of averages across the country and include all socio-demographic variables, such as sex, race, and rural/urban location. The norm group also represents averages of students from public, private and homeschools.

Test Reliability

Test reliability refers to the consistency of results if a test is given multiple times to the same student, assuming that all factors remain the same throughout the testing. To achieve greater reliability in measuring growth and progress over the course of your child's education, use the same standardized achievement test for a few years in a row.

Test Validity

When a trait is measured effectively, a test is considered valid. If a curriculum does not match well with questions on a nationally standardized test, then the results might not be valid. An example of this is when your homeschool is studying earth science, but the achievement test assumes a wide range of sciences like anatomy, chemistry, and biology. For this reason, many states do not require homeschools to do science and social studies testing.

Types of Statistics and Quantifiable Scores

Age and grade equivalent scores (AE and GE) indicate a child's level of performance. An AE score of "12-4" means that a student scored as well as an average child aged 12 years and 4 months. Likewise, a third-grade, two-month performance level would be "3-2". These scores are only rough guides and are dependent upon the actual material being taught. A more reliable and useful statistic in determining true strengths or weaknesses is the national percentile ranking.

National Percentile Ranking (NPR)

The national percentile ranking is the statistical difference between the individual's score and the norm population's average performance. If a child performs at the 86th percentile, he or she scored better than 86 out of 100 children in the norm group.

Low or High scores

Sometimes children are not good test takers, but this does not always reflect upon their intelligence. While test taking is a good skill, high scores might not indicate deep or creative thinking. In addition, many other talents and intelligences are not tested on national achievement tests. Remember that the world is filled with talented and successful people who did not achieve high scores on achievement tests.

It is important to understand what test scores do and do not mean. Some children are not good test takers because of test anxiety, inexperience with bubbles, slower work pace, attention or hyperactivity problems, illness, fatigue, hunger, temperature discomfort or any number of other things. A child might miss a line on a bubble sheet and throw off the rest of the answers. Remember that a test score is just a picture of someone's knowledge at one moment in time. Even the best tests are only approximations of true ability.

CAT/5, Iowa, and Stanford/10 score reports are very similar. The boldface titles represent major academic areas or clusters of scores, while the other labels represent subtests within each academic area. For example, the reading test is usually composed of two subtests: vocabulary and reading comprehension.

Typical scores are:

- **Standard Score (SS):** This score is used to obtain the other scores. It has no interpretive value for parents.

- **Grade Equivalent (GE):** This score is very often misinterpreted. If your child is a fourth- grader and earns a GE score of 5.7, this means your child scored about as well as a typical fifth grader (at the end of the seventh month). The GE does not represent a grade level in which a student should be placed, nor does it mean that a child is able to perform consistently at the level indicated by the GE.

- **National Stanine (NS):** Stanines range from 1 to 9 with an average value of 5. Stanines 1 through 3 are considered below average; 4 through 6, average; and scores 7 through 9 represent above-average performance.

- **Normal Curve Equivalent (NCE):** Educators use this score for statistical purposes and to measure change over time. This score has no interpretive value to parents.

- **National Percentile Rank (NPR):** Of all the scores reported, this one is most useful. The NPR represents the percentage of students in the norm group that scored lower than your child did. For example, a rank of 74 means that 74 percent of the national sample obtained scores lower than your student. It would also mean that 26 out of 100 students would have preformed better than your student did. An NPR between the 25th and75th percentiles is considered average and the 50th percentile is the national average.

%Students:	4	7	12	17	20	17	12	7	4
PR:	1-3	4-10	11-22	23-39	40-59	60-76	77-88	89-95	96-99
NS:	1	2	3	4	5	6	7	8	9

Test and/or Skill Statistics and Bar Graphs: These graphs (usually at the bottom of score reports) can give you a quick indication of your child's strengths and weaknesses, but use them with caution:

The percentage of questions answered correctly by students in the national sample (%C Nat.) and the percentage of questions answered correctly by your child (%C Stu) are compared. A negative number indicates that the percentage of questions answered correctly by your child was less than the percentage answered correctly by the national sample.

On the bar charts, a horizontal bar that extends to the right of the centerline suggests areas in which your child performed better than the national average, while a bar extending to the left indicates that your child performed below the national average.

- Use this information carefully. Look at the number of questions used to test each skill. If only a few items (less than 10 or so) are used to measure the skill, then be cautious in interpreting that information.

- If a student did not answer many questions in a certain skill area (Total Items – No. Attempted), try to determine why. Maybe your child has not been taught much in that skill area.

- If a student missed several questions in a skill area (%C Stu), review or different curriculum would typically be recommended. ♠

Debbie Thompson, (www.TriangleEd.com) is Director of Triangle Education Assessments, LLC, which helps thousands of homeschoolers each year with their achievement, ability, career and practice test needs.