

Handy Handouts®

Free informational handouts for educators, parents, and students

Testing Tips

Understanding Tests Jargon and Choosing the Best Tests for Your Caseload!

by Keri Spielvogel, M.C.D., CCC-SLP

Is choosing the best assessment for your student as difficult as administration and scoring combined? What's the difference between standard and raw scores anyway? What do norm- and criterion-referenced mean? Use the following information to choose the best assessments for your caseload and take some mystery out of the jargon used in testing protocols.



Making sense out of testing jargon!

Are all those terms, scores, confidence intervals, and equivalents really important? They are if you really want to understand the scores you report to the district, state, and parents. Help your students and their families by being educated. Below is a list of commonly used terms.

Testing Terms



- **Normed (Norm-referencing)** — The assessment was given to a group of normally developing children that represent the target population to which a child's scores are compared. Ideally, the group is made up of the same percentage of samples (geographic location, race, gender, ethnicity, ages, and socioeconomic status) that are found in the targeted population. For example, if 23% of the population in the Northeast are of Hispanic ethnicity, then the norming group in this region should consist of 23% of children with Hispanic ethnicity (i.e. If total sample is 500 (n=500), then 23% of these would be Hispanic children in the Northeast). The child's performance is then

Testing Terms (continued)



compared to the performance of others. The information derived from this process is used to formulate age equivalency and standard scores.

- **Criterion-referencing** — This refers to the measurement of mastery of specific skills. Unlike norm-referenced tests which measure performance against a group of others taking the test, an individual's performance is measured against a specific criteria or standards. Items are selected based on learning outcomes of the population they target and provide information about how a student has performed on each educational goal included on the test.

Testing Terms (continued)



- **Standard Score** — A score derived from a test that is administered to children in the same manner each time. Each cue is delivered precisely alike.



- **Reliability** — Does the test provide consistent results upon repeated administrations?



- **Validity** — Does the test measure what it is supposed to measure? This varies in accordance with what the test is used for.



- **Age Equivalency** — The age range of children who scored the same on an assessment. (e.g. If a four year old missed a number of questions, he/she might have an equivalency of a two-year-old. In other words, the two-year-olds who took the test scored within the same range.)



- **Confidence Intervals** — A range of scores in which a child's score falls. Confidence levels increase with probability (i.e. A confidence level of 90% means that there is a 90% probability that the child's score falls within that range).



- **Standard Deviation** — A measure of how disperse the data is in relation to the mean (i.e. average). Most children's scores will fall within 1.5 standard deviations of the mean. Score below 1.5 standard deviations may indicate the presence of a disorder.



Information adapted from:

Maxwell, D. L., & Satake, E. (1997). *Research and statistical methods in communication disorders*. Baltimore: Williams & Wilkins.