

Testing Tips!

Understanding Test Jargon and Choosing the Best Test for Your Caseload

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



Is choosing the best assessment for your caseload 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 assessment 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 children and their families by being educated. Below is a list of commonly used 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 compared to the performance of others. The information derived from this process is used to formulate age equivalency and standard scores.
- **Criterion-referencing:** 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. (i.e. Specific grade achievement tests are CRTs.)
- **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. (i.e. If a four year old missed a number of questions, he/she might have an equivalency of a two-year-old. Or, 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 fell within that range).
- **Standard Deviation:** A measure of variability derived by the variance. This is generally represented by a normal curve. Generally, children who fall 1.5 standard deviations below the mean (middle of the curve) qualify for speech and language services. In a normal curve, scores may fall one standard deviation below or above the mean for a child to be testing as "normal."

Now I know what it means, but what test do I choose?

Understanding the terminology within testing protocols is important. However, choosing an appropriate assessment for your caseload is essential. Moreover, it is important to understand administration and scoring procedures to decrease error. The following tests are organized by category and represent only a small sample of diagnostic materials in each area.

Articulation & Phonology:

- Clinical Assessment of Articulation and Phonology™ (CAAP™): (2002) Evaluation of articulation using simple words, cluster words, multi-syllabic words and sentences. Phonology is assessed using a checklist format that does not require phonetic transcription. Test is norm-referenced. Kit includes a stimuli easel, manual, forms (Articulation-50; Phonology Checklists I and II -30) five foam CAAP™ dolls, and a tote. Total administration and scoring time is approximately 30 minutes. Appropriate for ages 2.6 to 8.11.
- Goldman-Fristoe Test of Articulation-2 (GFTA-2): (2000) Evaluation of articulation using simple words, clusters, sounds in sentences, and stimulability in the initial, medial, and final positions. New normative data is provided for ages 2-21. Kit includes manual, stimuli easel, forms (25) and canvas carry bag. Evaluation time is approximately 10-15 minutes for Sounds-In-Words subtest. Can be used in conjunction with the KLPA-2 (Khan-Lewis Phonological Analysis, 2nd Edition) to derive a phonological process score.

Language:

- Test of Language Development – Primary Fourth Edition (TOLD-P:4) (2005) Assesses nine different areas of language (Picture, Relational, and Oral Vocabulary, Grammatical Understanding and Completion, Sentence Imitation, Word Articulation, Phonemic Analysis and Word Discrimination). Three of the subtests are supplemental, or optional. Testing time is approximately one hour. Test is normed for ages 4 – 8.11.
- Oral and Written Language Scales (OWLS) (1995/1996) Three parts including the Listening Comprehension Scale (LCS), Oral Expression Scale (OES) and the Written Expression Scale (WES). LCS assesses receptive language for ages 3– 21. The OES assesses expressive language skills ages 3 – 21 and doesn't require reading skills. The OES allows categorization by item type (lexical, syntactic, pragmatic, and supralinguistic). The WES, for ages 5 - 21 years, assesses language-based writing skills and is co-normed with Oral Language Scales. It can be administered individually or in groups and takes approximately 15-25 minutes to complete. Test is standardized and normed.

Fluency

- Stuttering Prediction Instrument for Young Children (SPI) – assesses dysfluency in children ages 3 - 8 years old using parental interview, observation, and tape recording of a child's speech. Analysis of the sample assesses reactions, part-word repetitions, prolongations, and frequency of stuttered words. Gives severity scales of dysfluency.
- Stuttering Severity Instrument for Children and Adults – Fourth Edition (SSI-4) – Assesses dysfluency in children and adults ages 2.10 and up. in four different areas (frequency of dysfluencies, duration of dysfluency, physical concomitants (secondary characteristics), distractibility, and severity. Form is easy-to-score and easy to follow. Provides pictures and grade-specific reading passages.



Choosing the most appropriate test for your caseload can be challenging, especially if the jargon used in most testing protocols is too cumbersome or difficult to understand. Good luck and happy testing!

Information adapted from:

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