



Language Assessment: A Hit-and-Miss Approach to “Finding Kids” March 2008

[This session presentation was given at the 2007 ACSLPA Conference by Eleanor Stewart, R.SLP, and Denyse Hayward, R.SLP, and the follow-up article was published in 2008.]

Session Synopsis

The business of ascertaining program service funding hinges on our ability to accurately identify children who have speech and/or language needs. In our zeal to secure funds for deserving children we may be inadvertently engaged in a hit-and-miss approach when we rely on test scores alone. Given that funders must make tough choices with limited resources, it is not surprising that cold hard numbers are enticing. But what about those test scores? Are the right children being identified? Did you know that you cannot use the norms of the PLS-4 unless a ceiling of seven (not five) failed items is achieved? Did you know that identifying children with language impairments on some tests is no better than chance? We can and should be able to identify the limits of standardized tests in order to be more confident in our test choices and clinical judgment in order to advocate both individually and collectively to funding agencies on behalf of children with communication disorders.

These questions are addressed as they relate to the tests we most often use to identify children with language needs (e.g., the Preschool Language Scale and Clinical Evaluation of Language Fundamentals). Do these tests do the job we ask? What do we really mean when we say “I like Test A better than Test B”? A closer look at these tests will assist clinicians in understanding whether such statements are objectively valid, under which circumstances using Test A over Test B is in the best interests of the children we serve, thus avoiding a hit-and-miss approach.

Follow-Up

We were very encouraged by the number of you who took the time to follow up with us about our presentation at the ACSLPA Conference in October 2007 and for discussing these issues with your colleagues. We had been hearing from some of them, too!

Our talk was aimed at supporting clinicians using formal testing measures to qualify children for early intervention quota programs. Our goal was to provide you with information about formal testing measures to enable you to justify and defend your decisions in terms of the tests you choose to use and your interpretation of the child’s test performance.

We chose to focus on the two tests most widely used in assessing young children who may be eligible for PUF; the PLS-4 and the CELF-P2. Here are the highlights of our presentation. We offer these key ideas to reinforce and support your clinical decision-making and to promote discussion with your colleagues:

1. **Prevalence estimates of language difficulty in preschoolers are 2% to 19% (ASHA, 2008).** It is important to note that prevalence estimates do

not address severity of language difficulties. Therefore, we can expect a smaller percentage of children will have *severe* language learning difficulties. Thinking about prevalence helps us to remember to expect variability in the number of children we will identify for quota programs each year. This will also result in variability in the number and distribution of children identified within and across school boards/regions each year. Finally, cut-off scores such as scoring at or below the first percentile are an inevitable aspect of quota programs due to a limited set of financial and personnel resources. This is important to keep this in mind, as it is very easy to adopt such criteria as clinical practice guidelines when in fact determining disorder severity requires consideration of factors above and beyond test scores.

2. **Testing standards allow for a test to continue to be used for only two years after the revised edition is published:** The PLS-4 replaced the PLS-3 five years ago and the CELF-P2 replaced its earlier edition three years ago. The transition time for clinicians to adopt the new test versions is long past, such that no clinician in Alberta should be currently using either the PLS-3 or CELF-P for any formal testing purpose.
3. **The ceiling rule for PLS-4 is seven consecutive errors:** In her 2006 review of the PLS-4 in the *Journal of Speech-Language Pathology and Audiology*, Pat Cleave states that clinicians should establish a ceiling of seven consecutive errors if they wish to report standard scores because if adequate psychometric properties could have been achieved with fewer ceiling items (i.e., five) the test developers would have used a lower ceiling to develop the norms. Since this is not the case, we must assume that in order to achieve acceptable psychometric properties a ceiling of seven consecutive errors was necessary.
4. **Measurement error occurs because no test is perfectly reliable.** In order to deal with this, test developers offer an estimate of error called the standard error of measurement (SEM). The SEM is the amount by which the obtained score may differ from the true score because of errors in measurement. The smaller the SEM, the more confident you can be in the obtained test score. The SEM is used to construct a confidence band that allows us to report how confident we are that the child's test score is actually in the range of the scores (e.g., 90%, 95%, 99% confident). In fact, the PLS-4 manual states, "You would never consider the score obtained on a standardized test as a perfectly precise measure of ability. It is most appropriate to consider a child's true score as lying within a range of scores. The confidence band" (pp. 127-128). When determining if a child's score on a test is at or below the first percentile confidence bands are the more appropriate and accurate scores to report. Using the PLS-4 normative tables indicates children between 3:0 and 3:5 with an EC standard score of less than or equal to 75 have confidence bands that include the first percentile.
5. **Language Sample analysis:** Almost anyone can give a test and even interpret it according to the examiner's manual. But, language sampling is the unique province of SLPs. Use samples to your best advantage and to the advantage of the child. A 50 utterance sample for a child with a severe

language delay will provide far more information than either the CELF-P2 or PLS-4 on language productivity in naturalistic contexts. Additionally, normative information is available to support eligibility for quota programs (see R. Paul's book for normative data, p. 319). Given that the MLU range for typically developing three-year-olds is 2.5 – 3.9 morphemes, the time needed to obtain a 50 utterance sample on a severely delayed three-year-old is neither as time-consuming or onerous as many clinicians believe. If a child is not able produce 50 consecutive utterances, remember that that is information, too (i.e., rate of communication).

Based on our study of formal testing measures to qualify children for quota programs we recommend the following:

1. Use only the current test editions (i.e., PLS-4 and CELF-P2).
2. Only report PLS-4 scores when ceilings of seven are reached. Otherwise, describe the child's performance in terms of items passed, behaviours that might have interfered, and the patterns evident (e.g. scattered performance between 19 and 24 months).
3. Report confidence band information to profile the child's performance range. For example, you can state, "X achieved a percentile rank of three with a true score range between 1-7 with 95% confidence."
4. Collect a 50 utterance spontaneous language sample and calculate the MLU.
5. Use your clinical knowledge of developmental stages to describe the child's language skill level. For example, an expressive vocabulary of 12 words at 24 months indicates a significant delay when you consider that on average a 24-month-old would have an expressive vocabulary of approximately 300 words. (Fenson, Dale, Reznick, Hartung, and Burgess, 1990 cited in Rhea Paul's 2001 book, *Language disorders from infancy to adolescence*).
6. Use more than one measure to create a "profile" of the child's language skills.

It is the profile that you create that is important. The profile reflects your clinical expertise as a speech-language pathologist.

References

- ASHA. (2008). Prevalence estimates of language difficulty in preschoolers 2% to 19%. *Communication Facts: Incidence and Prevalence of Communication Disorders*.
- Cleave, P. (2006). Resource review: Preschool language scale. (4th ed.). *Journal of Speech-Language Pathology and Audiology*, 30(4), 252–253.
- Paul, R. and C.F. Norbury (Eds.) (2012). *Language disorders from infancy to adolescence: Listening, Speaking, Reading, Writing and Communicating*. (4th ed.). St. Louis, MO: Elsevier Mosby.
- Zimmerman, I.L., Steiner, V.G., & Pond, R.E. (2002). *Preschool language scale-4*. San Antonio, TX: Harcourt Assessment.