

# PEARSON

## Answering Tough Questions About CELF-4 Interpretation



Elisabeth H. Wiig, Ph.D.

CELF-4 Author

April 12, 2011

## A Tough Lead-In Question

I am requesting clarification on whether one should administer the CELF-4 or the CELF-Preschool-2 to a child who is still in preschool, yet turned 5 years old, two or three weeks before the test was administered.

Would one be considered more valid and or reliable than the other? Should the test scores come out the same if they were both administered? I ask because a child of this profile was recently evaluated using the CELF-4. His scores came out significantly lower than we expected, and his preschool performance is not consistent with these scores.

# A Long Answer

- While structured like CELF-4, CELF Preschool-2 is much more child-friendly than CELF-4, which is designed for school-age children. In a validity study conducted with the two tests, correlations ranged from .81 to .86—high positive correlations, but not identical scores.
- Because you didn't give me the actual scores, I can only speak generally about the test results. What the information she obtained tells her is that this Kindergartener performs better when tasks are simpler and include more visual support. When there is less support (such as that on CELF-4), the child does not manage the language tasks as well.
- I would also consider the child's skills in his preschool classroom. Does he handle classroom tasks well? Is he having a lot of difficulty and need a lot of support? If the former, you may consider giving CELF 4. If the latter: CELF Preschool 2 which offers more support with content geared to preschoolers.

# Word Classes

Q: The subtest **Word classes for children 5-7** offers the child a visual of the words we offer them. Now **many of the children will identify the match visually (before we even mention the words), is this still considered to be a receptive language task?** The children could do this task by just looking at the pictures.

A: Yes, however you should still administer as per standard procedures (any verbal directions) when possible. When the child correctly identifies the similar pair they are demonstrating their understanding of a/the semantic relationship. If the child responded incorrectly, it would suggest/provide evidence of the lack of this knowledge. In the Item Analysis section you are able to record the Various Semantic Class Errors (conceptual errors).

# Word Classes

Q: Many of our children will score very low on Concepts and Following Directions and Sentence Structure, however very good on the task Word Classes, which impacts the total receptive score significantly. However if we test these children a year later their scores drop severely because as of the age of 8, we take away the visual and all of a sudden they have low scores on the subtest word classes as well. Why would that happen?

Now the biggest dilemma we have with this is that we need to use Total receptive language Scores to submit our children for coding, and so far children between 5-7 are being rejected (even though they clearly have a receptive language disorder) because the subtest Word Classes pulls their score up so much. [Note: Our children are from a culturally and linguistically diverse population]

A: We have not seen a consistent pattern of Word Classes scores being so high that children no longer qualify for services. That does sound like an issue for your students in your work setting (and I can see your concern if they are not qualifying for services based on high performance on one of the three subtests.)

# Word Classes

Q: If a child gives an incorrect pair (receptive) does the clinician give the correct pair and then present the second part of the item?

# Word Classes

A: CELF's Word Classes subtest includes two tasks, a receptive task and an expressive task.

- On the receptive task, examiners ask students to select two words (out of three or four stimulus items) that are related.
- On the expressive task, examiners ask students to explain how the two selected words are related.

Regarding the expressed confusion on how to administer the expressive task if students responded incorrectly to the receptive task... The examiner should ask how the child's responses go together, even if incorrect responses were given. Here's an example.

Item stimulus: cat, dog, balloon

Student's response on the receptive task: *cat* and *balloon* Correct administration, the examiner should ask: "How do the words *cat* and *balloon* go together?"

# Word Classes

Q: In the **Word Classes** subtest does the child have to say the word or is pointing enough?

Pointing is enough. Note as pointed correctly or incorrectly for observations.

Q: What if the child can't remember the word but points correctly? (same question I guess)

A: You could repeat the words and have them respond again. Pointing would be acceptable.

Q: What if the child points correctly but says a different word? (for example, life jacket for swimsuit)

A: This would be acceptable, if the student pointed to the correct response and the verbal response is similar to one of the responses. When a response is similar to the listed responses write it verbatim for later scoring consideration. You could also repeat the directions/words if necessary.

# Phonological Errors affecting Expressive Responses

Q: In Recalling Sentences, if a child presents with the phonological process--final consonant deletion and or syllable reduction--how do we score that?

Example The tractor was followed by the bus,

The child responds the tractor was follow by the bus--would we score this as 1 error, due to the missed morpheme -ed ,

A: Yes, but then you would look at the pattern of the child's errors on this subtest. If the child consistently omits final sounds, you will need to note that information in your report and decide if the child has an overall pattern of morphological errors, phonological errors, or both.

Q: Further on in the test the child responds with the following

The *rabbit* was not *put* in the cage by the girl. (two final consonant deletions) for -The rabbit was not put in the cage by the girl. Are those errors?

A: This is a phonological process error rather than a morphological or memory issue. You do not count misarticulations as errors (missing the /t/ at the end of *rabbit* and *put* are not errors. )

# Phonological Errors affecting Expressive Responses

**Q: On Formulated Sentences, how do we score if there is a clear phonological process going on.**

**A:** Misarticulations are not scored as errors. If the child has consistent final sound deletions, you can choose to score morpheme deletions such as plural s or past tense ed as errors or you can choose to skip that subtest due to inability to evaluate the child's skills on those items due to the phonological process issues.

Of course, if a subtest is missing, you cannot calculate an index score for Expressive Language or Structure. You will have to do a qualitative description of the child's skills.

# Phonological Errors affecting Expressive Responses

**Q:** What if the child pronounces a word in the Expressive Vocabulary test such as ‘fermodeber’ for thermometer, ‘pinando’ for piano, ‘skeletope’ for telescope? Is that considered acceptable or does it get 0?

**A:** Do not score 0 for articulation errors.

# Scores

Q: Why are subtest scores “scaled scores” and index scores “standard scores”?

A: We use the term “scaled scores” for subtests with a mean of 10 and standard deviation of 3.

We use the term “standard scores” for index scores with a mean of 100 and a standard deviation of 15. There are more items that factor into index scores, resulting in a wider range of scores and more precision in capturing a student’s skills.

# Scores

Q: I think there is an error in the test scores. I had a child who took the test when he just turned 9 and his standard scores came out very low. If I had tested him 3 days before when he was 8, he would have performed in the normal range with standard scores of 9s and 10s!

# Scores

A: There is a great deal of language growth that occurs between at ages 8 and 9.

In addition, when you test a child at the very end of the norm group range (e.g., a child who is 8 years, 11 months, and 27 days), you are comparing him to students who are mostly younger than he is. When you test a child at the very beginning of a norm group range (e.g., a child who is 9 years and 3 days), you are comparing him to students mostly older than he is.

If you believe scores are erroneously low or erroneously high because of the child's age at the time of testing, this information should be part of your interpretation of test results and recommendations to the team when making a placement decision.

## More on age ranges

Q: I have a question on chronological age and scoring. I administered half of the subtests for the CELF-4 to a child who was 5 yrs. 5 months 30 days of age on the first day of administration. I was not able to finish the test.

I administered the second half of the test a week later and the child was technically now 5 yrs. 6 months. There are two sets of score ranges in the manual, one for ages 5 yrs. 0 months to 5 yrs. 5 months, and the other for 5 yrs. 6 months to 5 yrs. 11 months.

This child's score would be different depending on which age range that I use. I was just wondering which age would be appropriate to use since the test was started at the younger age range, but completed at the upper age range.

# More on age ranges

A: Choose one norm group to use for all subtests. Don't use the younger norm group for the subtests administered used in the first session and the older norm group for subtests administered in the second test session.

If most subtests were administered in one session or the other, use the test date in which more subtests were administered.

If the examiner administered half of the subtests in the first session and half of the subtests in the second session, the examiner will have to choose one test date to use in calculating chronological age. If the examiner feels that the child's performance on the test in one test session more accurately reflected his or her language ability than at the other test session, the examiner should use the test session date that most accurately reflects the child's language ability.

If a child has really poor language skills, he or she will score low either way so it doesn't matter which norm group is used. Usually, though, it makes the difference between a child qualifying for services and a child not qualifying. The examiner should match the CELF results, then, to the other assessment information they have (from language sampling, teacher interview, Observational Rating Scale, dynamic assessment, etc.)

# Score Changes: The Other Way

- I have a client who has a child in her school district. This child was given the CELF-4 twice now- once when she was in Kindergarten and just now when she is in 5th grade. The scores for some of the subtests did not remain consistent, and this client is trying to interpret them. She wants to know if a child's scores would normally stay about the same over multiple assessments. I can't see that we have done any long term studies that show how you might expect a child's scores to change if you use the CELF-4 over a number of years. I did warn her that we cannot interpret test scores, because we don't know the child involved. But apparently none of their SLPs understand why the scores would change this way, and they would like any knowledge that we have. Do you have any information for them?

# Score Changes: the other way

## Child in K (6:4)

- C&FD -9
- WS -9
- RS -13
- WC -R & -T were both 12
- EV - 9

## Child in 5th grade (10:10)

- C&FD -12
- RS -12
- FS -14
- WC -R was 8
- WC -E & -T were both 7
- WD -10
- USP -7

# Score changes: the other way

- A: The scores for the test are expected to shift depending on how the child's skills evolve over the course of time, especially if the child has had therapy or classroom intervention. Otherwise, a clinician wouldn't need to ever conduct a retest. That is why the laws are written to require multiple assessments over time to check the child's status.
- From looking at the scores, the student is doing much better now in understanding concepts and following directions. He or she isn't keeping up as well as peers in understanding relationships among words. This could be attributed to the content of the therapy he or she has received—what skills did therapy focus on?
  - We can't do a long-distance interpretation of scores, but it is not unusual for a child's skills to shift over time, depending on the amount of therapy the child has received, the time elapsed, and focus of therapy activities.

# Index Score Differences

Q: I have a particular question in regards to a discrepancy between LCI and LMI. There was a +22 point difference, which is unusual. The Core language and RLI, ELI, LCI and LMI were all in the average range, with LC being the highest at 110 and LM being the lowest at 88. This is why the difference is unusual. What does this mean? The child is 9-5 and was referred for a full evaluation due to problems with written and narrative expression.

# Index Score Differences

- A: All scores are within normal limits, so I am not sure that there are clinically relevant implications to a 22 point difference. If memory was in the disordered range, there would be some practical application to the score difference (you could point out to teachers and family what strategies the child needs to perform well in the classroom). In this case, memory is still within normal limits (within 1 SD of the mean). The purpose of examining score differences is to identify a student's significant strengths and weaknesses so that you can leverage strengths to support weaknesses. A score of 88 does not reflect a "relative weakness", even if the score is significantly below the other scores. There is essentially no need to identify critical values and prevalence information with scores in the normal range.
- That being said, you may want to refer the child to a psychologist for a full evaluation to examine memory issues more fully (CELF-4 is not a comprehensive measure of memory). Language skills appear solid.

# CELF-4 Spanish

Q: I test a lot of students who have some skills in Spanish and some skills in English. Can I give credit for English responses on the Spanish test? What about giving credit for correct responses in Spanish on the English test?

# CELF-4 Spanish

A: You can definitely give credit for English responses on the Spanish test—the norms were collected this way.

We did not collect norms for the English edition by accepting other-language responses. I would mark responses as errors in English and calculate the scores on that basis. But I would note in my report that the student knew X number of additional items in Spanish and base my recommendations on what the student knows in both languages rather than strictly relying on the test scores.

# CELF-4 Spanish

Q: If an SLP wanted to use the CELF-4 to test a student in both English and Spanish, what practice effects might they expect? Would it even be appropriate to use both at the same time?

A: Practice effects should be very limited if the tests are not administered within a few days. There are very test items that overlap on the two tests.

## If you miss the live presentation...

- A recording of the presentation will be available on [www.speechandlanguage.com](http://www.speechandlanguage.com). (CEUs are not offered for the recorded presentation.)
- Look for Dr. Wiig's answers to questions submitted by the webinar participants on [www.speechandlanguage.com](http://www.speechandlanguage.com) over the next few weeks.