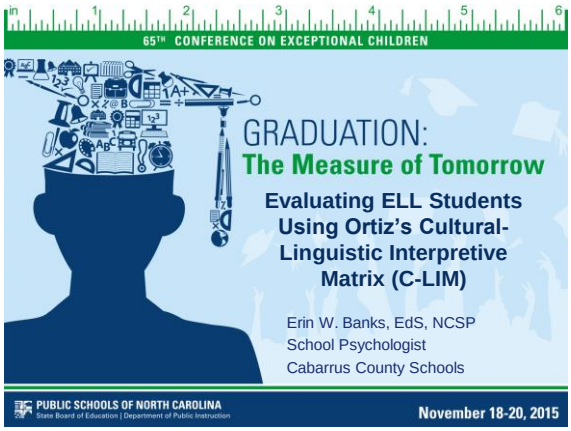


Session Objectives

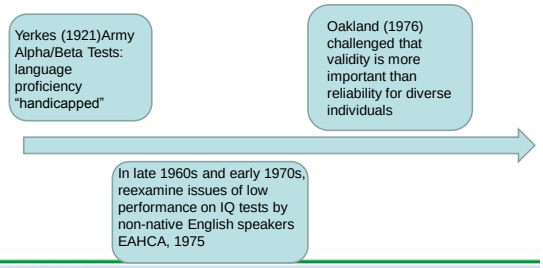
- Understand a brief history of assessment and impact on culturally and linguistically diverse learners
- Discuss different types of assessment techniques for evaluating culturally and linguistically diverse students
- Understand importance of validity and how the C-LIM can help determine validity of test results
- Understand the steps to entering data and interpreting results using the C-LIM and how to use such data for evaluating culturally and linguistically diverse students

Sattler's response

- Sattler (1992):

"Probably no test can be created that will entirely eliminate the influence of learning and cultural experiences. The test content and materials, the language in which the questions are phrased, the test directions, the categories for classifying the responses, the scoring criteria, and the validity criteria are all culture bound…In fact, all human experience is affected by the culture, from prenatal development on. (p.579)"

A historical framework



Validity

- Fair and equitable interpretation of test results is predicted on an understanding of the assumptions that underlie testing and the degree to which these assumptions are violated in the case of testing an individual whose background experiences and development are different from those of the individuals on whom the test was normed

Accommodations/modifications that compromise validity

- Use of Interpreter
  - Up to 20% of practitioners use this modification (Sotelo-Dynega, Cuskley, Geddes, McSwiggan & Soldano, 2011)
- Additional time
- Eliminating certain test items with presumed culturally biased content
- Repeating verbal instructions to ensure full comprehension
- Accepting responses in native language of student
- Administering only subtests that do not rely on oral expression

- ***Problem? Breaks standardization and significant error is introduced into testing situation***

Take-Home point

- If a test is administered with any type of alteration or modification not specifically allowed (i.e. translator, elimination of items, repeating instruction), the validity of scores from that test cannot be evaluated using the Culture-Language Interpretive Matrix (C-LIM).

Nonverbal testing

- Intend to eliminate the language barrier
- Popular in psychological practice
- Survey results:
  - (Sotelo-Dynega et al 2011): 88% practitioners give nonverbal assessment to ESL students
- Language-Reduced testing instead of Nonverbal Testing
  - Cannot eliminate all language

IQ/Achievement Discrepancy:

UNIT FSIQ = 98

WJ-III Basic Reading = 75 (-23)
WJ-III Reading Comp = 72 (-26)

SLD in Basic Reading and Reading Comprehension

Four Problems with NV Tests and SLD

- 1. NV intelligence tests may be no more fair or valid than those that include verbal abilities (Figueroa, 1989)
- 2. Range of abilities measured by NV composite is narrow than verbal batteries (Flanagan, Ortiz, & Alfonso; Ortiz, 2008).
- 3. Majority of referrals for SLD are language-based, particularly in reading. If you test for cognitive deficits related to reading: Auditory Processing (Ga), Crystallized Knowledge (Gc); not typically a part of nonverbal assessment.
- 4. NV tests are subject to the same problems with norm sample representation: cannot control for differences in acculturation or language development

Native language testing

- Inaccurate label of bilingual assessment
  - Implies evaluation that is to be conducted bilingually with use of two language
- Better described as Consecutive Monolingual Testing
  - Even in situations where a test is given in one language followed by testing in another language
- *Requires that psychologists speak the language of the test*
- Little research about performance of bilingual students on monolingual tests administered in native language (Esparza-Brown, 2007).
- Complicated by factors such as individual's age, level and type of prior education, current language of instruction, type of instructional problem (Goldenberg, 2008)
- WISC-IV Spanish publisher:
  - If student has been in US longer than 5 years, administer English WISC-IV

English-language testing

- Historical research shows that nonnative English speakers consistently perform more poorly (1 full standard deviation) as compared to their native English speaking counterparts on *subtests that have high language demands*
- Nonnative English speakers perform comparably to native English speakers that do not require much, if any, verbal or language-based development or skills
- Best Practice Guideline: Nondiscriminatory Comprehensive Assessment
- (Cummins, 1984; Feigueroa, 1989; McShane, 1980; Mercer, 1979; Naglieri, 1982; Valdes & Fiugueroa, 1994; Vukovich & Figueroa, 1982).

Nondiscriminatory Assessment: Processes and Procedures

I. Assess for the purpose of intervention

II. Assess initially with authentic and alternative procedures

III. Assess and evaluate the learning ecology

IV. Assess and evaluate language proficiency

V. Assess and evaluate opportunity for learning

VI. Assess and evaluate relevant cultural and linguistic factors

VII. Evaluate, revise, and re-test hypotheses

VIII. Determine the need for and language(s) of formal assessment

IX. Reduce bias in traditional assessment practices

X. Support conclusions via data convergence and multiple indicators

Integration of RTI Within General Education Framework

Pre-referral procedures (I. - VIII.)
Post-referral procedures (IX. - X.)

Unless otherwise indicated, information contained in this packet is Copyright © Samuel O. Ortiz, Ph.D. and may not be reproduced without permission.

Cultural and Linguistic Classification of Tests Addressing Validity in Diagnosis and Interpretation

Pattern of Scores on the Wechsler Subtests

Subtest	Monolingual	Bilingual	Difference
VOC	103.75	87.67	-16.08
INF	99.57	86.30	-13.27
SIM	103.68	91.12	-12.56
COM	100.66	89.88	-10.78
ARI	98.11	89.35	-8.76
CD	105.57	98.21	-7.36
PC	99.91	97.92	-1.99
PA	97.36	96.14	-1.22
OA	96.89	96.70	-0.19
BD	97.08	97.29	0.21

Source: Nieves, B., Ortiz, S.O., Flanagan, D.P., Chaplin, W. (2006), unpublished data..

Unless otherwise indicated, information contained in this packet is Copyright © Samuel O. Ortiz, Ph.D. and may not be reproduced without permission.

Introducing…Cultural-Linguistic Interpretive Matrix (C-LIM)

- Determine the extent a student's performance is influenced by cultural and linguistic factors; other than actual ability

Helps to determine ***VALIDITY***

- *Typically, tests designed to measure CHC abilities like Gc, Ga, Gsm may measure instead a student's lack of English proficiency or familiarity with cultural content*

- Helps to answer : "***Difference versus Disorder"***
  - The Culture-Language Test Classifications (C-LTC) and the Culture-Language Interpretive Matrix (C-LIM) were developed expressly for this purpose.

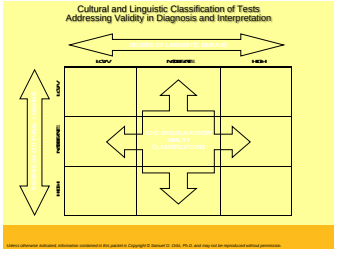
Culture-Language Test Classifications

- C-LTC and its applied companion, the C-LIM created by attempting to identify specific tests that would likely yield the most valid results with respect to measuring the full range of cognitive abilities specific by the McGrew integrated Gf-Gc model (Flanagan et al., 2000; McGrew & Flanagan, 2000) and later CHC theory (Flanagan & Ortiz 2001).

- C-LTC is viewed as an extension of the CHC theoretical classifications presented in part as the basis of the XBA approach
  - Not tied to XBA and can be used with an interpretive measure; however, strongly recommended to use with CHC theory interpretation and evaluation
- C-LTC
  - Seeks to identify tests with lowest levels of cultural loading and linguistic demands

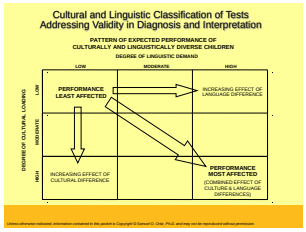
C-LTC organization

- 3X3 matrix table with degree of cultural loading along vertical axis
- Degree of linguistic demand along the horizontal axis
- Emphasis as a continuum rather nonverbal/verbal dichotomy (Ortiz et al, 2012).



C-LIM

- Based on consistent research finding for culturally and linguistically different students that demonstrates a pattern of lower performance on tests with higher levels of cultural content and higher degrees of linguistic demand, than on tests that are lower on these two dimensions
- When test data are entered, program automatically calculates an aggregate score for each cell
- Examine scores visually to determine if resulting pattern matches expected level for learner
- Research-based method to interpret test results



Step 1. Determine degree of difference

Clinical Judgment

The greater the difference, the greater the adverse effect on performance. Therefore, it is important to make this determination as accurately as possible and to use it as appropriate basis on which to evaluate the impact on test scores.

Use definitions for each to determine which the student best fits

XBASS Software System C-LIM

- Click one of the following:
- 1. Slightly Different
- 2. Moderately Different (default)
- 3. Markedly Different

General Guidelines for Expected Patterns of Test Performance for Diverse Individuals

DEGREE OF CULTURAL LOADING	DEGREE OF LINGUISTIC DEMAND: Low	Moderate	High
LOW	Slightly Different: 3-5 points Moderately Different: 5-7 points Markedly Different: 7-10 points	Slightly Different: 5-7 points Moderately Different: 7-10 points Markedly Different: 10-15 points	Slightly Different: 7-10 points Moderately Different: 10-15 points Markedly Different: 15-20 points
MOD	Slightly Different: 5-7 points Moderately Different: 7-10 points Markedly Different: 10-15 points	Slightly Different: 7-10 points Moderately Different: 10-15 points Markedly Different: 15-20 points	Slightly Different: 10-15 points Moderately Different: 15-20 points Markedly Different: 20-25 points
HIGH	Slightly Different: 7-10 points Moderately Different: 15-20 points Markedly Different: 20-25 points	Slightly Different: 10-15 points Moderately Different: 15-20 points Markedly Different: 20-25 points	Slightly Different: 15-20 points Moderately Different: 20-30 points Markedly Different: 25-35 points

Slightly Different: Includes individuals with very high levels of English language proficiency (e.g., advanced BICS/emerging CALP) and high acculturation but are not actually fully acculturated, e.g., third generation. Speaks English very well and has limited bilingual/bicultural experience.

Moderately Different: Includes individuals with moderate levels of English language proficiency (e.g., intermediate to advanced BICS) and moderate levels of acculturation, e.g., second generation. Speaks English well, appears well acculturated but has significant bicultural/bilingual experience.

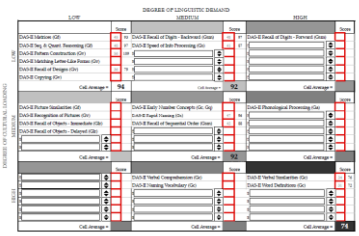
Markedly Different: Includes individuals with low to very low levels of English language proficiency (e.g., early BICS) and low or very low levels of acculturation, e.g., first generation. Does not speak English well yet, relatively new to the U.S., significant native culture and language experience.

Unless otherwise indicated, information contained in this packet is Copyright © Samuel O. Ortiz, Ph.D. and may not be reproduced without permission.

Step 2/3: data entry instructions

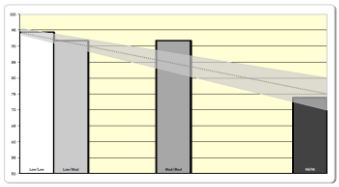
Enter subtest scores obtained from evaluations

- Use scaled scores or standard scores
- T-Scores on DAS-II subtests will be converted to standard scores
- Cell scores will be averaged for an overall cell aggregate



Step 4: interpretation

- Cells generate average
- Visually see the cells and graph
- Always start by looking at the top LEFT cell and analyze across, down, and diagonally
- Shaded area (trend line) – degree to which the individual was deemed to be different



GRADUATION: The Measure of Tomorrow
65TH CONFERENCE ON EXCEPTIONAL CHILDREN

CAUTION

- C-LIM is not a diagnostic tool and does not provide a mechanism for diagnosis. It is designed only to assist in determining whether results obtained from standardized testing are valid (and may therefore be interpreted) or not.
- Do not include graphs or tables in your psychological report.
- May use Interpretive Statements embedded in your report and on the DEC paperwork (handout)

Interpretation: Declining Pattern

Practitioners have to recognize that the invalidity of their results indicates no interpretation can be made and no direct inferences drawn regarding actual ability

If level of acculturation and language proficiency are manifested in pattern of test scores, they cannot be ignored or excluded and remain confounding influences

Student must be presumed average or normal in their abilities unless other data are found

Test scores cannot be used to bolster the presence of any disability if they follow the predicted declining pattern

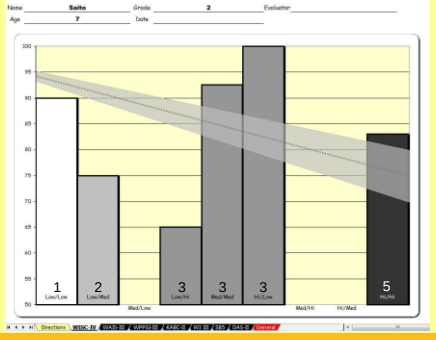
*****Remember: Using the C-LIM is only one part of TEN STEPS for a comprehensive nondiscriminatory evaluation*****

Interpretation: No Declining Pattern

- Absence of a declining pattern of test score performance does not automatically mean that the scores are valid or there is a disability
- ***C-LIM is not a diagnostic tool*
- Absence of a declining pattern indicates only that cultural and linguistic factors cannot account for the entirety of the obtained results
- Pattern suggests that there is some other factor(s) present in the results
- Analyze specific subtests causing the difference in predicted results

Culture-Language Interpretive Matrix: Case Study 6

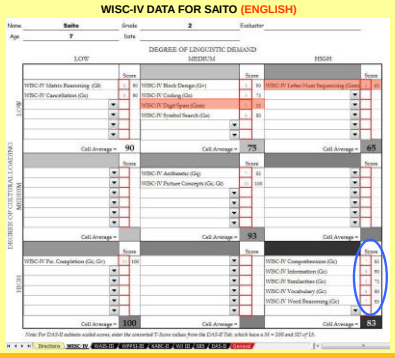
WISC-IV DATA FOR SAITO (ENGLISH)



Unless otherwise indicated, information contained in this packet is Copyright © Samuel O. Ortiz, Ph.D. and may not be reproduced without permission.

Culture-Language Interpretive Matrix: Case Study 6

WISC-IV DATA FOR SAITO (ENGLISH)



Unless otherwise indicated, information contained in this packet is Copyright © Samuel O. Ortiz, Ph.D. and may not be reproduced without permission.

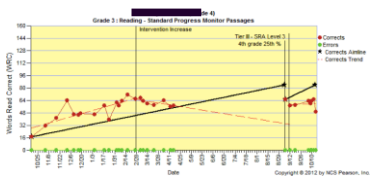
Let's talk about Comprehension and Knowledge (Gc)

Rule of thumb: CLD students typically perform one standard deviation or more below the mean of the norm group on Gc tasks (i.e. Verbal Comprehension, Vocabulary, Word Reasoning, Oral Expression).

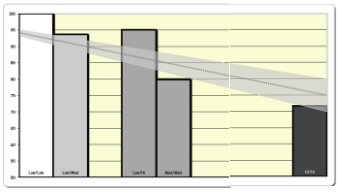
Crystallized Intelligence/Comprehension and Knowledge (Gc) is unique because it directly measures culture and language – impossible to separate the influence

If performance of CLD students in the area of Gc (lower right hand matrix cell) is a standard score of 80-85, then equitable interpretation of test scores are deemed to be valid (and Gc cannot be interpreted as a weakness even thought test interpretation deems it "Below Average"

Case Study: Howie:Tier III Data



Howie: C-LIM Graph



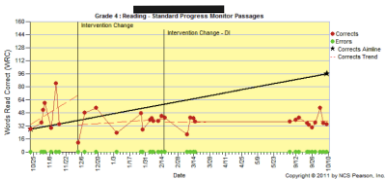
Language Testing

Howie	Core Language	Receptive Language	Expressive Language	Language Memory
CELF-English	75	79	75	75
CELF-Spanish	93	100	87	97

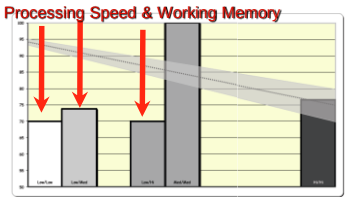
Eligibility Case: Juan

- Student: Juan (5th grade)
- RtI since 3rd grade - Tier III for 2 years
- Academic difficulties in reading, writing, and math
- Also dx ADHD, takes medication for symptom management

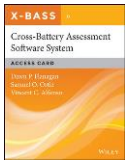
Juan's Tier III Data

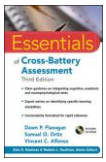


Juan's C-LIM Graph



I love it – I want more!





http://www.wiley.com/WileyCDA/WileyTitle/productCd-111905639X.html

http://www.wiley.com/WileyCDA/WileyTitle/productCd-0470621958.html

References

- Flanagan D.P., Ortiz, S.O., & Alfonso, V.C.; (2013). *Essentials of Cross-Battery Assessment, Third Edition.* Hoboken, New Jersey. John Wiley & Sons.
- Ortiz, S.O. (2011) Power Point Presentation: *Evidence Based Instruction and Evaluation of English Language Learners: Bridging research and practice to promote fairness and equity in assessment.* North Carolina School Psychologist Association Conference, Winston Salem, North Carolina.

GRADUATION: The Measure of Tomorrow
65TH CONFERENCE ON EXCEPTIONAL CHILDREN

GRADUATION: The Measure of Tomorrow
65TH CONFERENCE ON EXCEPTIONAL CHILDREN

Sample Interpretive Statements for use with the Culture-Language Interpretive Matrix (C-LIM)

Listed below are some sample validity statements that provide a general framework for how results from analysis with the C-LIM may be worded and prepared for inclusion in an evaluation report. There are four statements that may apply to four different evaluation scenarios. Statement 1 may be used in cases conducted for the purpose of suspected learning disability and where use of the C-LIM has resulted in a clear declining pattern that merits declaration of the scores as being invalid due to the primary influence of cultural and linguistic variables on test performance. Statement 2 is also written for cases conducted for the purpose of evaluating suspected learning disability but where the results have been declared valid and the results point to limited difficulties such as a learning disability. The third and fourth statements are similar to the second and apply in cases where the results have been declared to be invalid but the manifest patterns are more consistent with significant types of dysfunction including global cognitive impairment (Statement 3) and speech-language impairment (Statement 4).

The statements in this Word file are Copyright © 2013, Samuel O. Ortiz but have been released in the public domain to assist practitioners in describing their findings and designed for use in report writing. Therefore permission is granted to freely use and copy, in edited or un-edited form, any and all of the following statements to suit the user's needs as long as the purpose is educational and not for profit. In addition, it is not necessary to display any copyright statement and there is no need to provide any reference or citation the original source. These statements are offered only as guide and framework for crafting appropriate descriptive statements following analysis of the impact of cultural and linguistic factors on test performance. The statements can be easily personalized by changing the words, "the student" to the examinee's actual name. However, as no statement can apply specifically to all cases in all circumstances, users are cautioned to ensure that any necessary edits or modifications are utilized to ensure the accuracy of these statements as they may apply to any specific evaluation.

Statement 1. Evaluations of Suspected Learning Disability - Invalid Results

The following sample validity statement may be used in cases where a clear declining pattern is evident, that is, that there is a primary effect of culture and language thus results are NOT valid and there is no disability.

Because the student is not a native English speaker, it is necessary to establish the validity of the results obtained from testing to ensure that they are accurate estimates of ability or knowledge and not the manifestation of cultural or linguistic differences. To this end, a systematic evaluation of the possible effects of lack of acculturation and limited English proficiency was carried out via use of the Culture-Language Interpretive Matrix (C-LIM).

A careful review of the student's test data, as entered into the C-LIM, reveals a pattern of decline that is typical of and within the range that would be expected of other individuals with similar cultural and linguistic backgrounds. This overall, declining pattern of test performance suggests that test performance was due primarily to the influence of cultural and linguistic factors rather than lack of actual ability. Accordingly, the test results evaluated here cannot be considered valid and are not interpretable from a strict psychometric standpoint. However, given that the observed pattern is consistent with research-based performance that is typical of non-disabled, culturally and linguistically diverse individuals who are of average ability or higher, it can be reasonably concluded that the student's abilities are also within the average range of performance (or possibly higher) and does not suggest or support the presence of any type of disability.

In summary, the observed pattern of the student's test results is consistent with performance that is typical of non-disabled, culturally and linguistically diverse individuals who are of average ability or higher. Therefore, it can be reasonably concluded that the data evaluated with the C-LIM are invalid due to the presence of overarching cultural and linguistic influences and that the student's test performance and cannot be used to support the presence of any type of learning disability or other cognitive-based disorder.

Statement 2. Evaluations of Suspected Learning Disability - Valid Results

The following sample validity statement may be used in cases where a clear declining pattern is NOT evident, that is, there is no primary effect of culture and language thus the results ARE valid and there may be a disability.

Because the student is not a native English speaker, it is necessary to establish the validity of the results obtained from testing to ensure that they are accurate estimates of ability or knowledge and not the manifestation of cultural or linguistic differences. To this end, a systematic evaluation of the possible effects of lack of acculturation and limited English proficiency was carried out via use of the Culture-Language Interpretive Matrix (C-LIM).

A careful review of the student's test data as entered into the C-LIM does not appear to reveal a pattern of decline that is typical of or within the range that would be expected of other individuals with similar cultural and linguistic backgrounds. The overall pattern of test performance does not decline systematically and suggests that test performance was not due primarily to the influence of cultural and linguistic factors. Although such influences remain contributory factors, they cannot account for the resulting pattern of performance in its entirety and are, therefore, not believed to be the main or only reason for the reported learning difficulties. In addition, other extraneous factors that might account for the observed pattern (for example, lack of motivation, fatigue, incorrect administration/scoring, emotional/behavioral problems) have been excluded. This indicates that the test results can be considered valid, interpretable, and are likely to be good estimates of the student's actual ability or knowledge with the exception of Gc, which must be evaluated only against other ELLs due to the fact that it is a direct measure of cultural knowledge and language proficiency.

In summary, the observed pattern of the student's test results is not consistent with performance that is typical of non-disabled, culturally and linguistically diverse individuals who are of average ability or higher. Therefore, it can be reasonably concluded that the data evaluated with the C-LIM are likely valid and that, if supported by additional data, the student's test performance may be attributed primarily to the presence of a learning disability.

*(←*Note: a description of the data that support the presence of LD should follow here at this point in the report.)*

Statement 3. Evaluation of Global Cognitive Impairment - Valid Results

The following sample validity statement may be used in cases where although a declining pattern may be evident suggesting a primary influence of cultural and linguistic factors, the overall magnitude of the scores are well below the expected range and appear uniformly depressed.

Because the student is not a native English speaker, it is necessary to establish the validity of the results obtained from testing to ensure that they are accurate estimates of ability or knowledge and not the manifestation of cultural or linguistic differences. To this end, a systematic evaluation of the possible effects of lack of acculturation and limited English proficiency was carried out via use of the Culture-Language Interpretive Matrix (C-LIM).

A careful review of the student's test data as entered into the C-LIM indicates that one of the primary influences on testing was likely due to cultural and linguistic differences. Ordinarily, this might invalidate the results. However, in this case, the observed range of scores reveals an overall pattern that is considerably below what might generally be expected from individuals of similar cultural and linguistic backgrounds who possess average general cognitive ability. Therefore, although the results demonstrate a systematic decline as the demands of the tasks increase in terms of cultural loading and linguistic demand, the obtained values across the board are at least about a full standard deviation below expected performance. This pattern suggests the presence of a broad, pervasive influence that, in addition to the cultural and linguistic ones, is primarily responsible for lowering overall performance more or less uniformly across the entire range of functioning. Such a finding could include the possibility of a significant cognitive-based disability.

In summary, the observed pattern of the student's test results is not consistent with performance that is typical of non-disabled, culturally and linguistically diverse individuals who are of average ability or higher. Although the overall pattern of results in this case does decline, the results appear to be valid because the magnitude of the scores are much lower than what would be expected and indicate the presence of another influence. Therefore, it can be reasonably concluded that, if supported by additional data, the student's test performance may be attributed to some type of global cognitive impairment and intellectual functioning is at a level that could be considered significantly sub-average as compared to same age peers with similar cultural and linguistic backgrounds.

*(←*Note: a typical description of the data that support the presence of a global cognitive deficit or MR should follow here at this point in the report.)*

Statement 4. Evaluation of Speech-Language Impairment - Valid Results

The following sample validity statement may be used in cases where although a declining pattern is evident, the rate and severity of the decline is significantly more rapid and steeper than the expected range and cannot be attributed to cultural and linguistic factors only.

Because the student is not a native English speaker, it is necessary to establish the validity of the results obtained from testing to ensure that they are accurate estimates of ability or knowledge and not the manifestation of cultural or linguistic differences. To this end, a systematic

evaluation of the possible effects of lack of acculturation and limited English proficiency was carried out via use of the Culture-Language Interpretive Matrix (C-LIM).

A careful review of the student's test data as entered into the C-LIM indicates that one of the primary influences on testing was likely due to linguistic differences (not necessarily cultural differences). Ordinarily, this might invalidate the results. However, in this case, the observed pattern of scores reveals that performance on tests tends to decrease primarily as a function of the increasing demands of language and not as much in relation to increased cultural content. In addition, the rate of the decline in performance is much more rapid and severe as the demands of the tasks increase in terms of language as compared to the decline that would generally be expected from individuals of similar cultural and linguistic backgrounds who are of average ability or higher. This pattern suggests the presence of an additional factor, most likely related to language skills, that, in addition to the cultural and linguistic influences that remain contributory, is primarily responsible for lowering performance on these tasks to a degree that cannot be accounted for or attributed solely to language or cultural differences. This finding includes the possibility of a language-related disorder being present.

In summary, the observed pattern of the student's test results is not consistent with performance that is typical of non-disabled, culturally and linguistically diverse individuals who are of average ability or higher. Although the overall pattern of results in this case does decline, the results appear to be valid because the rate and magnitude of the decline is more rapid and severe than what would be expected and indicate the presence of another influence. Therefore, it can be reasonably concluded that, if supported by additional data, the student's test performance may be attributed to some type of language-related learning disability or speech-language disorder and performance in the domain of language, even when viewed within the context of the student's cultural and linguistic background and experiences, is a level that is significantly below that of same age peers with similar backgrounds.

*(←*Note: a description of the data that support the presence of a speech-language disorder/language-based LD should follow here at this point in the report.)*

Make sure to substitute the name of the student who is being evaluated in place of where I have written "the student's" to make the report more personalized. Also, if you use the statement above pertaining to LD/ID/SLI, where the results are deemed valid, you should then proceed to using the XBA DMIA and provide an analysis and interpretation of those results as they might pertain to the presence of a learning disability just as you would for any other student (with the noted exception of Gc).