

### Dual Identification of English Learners: Use of a PSW model for determining SLD with ELs.



COSA 2016 State English Learners Alliance Conference  
March 10, 2016  
Eugene, OR

Samuel O. Ortiz, Ph.D.  
St. John's University

---

---

---

---

---

---

---

---

---

---

### IDEA Regulations 34 CFR 300.307(a) Procedures for Identification of Specific Learning Disability

A State must adopt, consistent with 34 CFR 300.309, criteria for determining whether a child has a specific learning disability as defined in 34 CFR 300.8(c)(10). In addition, the criteria adopted by the State:

- Must not require the use of a severe discrepancy between intellectual ability and achievement for determining whether a child has a specific learning disability, as defined in 34 CFR 300.8(c)(10);
- Must permit the use of a process based on the child's response to scientific, research-based intervention; and
- *May permit the use of other alternative research-based procedures for determining whether a child has a specific learning disability, as defined in 34 CFR 300.8(c)(10).*

This includes use of approaches based on a Pattern of Strengths and Weaknesses (PSW)

---

---

---

---

---

---

---

---

---

---

### IDEA Regulations 34 CFR 300.307(a) Procedures for Identification of Specific Learning Disability

The group described in 34 CFR 300.306 may determine that a child has a specific learning disability, as defined in 34 CFR 300.8(c)(10), if...

- The child does not make sufficient progress to meet age or State-approved grade-level standards in one or more of the areas identified in 34 CFR 300.309(a)(1) when using a process based on the child's response to scientific, research-based intervention, or the child exhibits a pattern of strengths and weaknesses in performance, achievement, or both, relative to age, State-approved grade-level standards, or intellectual development, that is determined by the group to be relevant to the identification of a specific learning disability, using appropriate assessments, consistent with 34 CFR 300.304 and 300.305; and the group determines that its findings under 34 CFR 300.309(a)(1) and (2) are not primarily the result of:
  - A visual, hearing, or motor disability;
  - Mental retardation;
  - *Emotional disturbance;*
  - Cultural factors;
  - Environmental or economic disadvantage; or
  - Limited English proficiency.

Recognizes that lack of English proficiency or cultural difference cannot be the basis of a disability and cannot be the primary reason for observed academic problems.

Source: IDEA Statute and Regulations. Last retrieved on Feb. 5, 2016 from <http://www.wrightslaw.com/idea/law/section1414.pdf>

---

---

---

---

---

---

---

---

---

---

**20 U.S.C. 1414 Evaluations, Eligibility Determinations, Individualized Education Programs, and Educational Placements**

(b) EVALUATION PROCEDURES –

(3) ADDITIONAL REQUIREMENTS- Each local educational agency shall ensure that—

(A) ~~tests~~ assessments and other evaluation materials used to assess a child under this section—

(i) are selected and administered so as not to be discriminatory on a racial or cultural basis;

(ii) are provided and administered in the ~~child's native language or other mode of communication~~ language and form most likely to yield accurate information on what the child knows and can do academically, developmentally, and functionally, unless it is not feasible to so provide or administer;

(iii) are used for purposes for which the assessments or measures are valid and reliable;

Recognizes that validity is not automatically assured via native language testing.

Source: IDEA Statute and Regulations. Last retrieved on Feb. 5, 2016 from <http://www.wrightslaw.com/idea/law/section1414.pdf>

---

---

---

---

---

---

---

---

---

---

---

---

**General 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 potential bias in traditional assessment practices
- X. Support conclusions via data convergence and multiple indicators

Addresses concerns regarding fairness and equity in the assessment process

Addresses possible bias in use of test scores

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

---

---

---

---

---

---

---

---

---

---

---

---

**Main Threats to Test Score Validity for ELLs**

**NO BIAS**

- **Test items**  
(content, novelty)
- **Test structure**  
(sequence, order, difficulty)
- **Test reliability**  
(measurement error/accuracy)
- **Factor structure**  
(theoretical structure, relationship of variables to each other)
- **Predictive Validity**  
(correlation with academic success or achievement)

**BIAS**

- **Construct Validity**  
(nature and specificity of the intended/measured constructs)

When a test measures an unintended variable...

- **Incorrect Interpretation**  
(undermines accuracy of evaluative judgments and meaning assigned to scores)

"As long as tests do not at least sample in equal degree a state of saturation [assimilation of fundamental experiences and activities] that is equal for the 'norm children' and the particular bilingual child it cannot be assumed that the test is a valid one for the child."  
 Sanchez, 1934

---

---

---

---

---

---

---

---

---

---

---

---

### Main Threats to Test Score Validity for ELLs

**Acculturative Knowledge Acquisition – Not Race or Ethnicity**

*“When a child’s general background experiences differ from those of the children on whom a test was standardized, then the use of the norms of that test as an index for evaluating that child’s current performance or for predicting future performances may be inappropriate.”*

*Salvia & Ysseldyke, 1991*

**Developmental Language Proficiency – Not Language Dominance**

*“Most studies compare the performance of students from different ethnic groups...rather than ELL and non-ELL children within those ethnic groups...A major difficulty with all of these studies is that the category Hispanic includes students from diverse cultural backgrounds with markedly different English-language skills...This reinforces the need to separate the influences of ethnicity and ELL status on observed score differences.”*

*Lohman, Korb & Lakin, 2008*

---

---

---

---

---

---

---

---

---

---

### Processes and Procedures for Addressing Test Score Validity

**IX. REDUCE BIAS IN TRADITIONAL TESTING PRACTICES**

*Exactly how is evidence-based, nondiscriminatory assessment conducted and to what extent is there any research to support the use of any of these methods in being capable of establishing sufficient validity of the obtained results?*

- **Modified Methods of Evaluation**
  - Modified and altered assessment
- **Nonverbal Methods of Evaluation**
  - Language reduced assessment
- **Dominant Language Evaluation: L1**
  - Native language assessment
- **Dominant Language Evaluation: L2**
  - English language assessment

---

---

---

---

---

---

---

---

---

---

### Comparison of Methods for Addressing Main Threats to Validity

Evaluation Method	Norm sample representative of bilingual development	Measures full range of ability constructs	Does not require bilingual evaluator	Adheres to the test’s standardized protocol	Substantial research base on bilingual performance
Modified or Altered Assessment	✗	✓	✓	✗	✗
Reduced-language Assessment	✗	✗	✓	✓	✗
Dominant Language Assessment in L1: native	✗	✓	✗	✓	✗
Dominant Language Assessment in L2: English	✗	✓	✓	✓	✓

Addressing issues of fairness with respect to norm sample representation is an issue of validity and dependent on a sufficient research base.

---

---

---

---

---

---

---

---

---

---

**Evaluating and Defending Construct ELL Test Score Validity**

Whatever method or approach may be employed in evaluation of ELL's, the fundamental obstacle to nondiscriminatory interpretation rests on the degree to which the examiner is able to defend claims of test score construct validity. This is captured by and commonly referred to as a question of:

**"DIFFERENCE vs. DISORDER?"**

Simply absolving oneself from responsibility of doing so via wording such as, "all scores should be interpreted with extreme caution" does not in any way provide a defensible argument regarding the validity of obtained test results and does not permit interpretation.

At present, the only manner in which test score validity can be evaluated or established is via use of the existing research on the test performance of ELLs as reflected in the degree of "difference" the student displays relative to the norm samples of the tests being used, particularly for tests in English. This is the sole purpose of the C-LIM.

---

---

---

---

---

---

---

---

---

---

**Summary of Research on the Test Performance of English Language Learners**

Research conducted over the past 100 years on ELLs who are non-disabled, of average ability, possess moderate to high proficiency in English, and tested in English, has resulted in two robust and ubiquitous findings:

1. Native English speakers perform better than English learners at the broad ability level (e.g., FSIQ) on standardized, norm-referenced tests of intelligence and general cognitive ability.
2. English learners tend to perform significantly better on nonverbal type tests than they do on verbal tests (e.g., PIQ vs. VIQ).

So what explains these findings? Early explanations relied on genetic differences attributed to race even when data strongly indicated that the test performance of ELLs was moderated by the degree to which a given test relied on or required age- or grade-expected development in English and the acquisition of incidental acculturative knowledge.

---

---

---

---

---

---

---

---

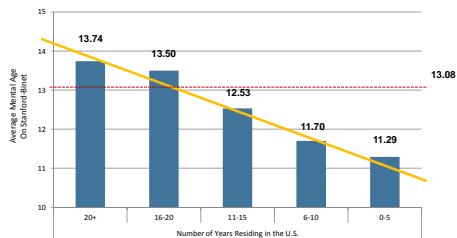
---

---

**Research Foundations for ELL Evaluation**

Principle 1: ELLs and non-ELL's perform differently at the broad ability level

Mean Mental Age (MA) from Binet Scales in a non-native English speaking sample from Yerkes' (1921) data as analyzed by C.C. Brigham (1923)



Average score for native English speakers on Beta = 101.6 (Very Superior; Grade A)  
Average score for non-native English speakers on Beta = 77.8 (Average; Grade C)

---

---

---

---

---

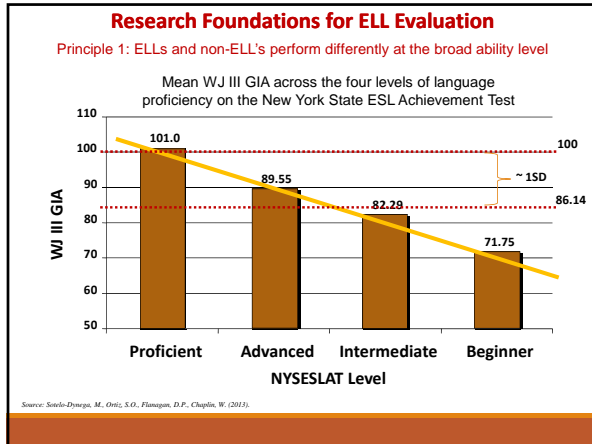
---

---

---

---

---




---

---

---

---

---

---

---

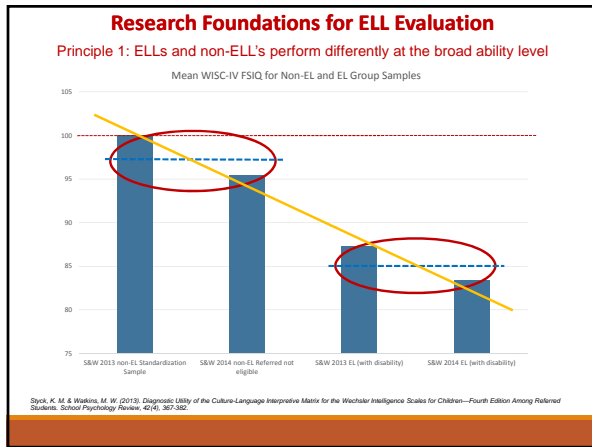
---

---

---

---

---




---

---

---

---

---

---

---

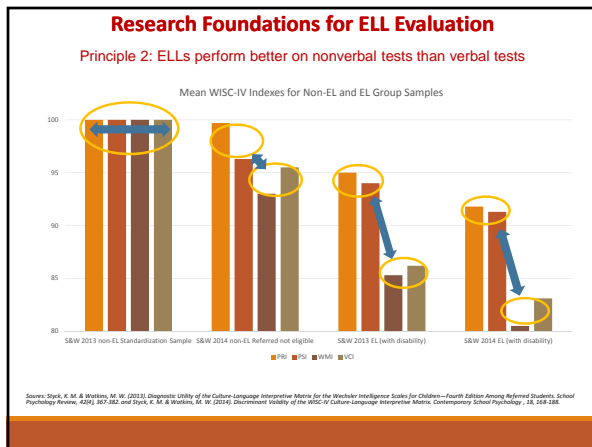
---

---

---

---

---




---

---

---

---

---

---

---

---

---

---

---

---

**Research Foundations for ELL Evaluation**

Historical and contemporary research has tended to ignore the fact that ELLs do not perform at the same level on ALL nonverbal tests any more than they perform at the same level on ALL verbal tests.

Instead, it appears that test performance of ELLs is not a dichotomy but rather a continuum formed by a linear, not dichotomous, attenuation of performance.

This means, a third principle is evident in the body of research on ELLs but has not been well understood or utilized in understanding test performance:

*3. Test performance of ELLs is moderated by the degree to which a given test relies on or requires age- or grade-expected English language development and the acquisition of incidental acculturative knowledge.*

---

---

---

---

---

---

---

---

---

---

**Research Foundations for ELL Evaluation**

ELL test performance is a linear, continuous pattern, not a dichotomy.

Cultural Loading and Linguistic Demand

Low

Moderate

High

Subtests can be arranged from high to low in accordance with the mean values reported by empirical studies for ELLs

SS = 100

95

90

85

80

→

Tests requiring lower levels of age/grade related acquisition of culture and language result in higher mean scores

Tests requiring higher levels of age/grade related acquisition of culture and language result in lower mean scores

---

---

---

---

---

---

---

---

---

---

**Research Foundations for ELL Evaluation**

Principle 3: ELL performance is moderated by linguistic/acculturative variables

Subtest Name	Hispanic Group (Mercer) (1972)	Hispanic Group (Vukovich & Figueroa) (1982)	ESL Group (Cummins) (1982)	Bilingual Group (Nieves-Brull) (2006)
Information	7.5	7.8	5.1	7.2
Vocabulary	8.0	8.3	6.1	7.5
Similarities	7.6	8.8	6.4	8.2
Comprehension	7.8	9.0	6.7	8.0
Digit Span	8.3	8.5	7.3	*
Arithmetic	8.7	9.4	7.4	7.8
Picture Arrangement	9.0	10.3	8.0	9.2
Block Design	9.5	10.8	8.0	9.4
Object Assembly	9.6	10.7	8.4	9.3
Picture Completion	9.7	9.9	8.7	9.5
Coding	9.6	10.9	8.9	9.6

\*Data for this subtest were not reported in the study.

---

---

---

---

---

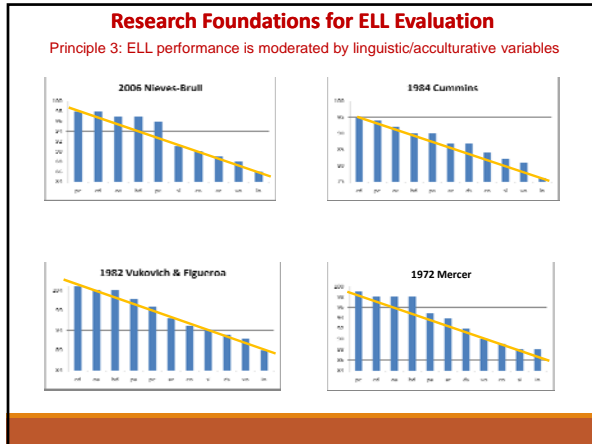
---

---

---

---

---




---

---

---

---

---

---

---

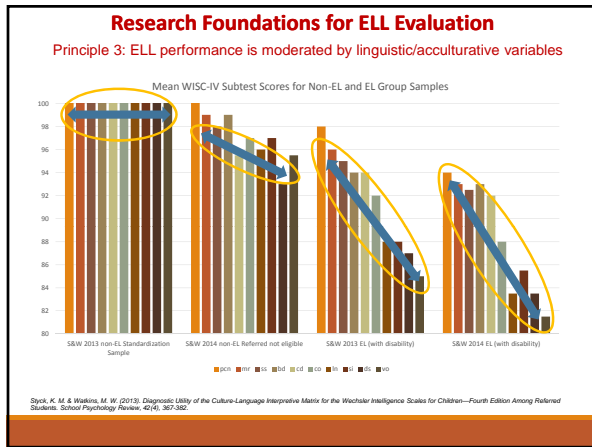
---

---

---

---

---




---

---

---

---

---

---

---

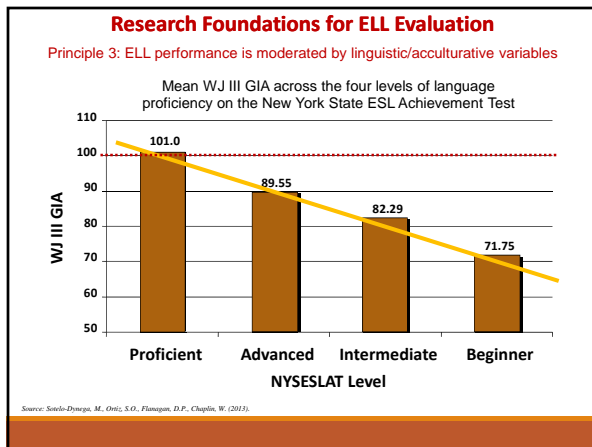
---

---

---

---

---




---

---

---

---

---

---

---

---

---

---

---

---

**Research Foundations for ELL Evaluation**

Principle 3: ELL performance is moderated by linguistic/acclulturative variables

Table 3. Variance Explained by Exogenous Variables (Individual Test Performance) by Age Group.

Individual test	Variance explained		
	7-10	11-14	15-18
Verbal Comprehension	.79 <sup>a</sup>	.86 <sup>a</sup>	.81 <sup>a</sup>
General Information	.71 <sup>a</sup>	.85 <sup>a</sup>	.86 <sup>a</sup>
Concept Formation	.67 <sup>a</sup>	.71 <sup>a</sup>	.67 <sup>a</sup>
Visual-Auditory Learning	.40 <sup>a</sup>	.37 <sup>a</sup>	.41 <sup>a</sup>
Delayed Recall Visual-Auditory Learning	.39 <sup>a</sup>	.32 <sup>a</sup>	.37 <sup>a</sup>
Analysis Synthesis	.29 <sup>a</sup>	.44 <sup>a</sup>	.47 <sup>a</sup>
Sound Blending	.25 <sup>a</sup>	.32 <sup>a</sup>	.35 <sup>a</sup>
Auditory Working Memory	.22 <sup>a</sup>	.44 <sup>a</sup>	.32 <sup>a</sup>
Retrieval Fluency	.22 <sup>a</sup>	.22 <sup>a</sup>	.28 <sup>a</sup>
Memory for Words	.18 <sup>a</sup>	.32 <sup>a</sup>	.22 <sup>a</sup>
Numbers Reversed	.17 <sup>a</sup>	.26 <sup>a</sup>	.30 <sup>a</sup>
Pair Cancellation	.17 <sup>a</sup>	.11 <sup>a</sup>	.11 <sup>a</sup>
Rapid Picture Naming	.16 <sup>a</sup>	.07 <sup>a</sup>	.16 <sup>a</sup>
Incomplete Words	.13 <sup>a</sup>	.31 <sup>a</sup>	.23 <sup>a</sup>
Visual Matching	.13 <sup>a</sup>	.15 <sup>a</sup>	.16 <sup>a</sup>
Decision Speed	.12 <sup>a</sup>	.15 <sup>a</sup>	.19 <sup>a</sup>
Auditory Attention	.10 <sup>a</sup>	.20 <sup>a</sup>	.15 <sup>a</sup>
Spatial Relations	.08 <sup>a</sup>	.16 <sup>a</sup>	.16 <sup>a</sup>
Planning	.07 <sup>a</sup>	.12 <sup>a</sup>	.11 <sup>a</sup>
Picture Recall	.02 <sup>a</sup>	.06 <sup>a</sup>	.10 <sup>a</sup>

\*Source: Comier, D.C., McGee, K.S. & Yuskis, J.E. (2014). The Influence of Linguistic Demand and Cultural Loading on Cognitive Test Scores. Journal of Psychoeducational Assessment, 22(7), 650-622.

---

---

---

---

---

---

---

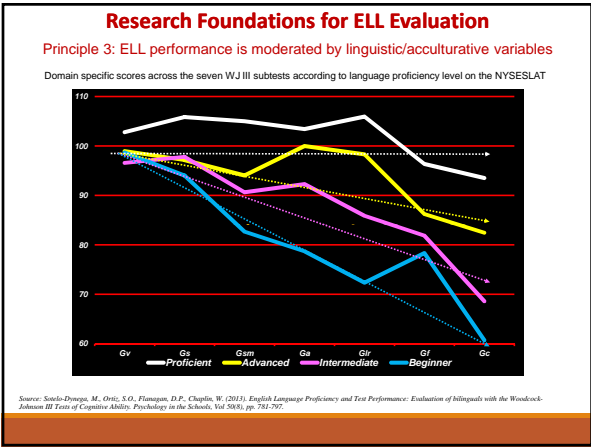
---

---

---

---

---




---

---

---

---

---

---

---

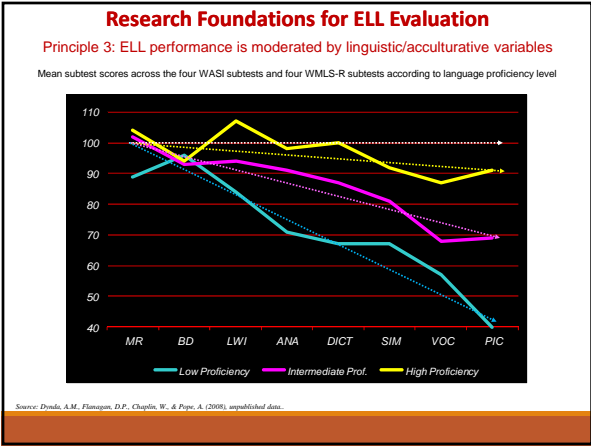
---

---

---

---

---




---

---

---

---

---

---

---

---

---

---

---

---



**Foundational Research Principles of the Culture-Language Interpretive Matrix**

- Principle 1: EL and non-ELs perform differently at the broad ability level on tests of cognitive ability.*
- Principle 2: ELs perform better on nonverbal tests than they do on verbal tests.*
- Principle 3: EL performance on both verbal and nonverbal tests is moderated by linguistic and acculturative variables.*

Because the basic research principles underlying the C-LIM are well supported, their operationalization within the C-LIM provides a substantive evidentiary base for evaluating the test performance of English language learners.

- This does not mean, however, that it cannot be improved. Productive research on EL test performance can assist in making any necessary "adjustments" to the order of the means as arranged in the C-LIM.
- Likewise, as new tests come out, new research is needed to determine the relative level of EL performance as compared to other tests with established values of expected average performance.
- Ultimately, only research that focuses on stratifying samples by relevant variables such as language proficiency, length and type of English and native language instruction, and developmental issues related to age and grade of first exposure to English, will serve useful in furthering knowledge in this area and assist in establishing appropriate expectations of test performance for specific populations of ELs.

---

---

---

---

---

---

---

---

---

---

---

---

**Practical Considerations for Addressing Validity in Evaluation Procedures for SLD with ELLs**

1. The usual purpose of testing is to identify deficits in ability (i.e., low scores)
2. Validity is more of a concern for low scores than average/higher scores because:
  - Test performances in the average range are NOT likely a chance finding and strongly suggests average ability (i.e., no deficits in ability)
  - Test performances that are below average MAY be a chance finding because of experiential or developmental differences and thus do not automatically confirm below average ability (i.e., possible deficits in ability)
3. Therefore, testing in one language only (English or native language) means that:
  - It can be determined that a student DOES NOT have a disability (i.e., if all scores are average or higher, they are very likely to be valid)
  - It CANNOT be determined if the student has a disability (i.e., low scores must be validated as true indicators of deficit ability)
4. Testing in both languages (English and native language) is necessary to determine disability
  - Testing requires confirmation that deficits are not language-specific and exist in both languages (although low performance in both can result from other factors)
5. All low test scores, whether in English or the native language, must be validated
  - Low scores from testing in English can be validated via research underlying the C-LIM
  - Low scores from testing in the native language cannot be validated with research

---

---

---

---

---

---

---

---

---

---

---

---

**Practical Considerations for Addressing Validity in Evaluation Procedures for SLD with ELLs**

Given the preceding considerations, the most practical and defensible general approach in evaluating ELLs would be:

- Test in English first and if all test scores indicate strengths (average or higher) a disability is not likely and thus no further testing is necessary
- If some scores from testing in English indicate weaknesses, re-test those areas in the native language to cross-validate as areas of true weakness

This approach provides the most efficient process and best use of available resources for evaluation since it permits ANY evaluator to begin and sometimes complete the testing without being bilingual or requiring assistance.

In addition, this approach is IDEA compliant and consistent with the specification that assessments "be provided and administered in the language and form most likely to yield accurate information" because it relies on an established body of research to guide examination of test score validity and ensures that that the results upon which decisions are based are in fact accurate.

---

---

---

---

---

---

---

---

---

---

---

---

### A Recommended Best Practice Approach for Using Tests with ELLs

**Step 1. Assessment of Bilinguals – validate all areas of performance (exclusion of cultural/linguistic factors)**

- Select or create an appropriate battery that is comprehensive and responds to the needs of the referral concerns, irrespective of language differences
- Administer all tests in standardized manner first in English only with no modifications
- Score tests and plot them for analysis via the C-LIM
- If analysis indicates expected range and pattern of decline, scores are invalid due to cultural and linguistic factors that cannot be excluded as primary reason for poor academic performance
- If analysis does not indicate expected range or pattern of decline, apply XBA (or other) interpretive methods to determine specific areas of weakness and difficulty and continue to Step 2

**Step 2. Bilingual Assessment – validate suspected areas of weakness (cross-language confirmation of deficit areas)**

- Review results and identify areas of suspected weakness or difficulty:
  - a. For Gc only, evaluate weakness according to high/high cell in C-LIM or in context of other data and information
  - b. For all other abilities, evaluate weakness using standard classifications (e.g., SS < 90)
- Except for Gc, re-test all other areas of suspected weakness using native language tests
- For Gc only:
  - a. If the high/high cell in C-LIM is within/above expected range, consider Gc a strength and assume it is at least average, thus re-testing is not necessary
  - b. If the high/high cell in C-LIM is below expected range, re-testing of Gc in the native language is recommended
- Administer native language tests or conduct re-testing using one of the following methods:
  - a. Native language test administered in the native language (e.g., WJ III/Bateria III or WISC-IV/WISC-IV Spanish)
  - b. Native language test administered via assistance of a trained interpreter
  - c. English language test translated and administered via assistance of a trained interpreter
- Administer tests in manner necessary to ensure full comprehension including use of any modifications and alterations necessary to reduce barriers to performance, while documenting approach to tasks, errors in responding, and behavior during testing, and analyze scores both quantitatively and qualitatively to confirm and validate areas as true weaknesses
- Except for Gc, if a score obtained in the native language validates/confirms a weakness score obtained in English (both SS < 90), use/interpret the score obtained in English as a weakness
- If a score obtained in the native language invalidates/discards a weakness score obtained in English (native SS ≥ 90), consider it as a strength and assume that it is at least in the average range
- Scores for Gc obtained in the native language and in English can only be interpreted relative to developmental and educational experiences of the examinee in each language and only as compared to others with similar developmental experiences

---

---

---

---

---

---

---

---

---

---

### The Culture-Language Interpretive Matrix (C-LIM) Addressing test score validity for ELLs

*Translation of Research into Practice*

1. The use of various traditional methods for evaluating ELLs, including testing in the dominant language, modified testing, nonverbal testing, or testing in the native language do not ensure valid results and provide no mechanism for determining whether results are valid, let alone what they might mean or signify.
2. The pattern of ELL test performance, when tests are administered in English, has been established by research and is predictable and based on the examinee's degree of English language proficiency and acculturative experiences/opportunities as compared to native English speakers.
3. The use of research on ELL test performance, when tests are administered in English, provides the only current method for applying evidence to determine the extent to which obtained results are valid (a minimal or only contributory influence of cultural and linguistic factors), possibly valid (minimal or contributory influence of cultural and linguistic factors but which requires additional evidence from native language evaluation), or invalid (a primary influence of cultural and linguistic factors).
4. The principles of ELL test performance as established by research are the foundations upon which the C-LIM is based and serve as a de facto norm sample for the purposes of comparing test results of individual ELLs to the performance of a group of average ELLs with a specific focus on the attenuating influence of cultural and linguistic factors.

---

---

---

---

---

---

---

---

---

---

### Application of Research as Foundations for the Cultural and Linguistic Classification of Tests and C-LIM

PATTERN OF EXPECTED PERFORMANCE FOR ENGLISH LANGUAGE LEARNERS  
DEGREE OF LINGUISTIC DEMAND

		DEGREE OF LINGUISTIC DEMAND		
		LOW	MODERATE	HIGH
DEGREE OF CULTURAL LOADING	LOW	<b>PERFORMANCE LEAST AFFECTED</b> (MINIMAL OR NO EFFECT OF CULTURE & LANGUAGE DIFFERENCES)		INCREASING EFFECT OF LANGUAGE DIFFERENCE
	MODERATE			
	HIGH	INCREASING EFFECT OF CULTURAL DIFFERENCE		<b>PERFORMANCE MOST AFFECTED</b> (LARGE COMBINED EFFECT OF CULTURE & LANGUAGE DIFFERENCES)

---

---

---

---

---

---

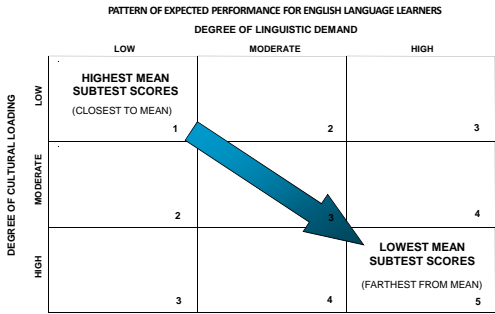
---

---

---

---

### Application of Research as Foundations for the Cultural and Linguistic Classification of Tests and C-LIM




---

---

---

---

---

---

---

---

---

---

### The Culture-Language Interpretive Matrix (C-LIM)

#### Important Considerations for Use and Practice

The C-LIM is not a test, scale, measure, or mechanism for making diagnoses. It is a visual representation of current and previous research on the test performance of English learners arranged by mean values to permit examination of the combined influence of acculturative knowledge acquisition and limited English proficiency and its impact on test score validity.

The C-LIM is not a language proficiency measure and will not distinguish native English speakers from English learners with high, native-like English proficiency and is not designed to determine if someone is or is not an English learner. Moreover, the C-LIM is not for use with individuals who are native English speakers.

The C-LIM is not designed or intended for diagnosing any particular disability but rather as a tool to assist clinician's in making decisions regarding whether ability test scores should be viewed as indications of actual disability or rather a reflection of differences in language proficiency and acculturative knowledge acquisition.

The primary purpose of the C-LIM is to assist evaluators in ruling out cultural and linguistic influences as exclusionary factors that may have undermined the validity of test scores, particularly in evaluations of SLD or other cognitive-based disorders. Being able to make this determination is the primary and main hurdle in evaluation of ELLs and the C-LIM's purpose is to provide an evidence-based method that assists clinician's regarding interpretation of test score data in a nondiscriminatory manner.

---

---

---

---

---

---

---

---

---

---

### The Culture-Language Interpretive Matrix (C-LIM)

#### GENERAL RULES AND GUIDANCE FOR EVALUATION OF TEST SCORE VALIDITY

There are two basic criteria that, when both are met, provide evidence to suggest that test performance reflects the primary influence of cultural and linguistic factors and not actual ability, or lack thereof. These criteria are:

1. *There exists a general, overall pattern of decline in the scores from left to right and diagonally across the matrix where performance is highest on the less linguistically demanding/culturally loaded tests (low/low cells) and performance is lowest on the more linguistically demanding/culturally loaded tests (high/high cells), and;*

2. *The magnitude of the aggregate test scores across the matrix for all cells fall within or above the expected range of difference (shaded area around the line) determined to be most representative of the examinee's background and development relative to the sample on whom the test was normed.*

When both criteria are observed, it may be concluded that the test scores are likely to have been influenced primarily by the presence of cultural/linguistic variables and therefore are not likely to be valid and should not be interpreted.

---

---

---

---

---

---

---

---

---

---

### The Culture-Language Interpretive Matrix (C-LIM)

**RANGE OF POSSIBLE OUTCOMES WHEN EVALUATING TEST SCORES WITHIN C-LIM**

**Condition A:** Overall pattern generally appears to decline across all cells and all cell aggregate scores within or above shaded range—test scores likely **invalid**, cultural/linguistic factors are primary influences, but examinee likely has average/higher ability as data do not support deficits, and further evaluation via testing is unnecessary.

**Condition B:** Overall pattern generally appears to decline across all cells but at least one cell aggregate (or more) is below shaded range—test scores **possibly valid**, cultural/linguistic factors are contributory influences, and further evaluation, including in the native language, is necessary to establish true weaknesses in a given domain.

**Condition C:** Overall pattern does not appear to decline across all cells and all cell aggregate scores within or above average range—test scores **likely valid**, cultural/linguistic factors are minimal influences, and further evaluation may be unnecessary if no weaknesses exist in any domain.

**Condition D:** Overall pattern does not appear to decline across all cells and at least one cell aggregate (or more) is below average range—test scores **possibly valid**, cultural/linguistic factors are minimal influences, and further evaluation, including in the native language, is necessary to establish true weaknesses in a given domain.

---

---

---

---

---

---

---

---

---

---

---

---

### The Culture-Language Interpretive Matrix (C-LIM)

**RANGE OF POSSIBLE OUTCOMES WHEN EVALUATING TEST SCORES WITHIN C-LIM**

	A general, overall pattern of decline exists?	All scores within or above the expected range?	All scores within or above the average range?	Degree of influence of cultural and linguistic factors	Likelihood that test scores are valid indicators of ability?
<b>Condition A</b>	Yes	Yes	No	Primary	Unlikely
<b>Condition B</b>	Yes	No	No	Contributory	Possibly*
<b>Condition C</b>	No	Yes	Yes	Minimal	Likely
<b>Condition D</b>	No	No	No	Minimal	Possibly*

\*Determination regarding the validity of test scores that are below the expected and average ranges requires additional data and information, particularly results from native language evaluation, qualitative evaluation and analysis, and data from a strong pre-referral process (e.g., progress monitoring data).

---

---

---

---

---

---

---

---

---

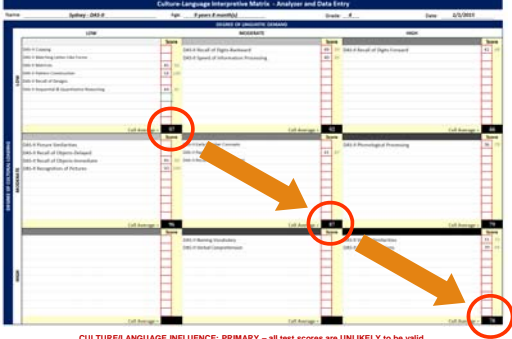
---

---

---

### Culture-Language Interpretive Matrix: Guidelines for evaluating test scores.

**CONDITION A: General declining pattern, all scores within or above expected range.**



**CULTURE/LANGUAGE INFLUENCE: PRIMARY – all test scores are UNLIKELY to be valid.**

---

---

---

---

---

---

---

---

---

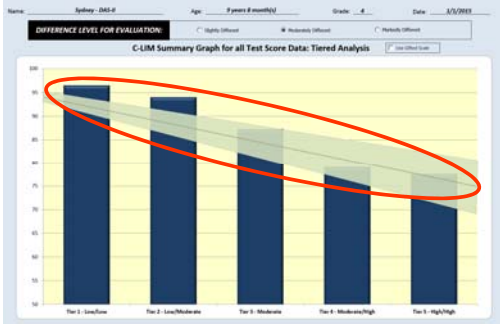
---

---

---

**Culture-Language Interpretive Matrix: Guidelines for evaluating test scores.**

**CONDITION A: General declining pattern, all scores within or above expected range.**



CULTURE/LANGUAGE INFLUENCE: PRIMARY – all test scores are UNLIKELY to be valid.

---

---

---

---

---

---

---

---

---

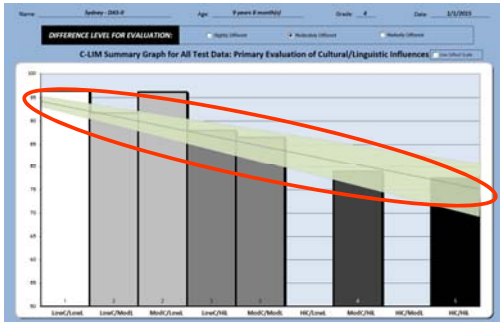
---

---

---

**Culture-Language Interpretive Matrix: Guidelines for evaluating test scores.**

**CONDITION A: General declining pattern, all scores within or above expected range.**



CULTURE/LANGUAGE INFLUENCE: PRIMARY – all test scores are UNLIKELY to be valid.

---

---

---

---

---

---

---

---

---

---

---

---

**Culture-Language Interpretive Matrix: Guidelines for evaluating test scores.**

**CONDITION B: Generally declining pattern, one or more scores below expected range.**

Category	Score	Expected Range
1	95	90-100
2	90	85-95
3	85	80-85
4	80	75-80
5	75	70-75
6	70	65-70
7	65	60-65
8	60	55-60
9	55	50-55
10	50	45-50
11	45	40-45
12	40	35-40
13	35	30-35
14	30	25-30
15	25	20-25
16	20	15-20
17	15	10-15
18	10	5-10
19	5	0-5
20	0	0-5

CULTURE/LANGUAGE INFLUENCE: CONTRIBUTORY – low test scores are POSSIBLY valid.

---

---

---

---

---

---

---

---

---

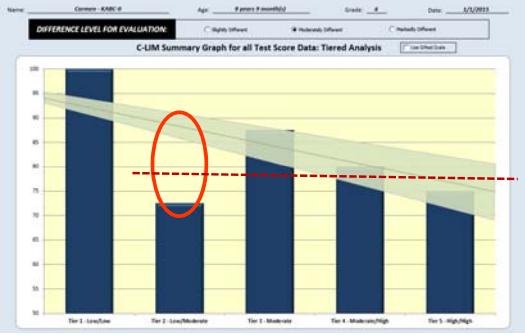
---

---

---

**Culture-Language Interpretive Matrix: Guidelines for evaluating test scores.**

**CONDITION B: Generally declining pattern, one or more scores below expected range.**



**CULTURE/LANGUAGE INFLUENCE: CONTRIBUTORY** – low test scores are POSSIBLY valid.

---

---

---

---

---

---

---

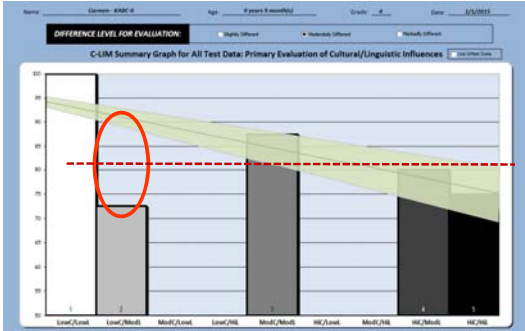
---

---

---

**Culture-Language Interpretive Matrix: Guidelines for evaluating test scores.**

**CONDITION B: Generally declining pattern, one or more scores below expected range.**



**CULTURE/LANGUAGE INFLUENCE: CONTRIBUTORY** – low test scores are POSSIBLY valid.

---

---

---

---

---

---

---

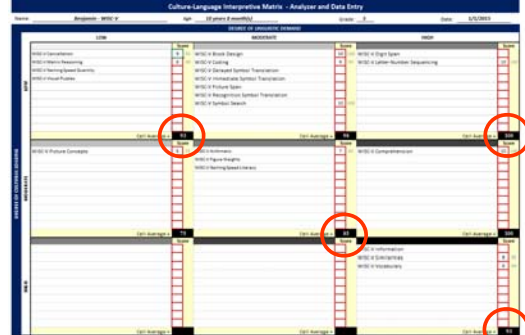
---

---

---

**Culture-Language Interpretive Matrix: Guidelines for evaluating test scores.**

**CONDITION C: No declining pattern, all scores within or above average range.**



**CULTURE/LANGUAGE INFLUENCE: MINIMAL** – all test scores are LIKELY to be valid.

---

---

---

---

---

---

---

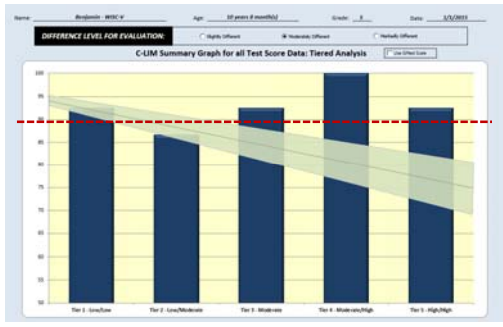
---

---

---

**Culture-Language Interpretive Matrix: Guidelines for evaluating test scores.**

**CONDITION C:** No declining pattern, all scores within or above average range.



CULTURE/LANGUAGE INFLUENCE: MINIMAL – all test scores are LIKELY to be valid.

---

---

---

---

---

---

---

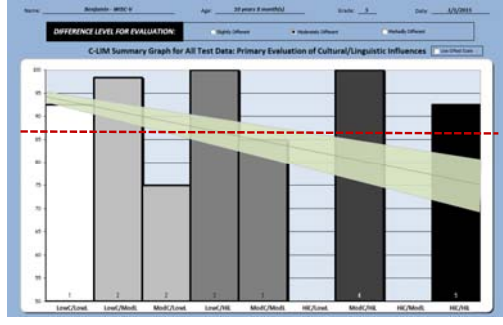
---

---

---

**Culture-Language Interpretive Matrix: Guidelines for evaluating test scores.**

**CONDITION C:** No declining pattern, all scores within or above average range.



CULTURE/LANGUAGE INFLUENCE: MINIMAL – all test scores are LIKELY to be valid.

---

---

---

---

---

---

---

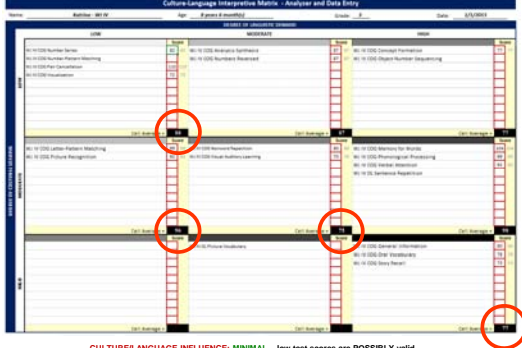
---

---

---

**Culture-Language Interpretive Matrix: Guidelines for evaluating test scores.**

**CONDITION D:** No declining pattern, one or more scores below average range.



CULTURE/LANGUAGE INFLUENCE: MINIMAL – low test scores are POSSIBLY valid.

---

---

---

---

---

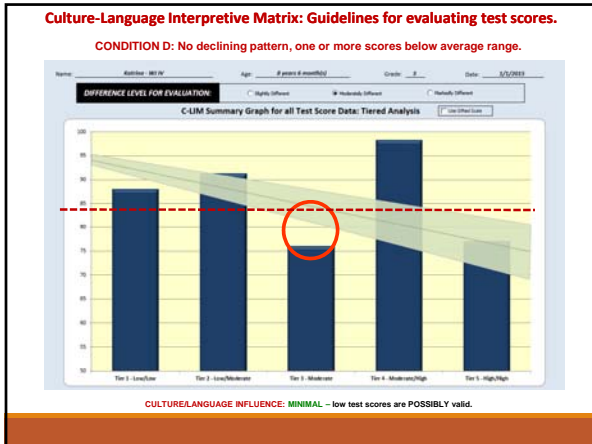
---

---

---

---

---




---

---

---

---

---

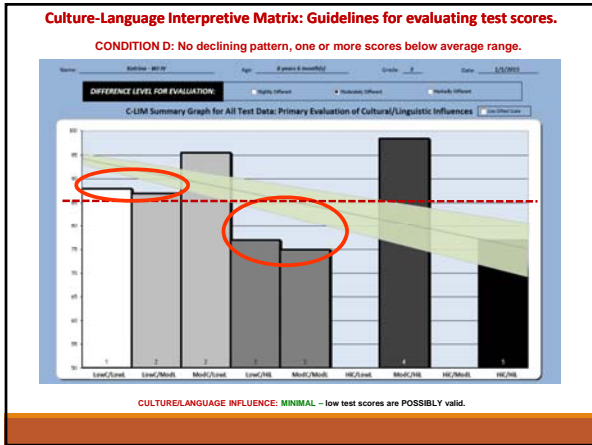
---

---

---

---

---




---

---

---

---

---

---

---

---

---

---

**Culture-Language Interpretive Matrix (C-LIM): Case Study - Elizabeth**

**Woodcock-Johnson IV: Tests of Cognitive Ability (English Administration)**

SS	PR		SS	PR		SS	PR	
Oral Vocabulary	69	2	General Information	79	3	Number Series	96	39
Concept Formation	87	19	Verbal Attention	84	14	Numbers Reversed	92	30
Letter-Pattern Matching	98	45	Pair Cancellation	94	34	Phonological Processing	81	10
Nonword Repetition	91	27	Story Recall	83	13	Visual-Auditory Learning	89	23
Visualization	102	55	Picture Recognition	91	27			

**Wechsler Intelligence Scale for Children - V (English Administration)**

Scaled Score	PR	Standard Score	Scaled Score	PR	Standard Score		
Information	5	9	80	Block Design	9	38	95
Similarities	4	2	70	Matrix Reasoning	10	50	100
Vocabulary	5	2	70	Symbol Search	10	50	100
Comprehension	6	16	85	Coding	8	25	90
Digit Span	9	38	95	Visual Puzzles	10	50	100

**Leiter-3 Nonverbal Intelligence Test (Nonverbal Administration)**

Scaled Score	PR	Standard Score	Scaled Score	PR	Standard Score		
Nonverbal Stroop	9	38	95	Sequential Order	8	25	90
Visual Patterns	9	38	95	Form Completion	8	25	90
Reverse Memory	10	50	100	Classification & Analog.	8	25	90
Figure Ground	8	25	90	Forward Memory	7	16	85

---

---

---

---

---

---

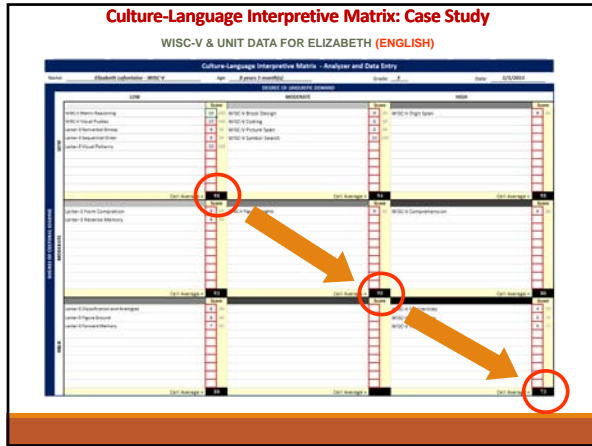
---

---

---

---






---

---

---

---

---

---

---

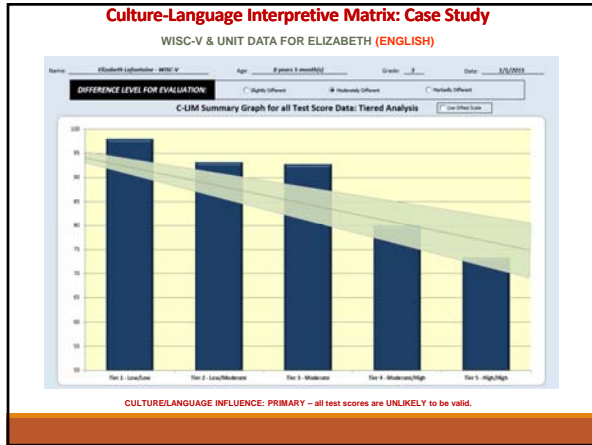
---

---

---

---

---




---

---

---

---

---

---

---

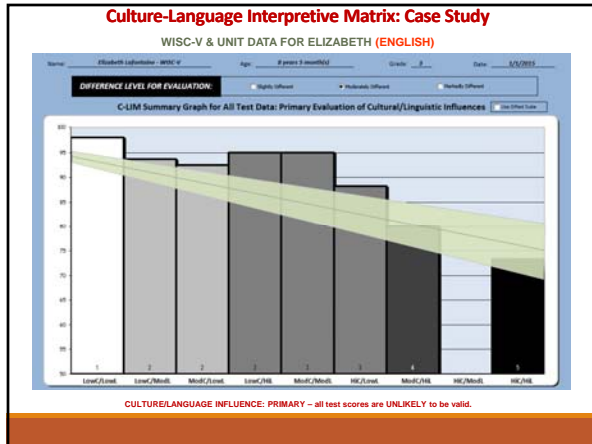
---

---

---

---

---




---

---

---

---

---

---

---

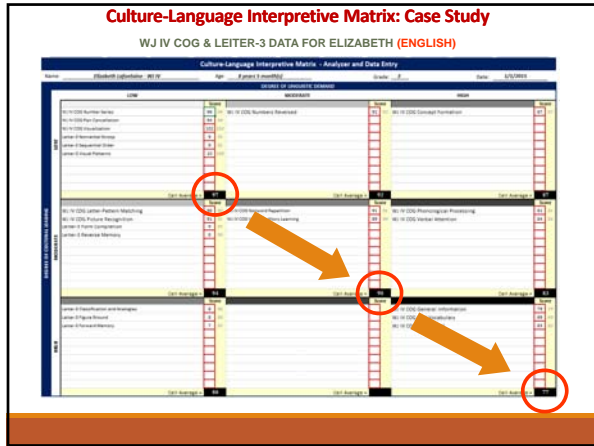
---

---

---

---

---




---

---

---

---

---

---

---

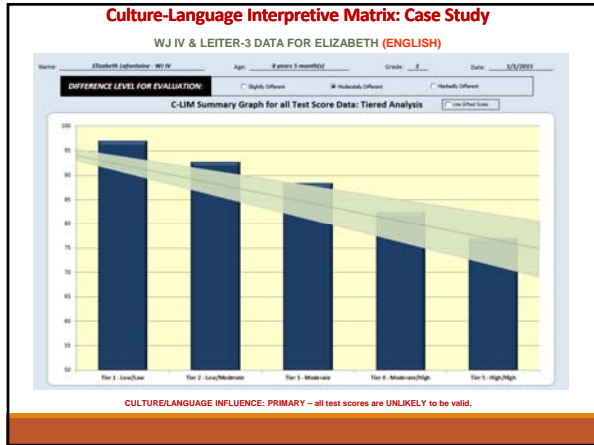
---

---

---

---

---




---

---

---

---

---

---

---

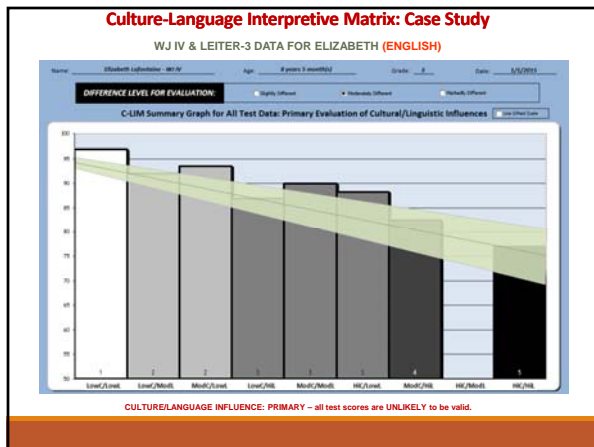
---

---

---

---

---




---

---

---

---

---

---

---

---

---

---

---

---

### Culture-Language Interpretive Matrix: Case Study

WISC-V ONLY DATA FOR YUQUITA (ENGLISH)

WISC-V Subtest	Score
Vocabulary	101
Block Design	101
Full Scale IQ	101
Matrix Reasoning	101
Similarities	101

---

---

---

---

---

---

---

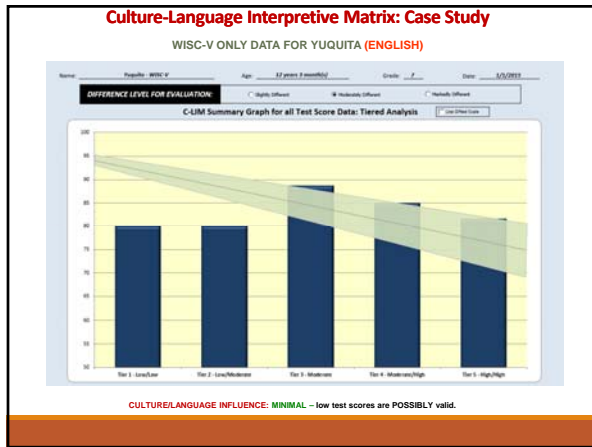
---

---

---

---

---



---

---

---

---

---

---

---

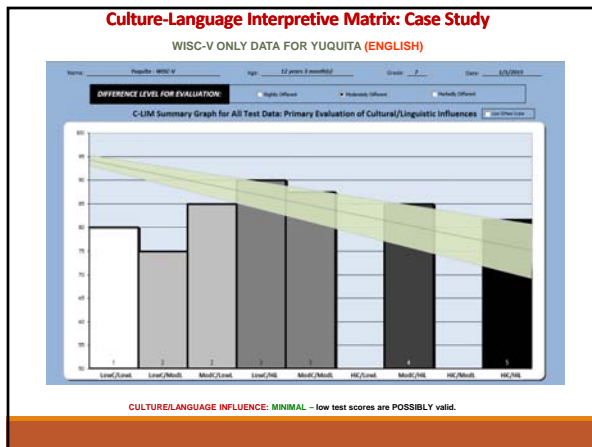
---

---

---

---

---



---

---

---

---

---

---

---

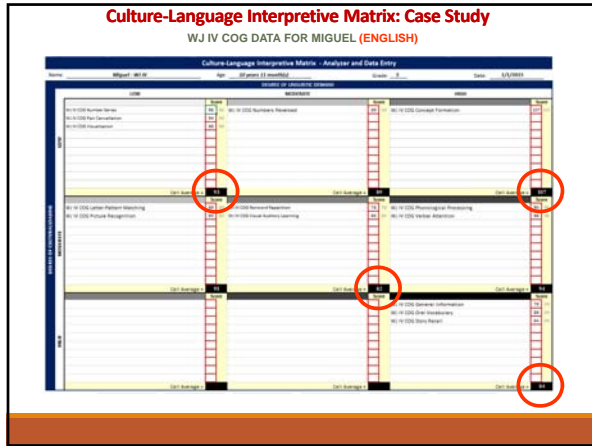
---

---

---

---

---




---

---

---

---

---

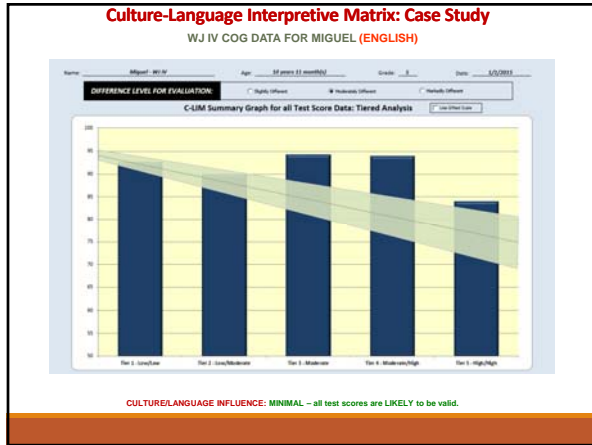
---

---

---

---

---




---

---

---

---

---

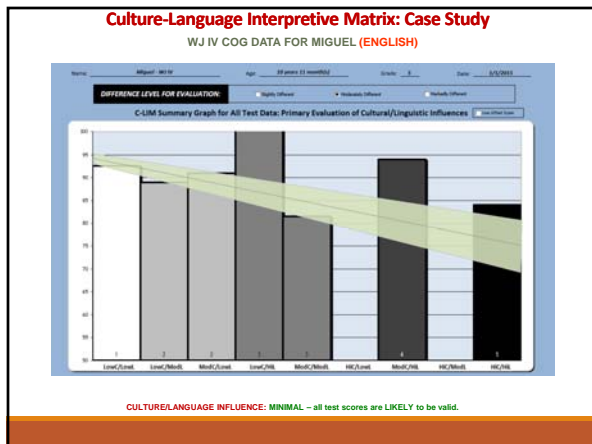
---

---

---

---

---




---

---

---

---

---

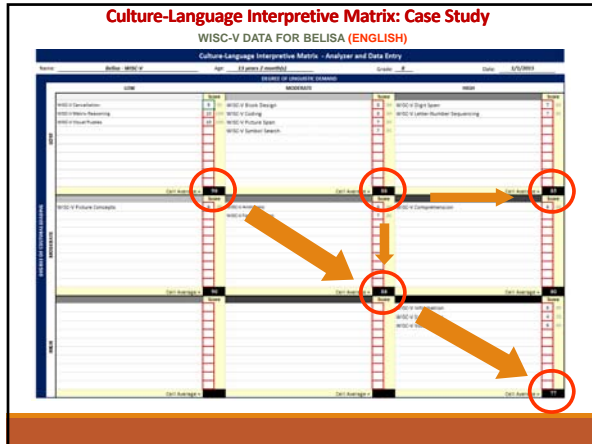
---

---

---

---

---



---

---

---

---

---

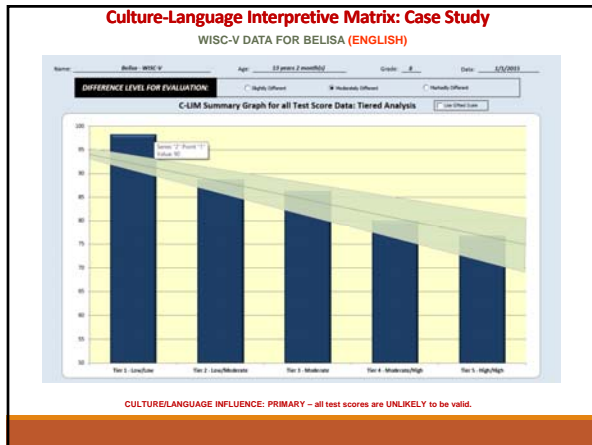
---

---

---

---

---



---

---

---

---

---

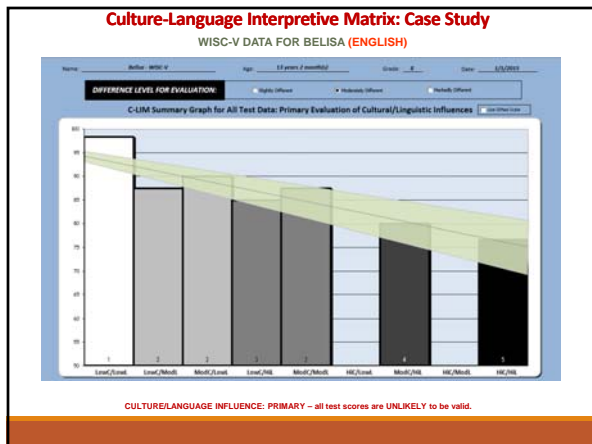
---

---

---

---

---



---

---

---

---

---

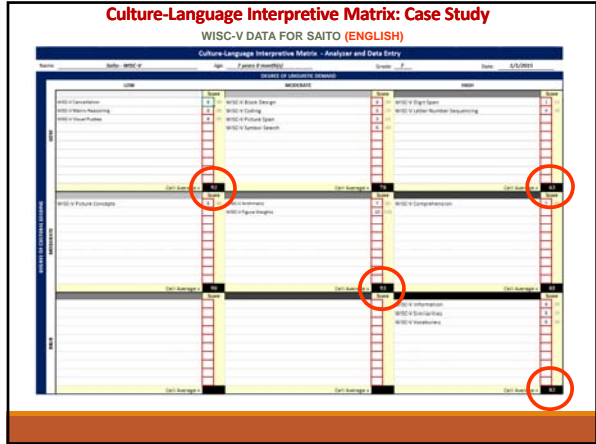
---

---

---

---

---




---

---

---

---

---

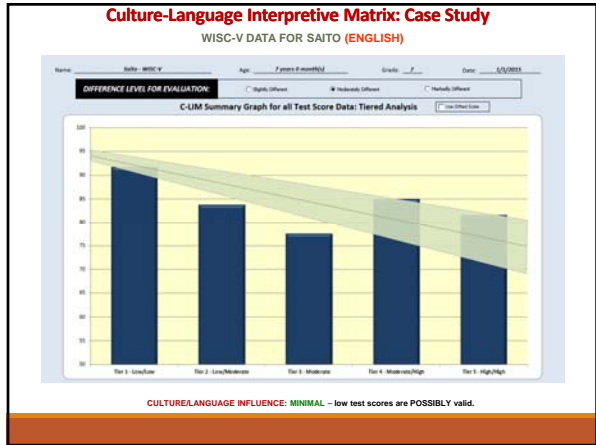
---

---

---

---

---




---

---

---

---

---

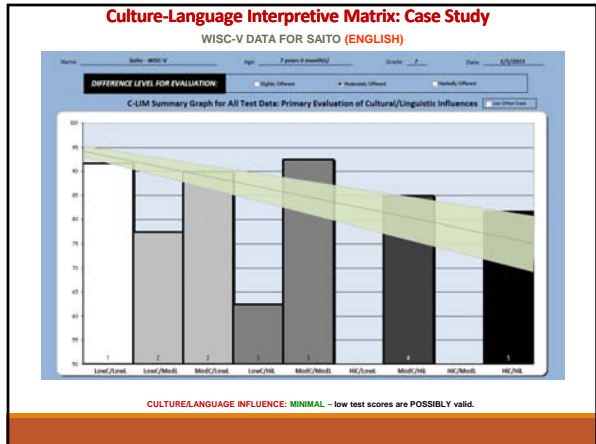
---

---

---

---

---




---

---

---

---

---

---

---

---

---

---

**Nondiscriminatory Interpretation of Test Scores**  
**The Gc caveat for English Language Learners**

Because Gc is, by definition, comprised of cultural knowledge and language development, the influence of cultural and linguistic differences cannot be separated from tests which are designed to measure culture and language. Thus, *Gc scores for ELLs, even when determined to be valid, remain at risk for inequitable interpretation and evaluation.*

Much like academic tests of manifest skills, Gc scores do reflect the examinee's current level of English language proficiency and acculturative knowledge. However, they do so as compared to native English speakers, not to other ELLs. This is discriminatory and *comparison of Gc performance using a test's actual norms remains unfair when assigning meaning to the value. It is necessary instead to ensure that both the magnitude and the interpretive "meaning" assigned to the obtained value is done in the least biased manner possible to maintain equity.*

For example, a Gc composite score of 76 would be viewed as "deficient" relative to the normative sample where the mean is equal to 100. However, for ELLs, interpretation of a Gc score of 76 should rightly be deemed as being indicative of "average" performance because it falls within the expected range on the C-LIM because it is instead being compared to other ELLs, not native English speakers. Interpreting Gc scores in this manner will help *ensure that ELLs are not unfairly regarded as having either deficient Gc ability or significantly lower overall cognitive ability*—conditions that may simultaneously decrease identification of SLD and increase suspicion of ID and speech impairment.

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

---

**Culture-Language Interpretive Matrix: Case Study**  
**KABC-II DATA FOR MARIO (ENGLISH)**

The screenshot shows a grid with columns for 'Cognitive', 'Language', and 'Cultural' domains, and rows for various subtests. Three red circles are drawn around specific cells: one in the 'Cognitive' column for 'Block Counting', one in the 'Language' column for 'Number Recall', and one in the 'Cultural' column for 'Picture Definition'.

---

---

---

---

---

---

---

---

---

---

---

---

---

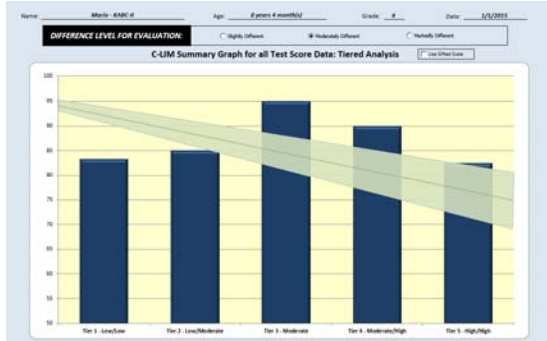
---

---

---

---

**Culture-Language Interpretive Matrix: Case Study**  
**KABC-II DATA FOR MARIO (ENGLISH)**




---

---

---

---

---

---

---

---

---

---

---

---

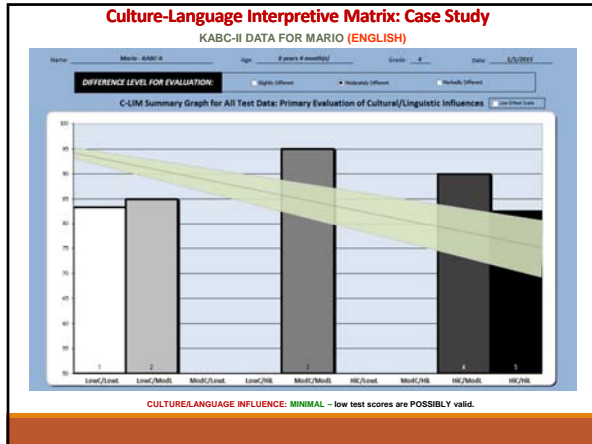
---

---

---

---

---



---

---

---

---

---

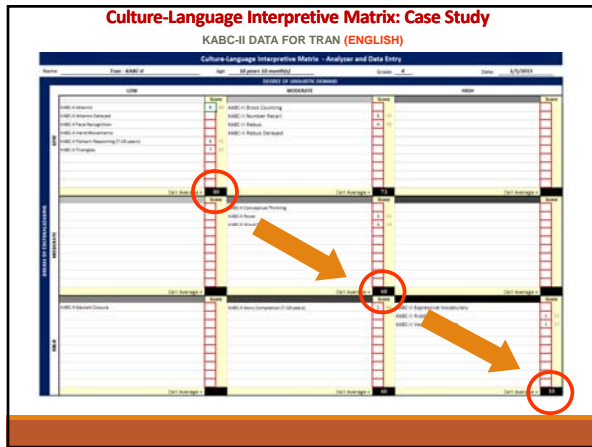
---

---

---

---

---



---

---

---

---

---

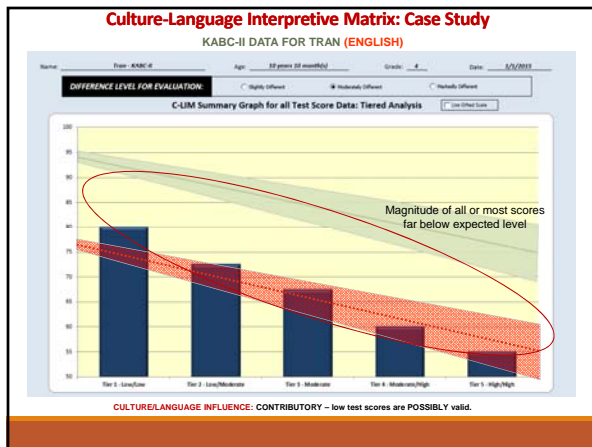
---

---

---

---

---



---

---

---

---

---

---

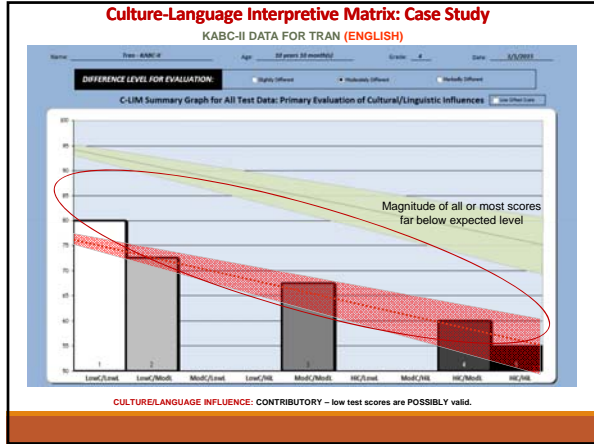
---

---

---

---






---

---

---

---

---

---

---

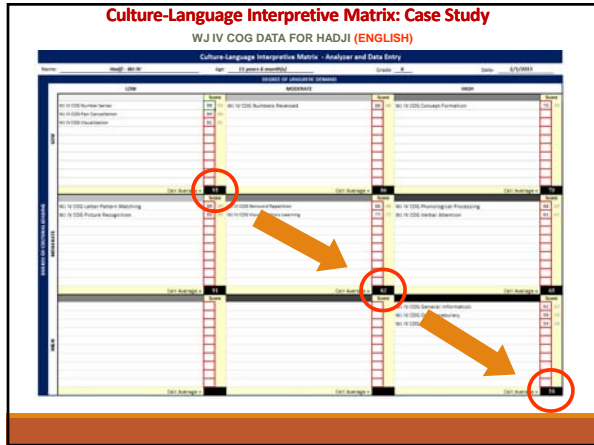
---

---

---

---

---




---

---

---

---

---

---

---

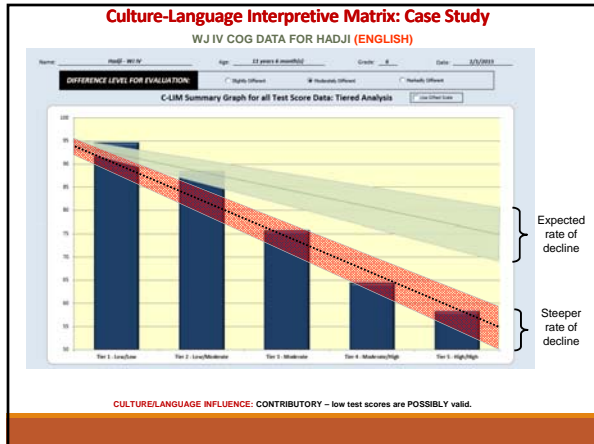
---

---

---

---

---




---

---

---

---

---

---

---

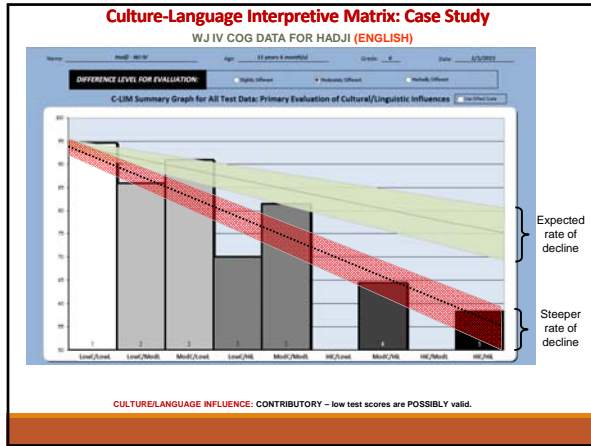
---

---

---

---

---




---

---

---

---

---

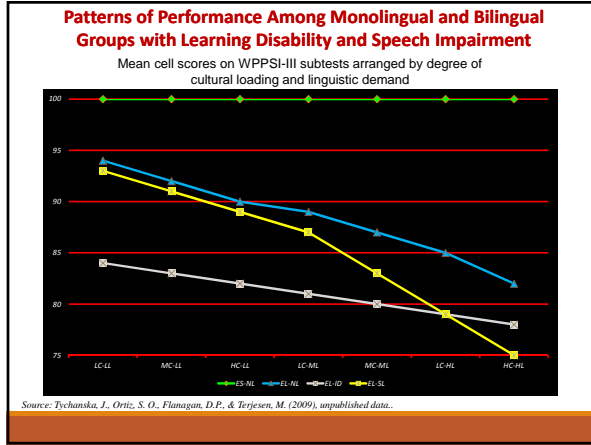
---

---

---

---

---




---

---

---

---

---

---

---

---

---

---

**Cross-Battery Assessment Software System (X-BASS® v1.2)**

**Culture-Language Interpretive Matrix - Notes**

#### Notes and Introduction

... (Text of notes) ...

#### Instructions for Use and Interpretation

... (Text of instructions) ...

Specific guidelines for determining degree of difference are included on the C-LIM Notes tab and are highlighted in yellow. The guidelines are not meant to be exhaustive or prescriptive but the determination is extremely critical and should be very well considered.

---

---

---

---

---

---

---

---

---

---

**Culture-Language Interpretive Matrix: The Importance of Difference**  
WJ III ONLY DATA

Subtests	Standard Score	Confidence Interval (95% Band)	Descriptions
Verbal Comprehension	64	56 – 72	Very Low
Visual-Auditory Learning	88	76 – 100	Low Average
Spatial Relations	98	91 – 107	Average
Sound Blending	75	64 – 87	Low
Concept Formation	70	62 – 78	Low
Visual Matching	86	76 – 97	Low Average
Numbers Reversed	80	67 – 93	Low
Incomplete Words	78	65 – 91	Low
Auditory Working Memory	85	76 – 94	Low Average
Analysis-Synthesis	78	66 – 90	Low
Auditory Attention	81	67 – 95	Low
Decision Speed	72	63 – 81	Low
Retrieval Fluency	82	69 – 95	Low
General Information	69	60 – 78	Very Low

---

---

---

---

---

---

---

---

---

---

---

---

**Culture-Language Interpretive Matrix: The Importance of Difference**  
XBA Culture-Language Interpretive Matrix (XBA C-LIM v2.0) for WJ III NU COG

Name: \_\_\_\_\_ Age: \_\_\_\_\_ Grade: \_\_\_\_\_ CLEAR DATA SAVE DATA

	DEGREE OF LINGUISTIC DEMAND			
	LOW	MODERATE	HIGH	
LOW	WJ III Spatial Relations	98	WJ III Analysis-Synthesis	78
	WJ III Numbers Reversed	80	WJ III Auditory Working Memory	85
	WJ III Visual Matching	86	WJ III Concept Formation	70
	Cell Averages	95	Cell Averages	83
MODERATE	WJ III Pair Cancellation	82	WJ III Auditory Attention	81
	WJ III Picture Recognition	82	WJ III Decision Speed	73
	WJ III Planning	82	WJ III Incomplete Words	78
	WJ III Visual Auditory Learning	88	WJ III Memory for Words	71
Cell Averages	85	Cell Averages	77	
HIGH	WJ III General Information	69	WJ III Verbal Comprehension	64
	Cell Averages	67		

---

---

---

---

---

---

---

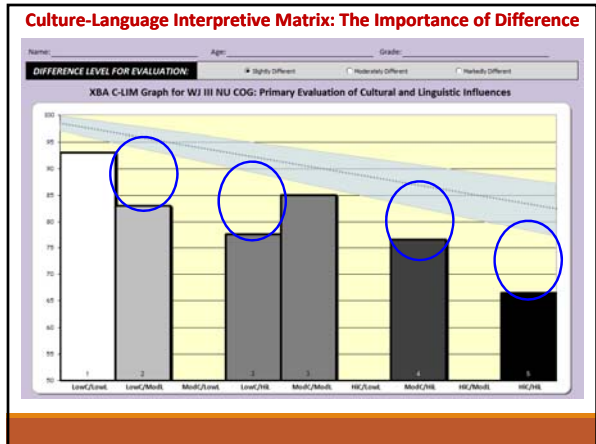
---

---

---

---

---




---

---

---

---

---

---

---

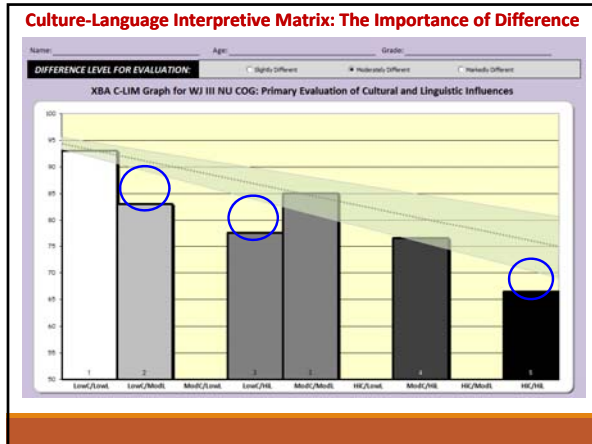
---

---

---

---

---




---

---

---

---

---

---

---

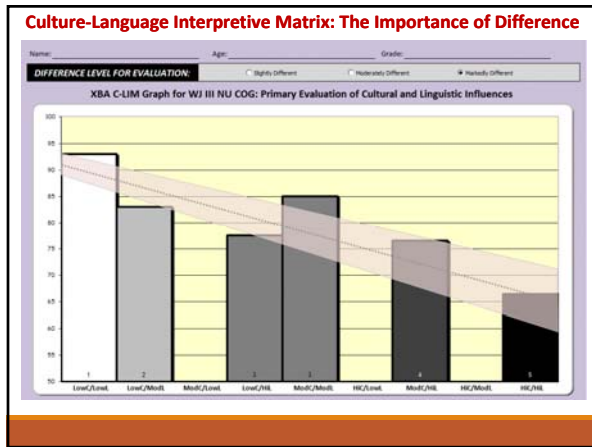
---

---

---

---

---




---

---

---

---

---

---

---

---

---

---

---

---

**The Culture-Language Interpretive Matrix (C-LIM)**

**Summary of Important Considerations for Use and Practice**

The C-LIM is not a test, scale, measure, or mechanism for making diagnoses. It is a visual representation of current and previous research on the test performance of English learners arranged by mean values to permit examination of the combined influence of acculturative knowledge acquisition and limited English proficiency and its impact on test score validity.

The C-LIM is not a language proficiency measure and will not distinguish native English speakers from English learners with high, native-like English proficiency and is not designed to determine if someone is or is not an English learner. Moreover, the C-LIM is not for use with individuals who are native English speakers.

The C-LIM is not designed or intended for diagnosing any particular disability but rather as a tool to assist clinician's in making decisions regarding whether ability test scores should be viewed as indications of actual disability or a mere reflection of differences in language proficiency and acculturative knowledge acquisition.

The C-LIM's primary purpose is to assist evaluators in ruling out cultural and linguistic influences as exclusionary factors that may have undermined the validity of test scores. Being able to make this determination is the primary and main hurdle in evaluation and the C-LIM can thus guide clinician's in their interpretation of test score data in a nondiscriminatory manner.

---

---

---

---

---

---

---

---

---

---

---

---

**The Culture-Language Test Classifications and Interpretive Matrix: Caveats and Conclusions**

Used in conjunction with other information relevant to appropriate bilingual, cross-cultural, nondiscriminatory assessment including...

- level of acculturation
- language proficiency
- socio-economic status
- academic history
- familial history
- developmental data
- work samples
- curriculum based data
- intervention results, etc.

...the C-LTC and C-LIM can be of practical value in helping establish credible and defensible validity for test data, thereby decreasing the potential for biased and discriminatory interpretation. Taken together with other assessment data, the C-LTC and C-LIM assist practitioners in answering the most basic question in ELL assessment:

*'Are the student's observed learning problems due primarily to cultural or linguistic differences or disorder?'*

---



---



---



---



---



---



---



---



---



---

**Assessment of English Language Learners - Resources**

**BOOKS:**

Rhodes, R., Ochoa, S. H. & Ortiz, S. O. (2005). Comprehensive Assessment of Culturally and Linguistically Diverse Students: A practical approach. New York: Guilford.

Flanagan, D. P., Ortiz, S.O. & Alfonso, V.C. (2013). Essentials of Cross-Battery Assessment, Third Edition. New York: Wiley & Sons, Inc.

Flanagan, D.P. & Ortiz, S.O. (2012). Essentials of Specific Learning Disability Identification. New York: Wiley & Sons, Inc.

Ortiz, S. O., Flanagan, D. P. & Alfonso, V. C. (2015). Cross-Battery Assessment Software System (X-BASS v1.0). New York: Wiley & Sons, Inc.



**ONLINE:**

New - Competency-based XBA Certification Program  
<https://www.schoolneuropsych.com/xba/>

CHC Cross-Battery Online  
<http://www.crossbattery.com/>




---



---



---



---



---



---



---



---



---



---