PRACTICAL APPLICATION OF CROSS BATTERY ASSESSMENT IN THE SCHOOL SETTING

PRESENTED BY:

AMY M. STRAUCH, LSSP
EVALUATION SPECIALIST
All information for this training was taken from the *Essentials of Cross-Battery Assessment – 3rd Edition* written by Dawn Flanagan, Samuel Ortiz, and Vincent Alfonso and subsequent trainings with the authors.
What is Cross Battery Assessment?

Founded in the research of CHC theory, XBA is a method of assessing cognitive abilities, academic achievement, and neuropsychological processes that allows the examiner to reliably & validly measure a wider range of skills and constructs than can be represented by a single stand-alone assessment.
CHC model v2.0 – Part 1 (Schneider & McGrew, 2012)

General Intelligence (g)

Broad
- Quantitative Knowledge (Gq)
  - Mathematical knowledge (MK)
  - Mathematical achievement (A3)
- Reading & Writing (Gw)
  - Reading decoding (RD)
  - Reading comprehension (RC)
  - Reading speed (RS)
  - Spelling ability (SG)
  - Writing ability (WA)
- Mathematics (Gm)
  - Knowledge of numbers (KN)
  - Calculations (CALC)
  - Problem-solving (PS)

Narrow
- Comprehension Knowledge (Gc)
  - General verbal information (KI)
  - Language development (LD)
  - General knowledge (GK)
  - Lexical knowledge (VL)
  - Phonological awareness (PA)
- Induction (I)
  - General sequential reasoning (Rg)
  - Quantitative reasoning (Rq)
- Fluid Reasoning (Gf)
  - Induction (I)
  - General sequential reasoning (Rg)
  - Quantitative reasoning (Rq)
  - Inductive reasoning (IR)
- Short-Term Memory (Gsm)
  - Working memory capacity (WC)
  - Immediate memory (IM)
  - Short-term memory (STM)
- Long-Term Storage & Retrieval (Glr)
  - Long-term memory (LTM)
  - Retrieval fluency (RF)
  - Retrieval efficiency (RE)
- Processing Speed (Gs)
  - Processing speed (PS)
  - Reaction time (RT)
  - Executive function (EF)

Acquired Knowledge

Functional groupings
- + = additional CHC abilities in groupings in Part 2 of model

Conceptual groupings

Domain-Independent General Capacities

Sensory-Motor Domain Specific Abilities (Sensory)

© Institute for Applied Psychometrics (IAP) 2-10-12
When is Cross-Battery Assessment Used?

- Whenever the constructs of interest cannot be assessed using a single battery
- When there is a need to follow up on inconsistent scores
- Comprehensive FIE
- Assessment of Specific Learning Disability
An operational definition of SLD is needed to provide any degree of confidence in the validity of a given SLD diagnosis.

(Flangan, Fiorello & Ortiz, 2010; Flangan, Ortiz, Alfonso, & Mascolo, 2006; Harrison & Holmes 2012; Kavale et al., 2009; Swanson, 1991)
Dual Discrepancy/Consistency (DD/C) Operational Definition of SLD

- If not for the presence of a cognitive deficiency affecting a specific academic skill, the student displays the capacity to perform satisfactorily.

- Definitions have almost always considered the overall cognitive ability of an individual.
Discrepancy #1

- There is an unexpected discrepancy between overall cognitive ability and academic achievement in a specific area.
Dual Discrepancy/Consistency (DD/C) Operational Definition of SLD

Discrepancy #2

- There is a discrepancy between overall cognitive ability and a specific deficit in linguistic competence, cognitive processes, or neuropsychological processes.
Dual Discrepancy/Consistency (DD/C)
Operational Definition of SLD

Consistency

- There is consistency between academic and cognitive deficits measured, demonstrated by a logical and empirical relationship that is confirmed with ecological validity.
Common Elements of “PSW Component” of Third Method Approaches to SLD Identification

COGNITIVE STRENGTH
Average or higher abilities and processes;
May also include strengths in academic skills

COGNITIVE WEAKNESS/DEFICIT
Cognitive Ability or Processing Disorder

ACADEMIC WEAKNESS/FAILURE
Academic Skills/Knowledge Deficits

Statistically significant difference between cognitive integrities and circumscribed cognitive ability or processing deficit(s)

Cognitive deficit(s) is specific, not general or pervasive, because overall cognitive ability is at least average

Consistent/Concordant

No Statistically significant Performance Difference (constructs are related empirically)

Statistically significant difference between cognitive integrities and academic skill deficit(s)

Academic deficit(s) is unexpected, not expected, because overall cognitive ability is at least average

Discrepant/Discordant

XBA Leveled Interpretation
The presence of a weakness or deficit of one or more areas of academic performance established through standardized testing and corroborated by other data sources such as CBM, clinical observations, work samples, etc.

Think in terms of the 8 areas of SLD qualification.
Level II involves evaluating the weaknesses or deficits found in Level 1 and determining whether they are or are NOT primarily the result of factors that may be largely external to the individual, noncognitive in nature, or the result of a condition other than SLD (called exclusionary factors)
What are the exclusionary factors that need to be ruled out?

<table>
<thead>
<tr>
<th>Exclusionary Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vision</td>
</tr>
<tr>
<td>2. Hearing</td>
</tr>
<tr>
<td>3. Motor Functioning</td>
</tr>
<tr>
<td>4. Cognitive and Adaptive Functioning</td>
</tr>
<tr>
<td>5. Social-Emotional/Psychological Factors</td>
</tr>
<tr>
<td>6. Environmental/Economic Factors</td>
</tr>
<tr>
<td>7. Cultural/Linguistic Factors</td>
</tr>
<tr>
<td>8. Physical/Health Factors</td>
</tr>
<tr>
<td>9. Instructional Factors</td>
</tr>
</tbody>
</table>
Weaknesses or Deficits in Cognitive Abilities or Neuropsychological Processes
Weakness vs. Deficit

- Weakness – performance on standardized norm-referenced tests that fall “below average” (associated with scores of 85-89)

- Deficit - performance on standardized norm-referenced tests that falls greater than 1 SD below the mean (associated with scores <85)
Level IV: A Pattern of Strengths and Weaknesses Characterized by Dual Discrepancy/Consistency

An examiner determines:

1. Whether there is a relationship between cognitive and academic weaknesses
2. Whether the student displays generally average ability
3. Whether the learning difficulty is domain specific
4. Whether the individual’s underachievement is unexpected.

Use of the XBA Software is critical in the determination.
Determine the degree in which the identified learning deficit impacts educational performance.

Determine whether the identified condition of SLD actually impairs academic functioning to such an extent that special education services are necessary.
Guiding Principles of XBA:

Principle 1 – Selecting a comprehensive ability battery as the core of the assessment.
- Tests that are most responsive to referral concerns

Principle 2 – Use Norm based composites from a single battery whenever possible to represent broad CHC abilities.
- Use test composites when they are interpretable. Use XBA composites as an alternative when needed.
Guiding Principles of XBA:

Principle 3 – When constructing Broad & Narrow CHC clusters, select tests that have been classified using an acceptable method.

- Use tests that validly measure what you think they measure.

Principle 4 – Use 2 different indicators of a broad ability to create a composite.

- If the core battery selected has only 1 or NO measures of the broad ability, select another test to supplement.
Guiding Principles of XBA:

Principle 5 – When crossing batteries select tests normed and developed within a few years of one another to reduce the Flynn effect.

- **Flynn effect** – The substantial increase in average scores on intelligence tests globally as measured since 1930.
- All subtests within the XBA software have been normed within 10 years of one another.

Principle 6 – Select tests from the smallest number of batteries to avoid the effects of different norm samples.

- Minimize unintended errors in measurement & comparison
**Guiding Principles of XBA:**

**Principle 7** – Establish ecological validity for any and all test performance that suggest weakness or deficit.

- Find evidence of how the weakness manifests in daily living – most likely with academic achievement.
- Make clear connections between cognitive dysfunction and the educational impact of the dysfunction in classroom performance.
Caution: If a test is not included, there’s a good reason!

1. Normed or published before 2001
2. Year of normative data was gathered was not reported.
3. Test not normed in the U.S.
4. Test did not include age-based norms.
5. Test was not norm referenced or provide normative scores (standard, scaled, etc.)
The process of interpreting test data should be conducted within a broader conceptual framework of hypothesis generation in order to measure the functional assumptions about the expected performance outcome.
Step 1 – Select a battery that is appropriate after giving consideration to the following:

- Age
- Developmental level
- Acculturative experiences and background
- English proficiency (not just dominance!)
- Referral concerns
Construction of a Cross-Battery Assessment

1. Make connections between presenting difficulties in the referral and related cognitive/neuropsychological processes.
2. Determine the scope of assessment and purpose.
3. Generate the Hypothesis
4. Ensure **Referral Relevant** abilities are measured & data gathered is sufficient to address concerns and draw conclusions.
Step 2: Identify broad abilities that are/are not measured by the selected battery

- Ensure the broad abilities that are measured within the battery are well represented.
- Each broad ability must be represented by at least two narrow abilities.
- Each of the narrow abilities represented must be **qualitatively different** from the other in order to use the broad ability as a valid estimate of ability.
Step 3: Identify narrow abilities that are/are not measured by the battery

- Necessary when the nature of the referral requires measurement of the specifics of a cognitive ability and or neuropsychological process.

- When referrals are specific to reading, math and written language, the narrows that best measure these should be utilized.
Step 4: Administer/Scores selected Battery and Supplemental Test

Tests should be administered and scored as instructed by the test manual.
Interpretation of Broad CHC Ability

- Broad Ability Constructs are represented by at least 2 qualitatively different narrow subtests.
- The Broad Ability Composite must be considered COHESIVE to be a likely indicator of the abilities it represents.
- Composite Cohesion can be calculated from 2, 3, or 4, subtests entered.
- Composite Cohesion is based on:
  1. The magnitude of the Standard Deviation score difference
  2. The rarity of the difference that occurred
Narrow CHC Ability Interpretation

- Occurs when a Non-Cohesive Broad Ability is obtained & the examiner chooses to follow up on the lower score, by giving another measure of that narrow ability

- If the third score obtained forms a cohesive score with the lower narrow ability, then a Narrow Ability Composite can be reported.

- The Divergent score is considered a relative strength
Cohesive and Not Cohesive
have replaced the terms
Unitary and Nonunitary

The term unitary was sometimes
misunderstood to mean that there
was **only one** theoretical dimension
underlying the broad score, which
was not the intent.
Rules of Thumb have been replaced by Statistical Analysis included in the XBA DMIA.

Rules may have been applied too stringently or not stringently enough. Analysis now relies on statistical calculation within the XBA DMIA based on Standard Deviation difference & base rate data.
Outlier has been replaced by Divergent Score

For some, the term outlier meant that the obtained score could be disregarded or ignored. That was never the intent. Therefore, divergent was chosen to describe scores that are significantly different from the composite.
When the XBA DMIA gives an Interpretive Statement indicating Follow Up is needed, that does not always mean more tests.

Evaluators should consider all available data sources in follow-up:

- Previous evaluation
- Work Samples
- Error Analysis
- Parent/Teacher/Student report
- Intervention Data
- Additional Testing
Cohesion asks two primary questions:

1. Is the variability between the subtests making up the composite significant or substantial? (more than 2/3 SD 10 points)

2. Is the variability between the subtests making up the composite infrequent or uncommon? (occurs in less than 10% of the population)
Pattern of Strengths & Weaknesses Analyzer

**CHC Ability Composite**
- Broad Ability Composite
- XBA Composite from CHC Analyzer
- Composite from specific test

**Sufficiency**

**Determining Sufficiency:**
An ability is considered "sufficient" when it is judged by the evaluator to contribute meaningfully to the individual's overall cognitive functioning, particularly for the purpose of facilitating academic performance (e.g., acquisition and development of academic skills).

Typically, standard scores around 90 or higher are sufficient, as abilities associated with scores in this range (≥ 90) often contribute meaningfully to the individual's overall cognitive functioning and, therefore, support learning.

When standard scores are around 90 or lower, clinical judgment is necessary to determine if the broad ability constrains or inhibits learning and achievement.
PSW Data Entry

- Enter Scores
- Select Sufficiency
- Proceed to PSW!

<table>
<thead>
<tr>
<th>CHC ABILITY COMPOSITES</th>
<th>Enter Standard Scores (Range 40 - 160)*</th>
<th>Select Yes or No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gc - Crystallized Knowledge</td>
<td>84</td>
<td>Yes</td>
</tr>
<tr>
<td>Gf - Fluid Reasoning</td>
<td>93</td>
<td>Yes</td>
</tr>
<tr>
<td>Glr - Long-Term Storage &amp; Retrieval</td>
<td>90</td>
<td>Yes</td>
</tr>
<tr>
<td>Gsm - Short-Term Memory</td>
<td>89</td>
<td>Yes</td>
</tr>
<tr>
<td>Gv - Visual Processing</td>
<td>90</td>
<td>Yes</td>
</tr>
<tr>
<td>Ga - Auditory Processing</td>
<td>93</td>
<td>Yes</td>
</tr>
<tr>
<td>Gs - Processing Speed</td>
<td>71</td>
<td>Yes</td>
</tr>
<tr>
<td>CHC ABILITY COMPOSITES</td>
<td>Enter Standard Scores (Range 40 - 160)*</td>
<td>Select Yes or No</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>----------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Gc - Crystallized Knowledge</td>
<td>90</td>
<td>☐ Yes ☐ No</td>
</tr>
<tr>
<td>Gf - Fluid Reasoning</td>
<td>82</td>
<td>☐ Yes ☐ No</td>
</tr>
<tr>
<td>Glr - Long-Term Storage &amp; Retrieval</td>
<td>90</td>
<td>☐ Yes ☐ No</td>
</tr>
<tr>
<td>Gsm - Short-Term Memory</td>
<td>85</td>
<td>☐ Yes ☐ No</td>
</tr>
<tr>
<td>Gv - Visual Processing</td>
<td>90</td>
<td>☐ Yes ☐ No</td>
</tr>
<tr>
<td>Ga - Auditory Processing</td>
<td>93</td>
<td>☐ Yes ☐ No</td>
</tr>
<tr>
<td>Gs - Processing Speed</td>
<td>71</td>
<td>☐ Yes ☐ No</td>
</tr>
</tbody>
</table>
The $g$-Value reflects overall cognitive ability based on the broad CHC abilities judged by the evaluator to be "sufficient." The $g$-Value is interpreted according to the likelihood that an individual possesses at least average overall cognitive ability.

- $\geq 0.60 = $ average overall ability is very likely
- $0.51 - 0.59 = $ more information needed
- $\leq 0.50 = $ average overall ability is unlikely

**Note:** An asterisk (*) next to a broad ability code indicates that the ability was judged as "insufficient" by the evaluator.
PSW Data Entry

- g-Value will show green, yellow, or red
- More info does not have to mean more testing – verification thru multiple data sources

*Note:* An asterisk (*) next to a broad ability code indicates that the ability was judged as "insufficient" by the evaluator.
The g-value remains an indication of the likelihood that the individual has at least average overall ability to think and reason. The value has been rescaled from 0-1.0 where scores > or = 0.60 are considered indicative of average ability.
Intact Ability Estimate (IA-e)

This score is generated by the PSW-A and provides an estimate of overall intellectual ability. It is similar to a full scale IQ score...BUT it is the aggregate of ONLY the intact cognitive abilities measured, factoring out the potential negative impact of the identified cognitive weaknesses.
1a. Intact Ability Estimate
This composite is calculated using median reliabilities and intercorrelations among the CHC broad ability scores judged as sufficient on the g-Value tab.

The Intact Ability Estimate (IA-e) appears in green when it is ≥ 90 and the g-Value ≥ .60. The IA-e appears in yellow when it is between 85–89, inclusive, or the g-Value is between .51–.59, inclusive. "N/A" will appear if the IA-e is < 85 or the g-Value < .50, or if there were too few abilities judged to be sufficient (i.e., < 3). When "N/A" appears in the box no further analyses can or should be performed. When an alternative score is entered below, double dashes "--" are displayed.

Note: If you would like the program to use a value other than the IA-e, you may enter an alternative score here. Be sure that the value you enter here is an adequate representation of the individual's overall cognitive ability and is greater than or equal to 85. Simply delete this value if you wish to return to using the IA-e.

1b. Alternative Ability Estimate
You may enter an alternative value if desired or when the IA-e is not believed to be a good estimate of general ability.

2a. Cognitive Weakness
Enter the scaled/standard score and subtest or composite name in the boxes on the right that best represents the student's cognitive weakness or deficit. If using T-Scores, convert to Standard Scores before entering (use Tab2A).

This score should be the best estimate of a cognitive weakness or deficit. Indicate whether the score is a composite/subtest and select the cognitive area it represents. For example, if you entered a "working memory" composite, check "Composite" and select "Gsm-Short-Term Memory" from the drop down menu.

2b. Frequency of Difference
Select the level to be used in PSW analysis for determining if the size of a difference is infrequent or uncommon. The default value is 5% and will be adjusted for test unreliability. A more conservative or liberal value may be selected. If a second comparison is being made or a subtest is used, consider using a stricter value.

3a. Academic Weakness
Enter a scaled/standard score (required) and the name of the subtest or composite (optional) in the boxes at the right that represents a significant area of academic weakness or deficit for the individual.

3b. Frequency of Difference
Select the level to be used in PSW analysis for determining if the size of a difference is infrequent or uncommon. The default value is 5% and will be adjusted for test unreliability.
Cross-Battery Pattern of Strengths and Weaknesses Analyzer (XBA PSW-A® v1.0)
Conceptualization by D.P. Flanagan, S.O. Ortiz, and V.C. Alfonso
Copyright © 2013 Wiley. All Rights Reserved

Are weaknesses domain specific?
Using the IA-e as the predictor, if the difference between Actual and Predicted specific cognitive performance equals or exceeds the Critical Value, then the size of the difference is unusually large and infrequent and the weakness is domain specific.

Difference: 20.72
Critical Value: 12.88
Yes, domain specific

Is the difference statistically significant?
p < .05

AGGREGATE OF COGNITIVE STRENGTHS
The aggregate is either the Intact Ability Estimate (IA-e) or a user-entered alternative value that represents the individual’s overall ability.
g-Value = 0.69
88

Is underachievement unexpected?
Using the IA-e as the predictor, if the difference between Actual and Predicted specific academic performance equals or exceeds the Critical Value, then the size of the difference is unusually large and infrequent and underachievement is unexpected.

Difference: 23.04
Critical Value: 16.91
Yes, unexpected underachievement

Cognitive Weakness
Below are the individual’s Actual and Predicted performances in the area of:
Processing Speed
Actual: 71
Predicted: 92

Academic Weakness
Below are the individual’s Actual and Predicted performances in the area of:
Reading Fluency
Actual: 70
Predicted: 93

Is there a BELOW AVERAGE aptitude-achievement consistency?
The box above addresses this component of learning disability through consideration of the degree to which the meaning of the scores is similar (e.g., indicative of weakness or deficit) and the position of the scores is consistent.
Conclusion.....

It’s important to understand that if:

a. A student did not respond well to quality instruction/interventions,

b. The inclusionary PSW criteria are met, and

c. Exclusionary factors are ruled out as the PRIMARY cause of academic deficits,

then SLD is *highly probable*. 
Amy M. Strauch

amy.strauch@esc20.net

210-370-5440