Designing and Implementing data driven mental health services From Universal screening to progress monitoring behavior interventions.

PART B

Achilles N. Bardos, Ph.D.
Univ. of Northern Colorado
School Psychology Programs
Greeley, Colorado, USA
www.achillesbardos.com

INTRODUCING THE

BIMAS
Behavior Intervention Monitoring Assessment System

BIMAS2
Behavior Intervention Monitoring Assessment System
What is the BIMAS?

1. **Screening**- as a brief screening device to detect students in need of further assessment and to identify their respective areas of strengths and needs.

2. **Student Progress Monitoring**- To provide feedback about the progress of individual students or clients.

3. **Program Evaluation** - To gather evidence that intervention services are effective.

USES OF THE BIMAS

For those are required (or wish) to have an outcome measure sensitive to short term therapeutic gains

- school-based mental health providers
- Public/private organizations providing school or community-based intervention programs
- community mental health agencies
- managed care agencies (HMOs)
- Private practitioners
FORMAT OF THE BIMAS

• A multi-informant assessment system
  – Parent
  – Teacher
  – Self (12 -18 yrs old)
  – Clinician

Background & Development
BIMAS foundation - Earlier Studies

Intervention Item Selection Rules (IIRS)

- Meier studied scales constructed with both traditional and IISR procedures in a variety of clinical and school settings (Meier, 2004, 2000, 1998).
- Overall, scales constructed with IISR procedures demonstrated larger treatment effect sizes than traditional scales and adequate reliability estimates.

Central philosophy of the IISRs...

- Intervention-sensitive items should change in response to an intervention and behave in a theoretically expected manner in other conditions (e.g., remain stable over time when no intervention is present).
- Tests are likely to be shorter in length than traditional screening instruments because IISR scales contain only change-sensitive items.
Central philosophy of the IISRs…

• Items will share some characteristics with traditional, trait-sensitive tests.
  – theoretically based,
  – reliable,
  – unrelated to systematic error sources.

• However, intervention-sensitive items should possess additional properties, foremost of which is that they change in response to an intervention.

Dr. Scott Meier
Intervention Item Selection Rules

<table>
<thead>
<tr>
<th>Rule</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ground scale items in theoretical and empirical literature relevant to applicable interventions and target problems</td>
</tr>
<tr>
<td>2</td>
<td>Aggregate at appropriate levels</td>
</tr>
<tr>
<td>3</td>
<td>Assess range of item scores at pretest</td>
</tr>
<tr>
<td>4</td>
<td>Detect change in an item’s score after an intervention</td>
</tr>
<tr>
<td>5</td>
<td>Assess whether change occurs in the expected direction</td>
</tr>
<tr>
<td>6</td>
<td>Examine whether differences in change exist between intervention and comparison groups</td>
</tr>
<tr>
<td>7</td>
<td>Examine whether intake differences exist between comparison groups</td>
</tr>
<tr>
<td>8</td>
<td>Examination relations between item scores and systematic error sources</td>
</tr>
<tr>
<td>9</td>
<td>Aggregate selected items into scale(s) and cross-validate</td>
</tr>
</tbody>
</table>
1. Ground items in theory

1. *Ground items in previous research and theory.* Relevant research and theory provide a context for understanding the meaning of changing scores on an intervention-sensitive measure. In the area of child and adolescent psychotherapy, Kazdin (2000) noted that more than 1,000 controlled studies of psychosocial interventions for children and adolescents exist. Kazdin maintained that because ESs for all interventions averaged about .70 for children and adolescents, maturation alone cannot account for such gains. Meta-analytic studies also indicate that adolescents appear to benefit more from psychotherapy than children, although most of the difference can be attributed to the benefits received by adolescent girls (Weisz, Huey, & Weersing, 1998). Applied to this study, these findings suggest that (a) some PE-BIMAS items

2. Aggregate Items at an appropriate level.

2. *Aggregate items at an appropriate level.* Because an item response contributed by an individual on one occasion may be influenced by random error (Messick, 1989), item responses should first be aggregated across individuals before further analyses are conducted. Similarly, test developers have long recognized that aggregation of individual item responses into scales increases the reliability and validity of measurement of the studied construct. Intervention-sensitive items are not aggregated across occasions, however, but summed across individuals and items. As was done in this study, item scores are then compared across time periods in which interventions take place to determine if change effects are present at the level of aggregated item responses.
3. Avoid ceiling, floor & under-estimation effects

3. Assess range of item scores at pretest. Ceiling and floor effects inhibit detection of desired changes in intervention-sensitive tests because they can restrict the potential range of scores. In this study, a ceiling effect occurred when an item’s standard deviation was added to the item mean and the resulting sum exceeded the highest value of the scale (3); a floor effect occurred when the item’s standard deviation was subtracted from the item mean and the result was less than the bottom range of the scale (0). Three Strengths items in both subsamples had a ceiling effect: communicates clearly, starts conversations, and limits set with children. No floor effects were found.

4. Demonstrate Change in Interventions

4. Items should evidence change in intervention conditions. Intervention-sensitive items should demonstrate change, from baseline to follow-up periods, with clients who receive psychosocial interventions (cf. Cronbach et al., 1980). For the current study, paired r tests were computed to examine change in item scores from intake to follow-up. Because these analyses are exploratory in nature, and the expected effects at the level of an individual item are likely to be small, an α level of .10 was set to detect statistically significant change (cf. Meier, 2000). As shown in Table 2, 12 of 19 items evidenced statistically significant change in one or both subsamples: controls temper, pays attention to speakers, stays out of trouble, communicates clearly, shares thinking, feels depressed, behaves differently, acts impulsively, fights with others, family members fight, lies or cheats, and gets failing grades.
5. Change in the direction expected

5. *Items should evidence change in the theoretically expected direction.* All 12 items that evidenced significant change in Table 2 improved from intake to follow-up. Although clients worsened in at least one subsample on the items makes friends easily, limits set with children, and helps with household tasks, these changes did not reach statistical significance.

6. Evaluate item change in intervention and control groups

6. *Evaluate item change in intervention and comparison groups.* Item change in intervention groups can be compared to change in items completed by available comparison groups. As noted above, Kazdin’s (2000; see also Weisz, Weiss, Han, Granger, & Morton, 1995; Webster-Stratton, 1996) review found that girls evidence more improvement than boys as a result of psychosocial interventions. Meta-analytic
7. Examine equivalence of item scores at intake between groups

7. Examine the equivalence of items scores at intake between groups. In the PE-BIMAS data set, intake equivalence could be examined between the two randomly created subsamples A and B. Paired t tests were used to assess differences between item means, and two items differed at intake: makes friends easily ($t = -1.78$, $p < .10$) and family members fight ($t = -2.37$, $p < .05$). Overall, random assignment resulted in statistically equivalent groups, providing confidence that the subsequent cross-validation analyses (IISR 9) of intervention-sensitive items can be interpreted appropriately.

9. Aggregate selected items into scales and cross-validate.

8. Examine the relationship between scale items and systematic error sources. No data were available for addressing this IISR.

9. Aggregate selected items into scale(s) and cross-validate. The preceding IISR analyses provide a basis for understanding the relevant properties of scale items and lay the foundation for subsequent decisions about inclusion in multi-item scales.
The BIMAS Scale Structure
BIMAS Standard Form

Behavioral Concern Scales:
- **Conduct**: anger management problems, bullying behaviors, substance abuse, deviance
- **Negative Affect**: anxiety, depression
- **Cognitive/Attention**: attention, focus, memory, planning, organization

Adaptive Scales:
- **Social**: social functioning, friendship maintenance, communication
- **Academic Functioning**: academic performance, attendance, ability to follow directions

The Conduct scale items
- ✓ appeared angry.
- ✓ engaged in risk taking behavior(s).
- ✓ fought with others (verbally, physically, or both).
- ✓ lied or cheated.
- ✓ lost his/her temper when upset.
- ✓ was aggressive (threatened or bullied others).
- ✓ was suspected of using alcohol and/or drugs.
- ✓ was sent to an authority for disciplinary reasons.
- ✓ was suspected of smoking or chewing tobacco.
The Negative Affect scale

✓ appeared sleepy or tired.
✓ appeared depressed.
✓ acted sad or withdrawn.
✓ was easily embarrassed or felt ashamed
✓ appeared anxious.
✓ expressed thoughts of hurting self.
✓ was emotional or upset.

The Cognitive/Attention scale

✓ had trouble paying attention.
✓ was impulsive.
✓ had problems staying on task.
✓ acted without thinking.
✓ had trouble remembering.
✓ had difficulties with organizing things.
✓ fidgeted.
✓ had trouble planning.
The Social Scale

✓ shared what he/she was thinking about.
✓ spoke clearly with others.
✓ maintained friendships.
✓ appeared comfortable when relating to others.
✓ was generally friendly with others.
✓ worked out problems with others.
✓ attended his/her scheduled therapy appointments. (Clinician Form)

The Academic Functioning Scale

(parent & teacher form)

✓ Followed directions
✓ Received failing grades
✓ Worked up to his/her academic potential
✓ Went prepared to class
✓ Was absent from school
The BIMAS-Flex

• 10 extra items were created for each of the Standard Form screener items with specific to or closely related behaviors /emotions.

• Flex items can be selected by the intervention team (Parent, school, clinician) and be customized for each child as needed.

BIMAS Flex Example

Standard Item:
Fought with others (verbally, physically, or both)

Negatively worded:
• Argued with peers
• Argued with teachers
• Argued with parents
• Argued with siblings
• Talked back to parents
• Talked back to teachers
• Physically hurt peers
• Physically hurt parents
• Physically hurt teachers
• Physically hurt siblings
• Threatened peers
• Threatened teachers
• Threatened parents
• Threatened siblings

Positively worded:
• Showed regret after a fight
• Was respectful to adults
• Walked away from a fight
• Prevented a fight
• Stopped an argument
• Found a positive outlet for frustration
• Avoided a verbal confrontation

Or... custom create your own!
Best use of BIMAS Flex items

- Select items based on elevated Standard scale score for an individual student
- make notes to describe specific behaviors, response to services, or to add other comments.
- If student was previously tested or diagnosed using any child behavioral assessment instrument (e.g., Conners, BASC-2, etc.), then BIMAS-Standard administration might not be necessary for BIMAS-Flex administration but highly recommended.
- Parent, Teacher, Self and Clinician forms

---

How to use the BIMAS

UNIVERSAL SCREENING

AND

PROGRESS MONITORING
Universal Screening and Progress Monitoring

**Universal Screening**
Universal screening is a type of assessment that is characterized by the administration (usually three times a year) of quick, low-cost, repeatable data collection of academic and behavioral skills of all students. It shows how functional the curriculum and instruction are in the school and detects whether or not students are making acceptable progress in the curriculum.

**Progress Monitoring**
Progress monitoring is a systematic approach to gathering academic and behavioral data using a variety of data collection methods. Student performance is examined frequently, over time, to evaluate response to instruction and intervention.

---

**UNIVERSAL SCREENING**  
**(Tier I)**

- Use the BIMAS-Standard form....
  - Early Fall
  - Mid year
  - Late Spring

- Outcome!!! receive data for decision making at:
  - System level (Schools, grades, classrooms)
  - Individual data (specific students)
BIMAS System level data -- district

---

BIMAS System level data across schools

---
BIMAS system level data across classrooms

PROGRESS MONITORING Tier I

- Screening is “benchmarking” not a one-shot approach or practice.
- Universal Screenings should be done across time for the....
  - Entire school district
  - Across Buildings
  - Across Grades
**PROGRESS MONITORING**
**Tiers II & III**

- Includes a smaller group or individual students
  - Identified from Tier I screening using a combo of…
    - BIMAS - Standard form data
    - Other school data (PBIS-ODR, referrals)

**PROGRESS MONITORING**
**Individual student across any Tier**
Administration & Scoring

COMPLETELY WEB_BASED

BIMAS
Technical Information

• Norms development
• Psychometric properties
  – Reliability
  – Validity
Normative Sample

- **Total Sample**
  - Parent: $N = 1,867$ (Normative: $N = 1,400$, Clinical: $N = 467$)
  - Teacher: $N = 1,938$ (Normative: $N = 1,400$, Clinical: $N = 538$)
  - Self-Report: $N = 1,050$ (Normative: $N = 700$, Clinical: $N = 350$)

Psychometric Properties

- **Reliability**
  - Internal Consistency
  - Test-Retest (stability)

- **Validity**
  - Content and sources of information for decision making
  - Construct
    - Scale structure
    - Screening accuracy
    - Concurrent validity
    - Progress monitoring
Internal Consistency
Cronbach’s Alpha

<table>
<thead>
<tr>
<th>Form</th>
<th>Behavioral Concern Scales</th>
<th>Adaptive Scales</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Conduct</td>
<td>Negative Affect</td>
</tr>
<tr>
<td>Parent</td>
<td>.87</td>
<td>.82</td>
</tr>
<tr>
<td>Teacher</td>
<td>.91</td>
<td>.85</td>
</tr>
<tr>
<td>Self-Report</td>
<td>.88</td>
<td>.85</td>
</tr>
</tbody>
</table>

Validity

The validity of a test refers to the quality of inferences that can be made by the test’s scores, that is, how well does the test measures and supports with empirical evidence the claims it makes for its use and applications.
CONTENT VALIDITY

• Behaviors included in the BIMAS Standard and BIMAS Flex
  – Meier's work presented earlier on change sensitive item selection
  – Input from colleagues in field testing studies over an 8 year period
• Structure of items into scales
  – Exploratory factor analysis
  – Rational/clinical analysis

BIMAS CLAIMS & EVIDENCE

• The BIMAS that can be used to identify emotional and behavior concerns of students using multiple sources of data..
• a multi-informant screening tool
  – Teacher
  – Parent
  – Self
• A progress monitoring tool
BIMAS as a Screening Tool

- Ratings offered by parents, teachers, students (self)
- Clinical samples were identified during the standardization process.
  - Screening criteria were applied thru the use of a Clinical Diagnostic Information Form.

THE BIMAS Clinical Samples
(N=1,355)

<table>
<thead>
<tr>
<th>Clinical Group</th>
<th>Teacher</th>
<th>Parent</th>
<th>Self</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>DB</td>
<td>123</td>
<td>22.9</td>
<td>70</td>
<td>15.0</td>
</tr>
<tr>
<td>ADHD</td>
<td>109</td>
<td>20.3</td>
<td>117</td>
<td>25.1</td>
</tr>
<tr>
<td>Anxiety</td>
<td>55</td>
<td>10.2</td>
<td>67</td>
<td>14.3</td>
</tr>
<tr>
<td>Depression</td>
<td>60</td>
<td>11.2</td>
<td>73</td>
<td>15.6</td>
</tr>
<tr>
<td>PDD</td>
<td>95</td>
<td>17.7</td>
<td>86</td>
<td>18.4</td>
</tr>
<tr>
<td>LD</td>
<td>45</td>
<td>8.4</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>DD</td>
<td>30</td>
<td>5.6</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Other</td>
<td>21</td>
<td>3.9</td>
<td>54</td>
<td>11.6</td>
</tr>
<tr>
<td>Total</td>
<td>538</td>
<td>100.0</td>
<td>467</td>
<td>100.0</td>
</tr>
</tbody>
</table>
The BIMAS as a Screening Tool
How were the data analyzed?......

• What is the % correct classification estimates for the....
  – Clinical
  – Non-clinical
  – Total sample
• Calculate other accuracy classification statistics

The Teachers as screening agents
BIMAS–T scores for Clinical sample

<table>
<thead>
<tr>
<th>BIMAS-T Standard Scales</th>
<th>Clinical Sample</th>
<th>Cohen's $d$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>$M$</td>
</tr>
<tr>
<td>Conduct</td>
<td>516</td>
<td>63.5</td>
</tr>
<tr>
<td>Negative Affect</td>
<td>537</td>
<td>66.4</td>
</tr>
<tr>
<td>Cognitive/Attention</td>
<td>538</td>
<td>66.6</td>
</tr>
<tr>
<td>Social</td>
<td>538</td>
<td>35.6</td>
</tr>
<tr>
<td>Academic Functioning</td>
<td>538</td>
<td>40.2</td>
</tr>
</tbody>
</table>

Note. Clinical $M$s ($SD$s) compared to values from the normative sample ($N = 1,361, M = 50, SD = 10$).

Cohen’s $d$ values of $|0.2| = \text{small effect}$, $|0.5| = \text{medium effect}$, and $|0.8| = \text{large effect}$.

Classification Accuracy of BIMAS–Teacher Scales

<table>
<thead>
<tr>
<th>Classification Accuracy Statistic</th>
<th>Full Range of Scores</th>
<th>Cut-Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Correct Classification</td>
<td>85.2%</td>
<td>82.5%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>83.5%</td>
<td>80.1%</td>
</tr>
<tr>
<td>Specificity</td>
<td>85.8%</td>
<td>83.4%</td>
</tr>
<tr>
<td>Positive Predictive Power</td>
<td>68.4%</td>
<td>64.9%</td>
</tr>
<tr>
<td>Negative Predictive Power</td>
<td>93.4%</td>
<td>91.6%</td>
</tr>
</tbody>
</table>
The Parents as screening agents

### BIMAS–P
Clinical vs. Non-Clinical samples

<table>
<thead>
<tr>
<th>BIMAS–P Standard Scales</th>
<th>Clinical Sample</th>
<th>Cohen’s $d$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>$M$</td>
</tr>
<tr>
<td>Conduct</td>
<td>467</td>
<td>60.3</td>
</tr>
<tr>
<td>Negative Affect</td>
<td>467</td>
<td>61.5</td>
</tr>
<tr>
<td>Cognitive/Attention</td>
<td>467</td>
<td>60.7</td>
</tr>
<tr>
<td>Social</td>
<td>467</td>
<td>38.4</td>
</tr>
<tr>
<td>Academic Functioning</td>
<td>467</td>
<td>40.4</td>
</tr>
</tbody>
</table>

Note. Clinical Ms (SDs) compared to values from the normative sample ($N = 1,400, M = 50$, $SD = 10$).

Cohen’s $d$ values of $|0.2| =$ small effect, $|0.5| =$ medium effect, and $|0.8| =$ large effect.
## Classification Accuracy of BIMAS–Parent Scales

<table>
<thead>
<tr>
<th>Classification Accuracy Statistic</th>
<th>Full Range of Scores</th>
<th>Cut-Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Correct Classification</td>
<td>78.3%</td>
<td>78.6%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>80.1%</td>
<td>73.4%</td>
</tr>
<tr>
<td>Specificity</td>
<td>77.7%</td>
<td>80.3%</td>
</tr>
<tr>
<td>Positive Predictive Power</td>
<td>54.6%</td>
<td>55.4%</td>
</tr>
<tr>
<td>Negative Predictive Power</td>
<td>92.1%</td>
<td>90.1%</td>
</tr>
</tbody>
</table>

The Students as screening agents
**BIMAS–Self ratings**  
**Clinical vs. Non-Clinical**

<table>
<thead>
<tr>
<th>BIMAS-P Standard Scales</th>
<th>Clinical Sample</th>
<th>Cohen's d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
</tr>
<tr>
<td>Conduct</td>
<td>350</td>
<td>57.3</td>
</tr>
<tr>
<td>Negative Affect</td>
<td>350</td>
<td>59.2</td>
</tr>
<tr>
<td>Cognitive/Attention</td>
<td>350</td>
<td>57.3</td>
</tr>
<tr>
<td>Social</td>
<td>350</td>
<td>41.4</td>
</tr>
<tr>
<td>Academic Functioning</td>
<td>350</td>
<td>42.3</td>
</tr>
</tbody>
</table>

Note. Clinical Ms (SDs) compared to values from the normative sample (N = 703, M = 50, SD = 10).

Cohen’s d values of |0.2| = small effect, |0.5| = medium effect, and |0.8| = large effect.

---

**Classification Accuracy of BIMAS–Self-Report Scales**

<table>
<thead>
<tr>
<th>Classification Accuracy Statistic</th>
<th>Full Range of Scores</th>
<th>Cut-Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Correct Classification</td>
<td>71.5%</td>
<td>71.8%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>76.3%</td>
<td>67.1%</td>
</tr>
<tr>
<td>Specificity</td>
<td>69.1%</td>
<td>74.1%</td>
</tr>
<tr>
<td>Positive Predictive Power</td>
<td>55.3%</td>
<td>56.5%</td>
</tr>
<tr>
<td>Negative Predictive Power</td>
<td>85.3%</td>
<td>81.9%</td>
</tr>
</tbody>
</table>
The BIMAS as a Progress Monitoring Tool

Progress Monitoring

- Documenting and Measuring Change/progress
  - BIMAS Standard
  - BIMAS Flex
Progress Monitoring with the BIMAS Standard

- Numerous Group reports for each BIMAS scale…across Universal Assessments by:
  - School;
  - Grade ;
  - Rater
  - Service Code (reg educ, spec educ, Title 1)
  - Risk level across Universal assessments

More on this during the demonstration of the BIMAS.

BIMAS-Scores for Progress and Outcome Monitoring

- Type of scores
  - % percentages for risk categories
  - % percentiles
  - T-scores for all 5 scales

  - GOAL...
    - DECREASE Behavior Concerns scores
    - INCREASE Adaptive behavior scores
BIMAS-Scores for Progress and Outcome Monitoring

- Progress vs. Outcome monitoring
- Several methods, but no consensus
- BIMAS indexes of CHANGE
  - Tabular presentation or visual displays,
  - Effect size (ES) estimates,
  - the reliable change index (RCI).

BIMAS Progress Monitoring data

<table>
<thead>
<tr>
<th>Scales</th>
<th>Fall 2016 09/2016 Results</th>
<th>Spring 2016 03/15/16 Results</th>
<th>Winter 2016 01/04/16 Results</th>
<th>Fall 2015 10/01/15 Results</th>
<th>Spring 2015 03/01/15 Results</th>
<th>Winter 2015 01/03/15 Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conduct</td>
<td>66</td>
<td>64</td>
<td>57</td>
<td>61</td>
<td>61</td>
<td>61</td>
</tr>
<tr>
<td>Negative Affect</td>
<td>77</td>
<td>75</td>
<td>76</td>
<td>75</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>Cognitive/Attention</td>
<td>58</td>
<td>61</td>
<td>52</td>
<td>61</td>
<td>61</td>
<td>61</td>
</tr>
<tr>
<td>Social</td>
<td>24</td>
<td>26</td>
<td>25</td>
<td>24</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Academic Functioning</td>
<td>30</td>
<td>31</td>
<td>32</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
</tbody>
</table>

BIMAS-2 will be able to maintain data for as many years as the student attends the school district.
BIMAS Visual Displays

BIMAS Effect size estimates

Table 5.8. Effect Size Interpretations for Individual Clients on the BIMAS Standard

<table>
<thead>
<tr>
<th>Effect Size</th>
<th>Interpretation for Behavioral Concern Scales</th>
<th>Interpretation for Adaptive Scales</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ -1.50</td>
<td>Much Improved</td>
<td>Much Worse</td>
</tr>
<tr>
<td>-1.50 to -1.49</td>
<td>Improved</td>
<td>Worse</td>
</tr>
<tr>
<td>-1.50 to 0.50</td>
<td>No Change</td>
<td>No Change</td>
</tr>
<tr>
<td>0.51 to 1.49</td>
<td>Worse</td>
<td>Improved</td>
</tr>
<tr>
<td>≥ 1.50</td>
<td>Much Worse</td>
<td>Much Improved</td>
</tr>
</tbody>
</table>
BIMAS Effect size report

Reliable Change Index (RCI) (Jacobson & Truax, 1991).

- Has a clinically significant change occurred for a student?
- the RCI formula employs an individual’s pre and post scores, the pretest standard deviation for a group of scores, and a reliability estimate for the test.
Progress Monitoring with the BIMAS

Anger Management Study

Anger Management Treatment Study

\( N = 46 \) (ages 12 to 18 years)
Gender: 32 males and 14 females.
Race/Ethnicity:

- 30 African American,
- 2 Hispanic &
- 14 Caucasian students

- BIMAS scores showed **good sensitivity to change** in response to intervention in the theoretically expected direction.
### Pre-Post Intervention Performance of an Anger Management Treatment Group: BIMAS–Teacher T-scores

- Statistically significant change in theoretically expected direction

<table>
<thead>
<tr>
<th>BIMAS-T Scale</th>
<th>Pre-Test</th>
<th>Post-Test</th>
<th>t</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conduct</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>65.9</td>
<td>59.3</td>
<td>9.2</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>4.8</td>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td>Negative Affect</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>63.0</td>
<td>53.9</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>10.7</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td>Cognitive/ Attention</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>63.3</td>
<td>55.3</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>6.6</td>
<td>6.9</td>
<td></td>
</tr>
<tr>
<td>Social</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>30.0</td>
<td>34.4</td>
<td>-3.4</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>5.5</td>
<td>7.2</td>
<td></td>
</tr>
<tr>
<td>Academic Functioning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>41.9</td>
<td>45.7</td>
<td>-5.2</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>4.9</td>
<td>4.1</td>
<td></td>
</tr>
</tbody>
</table>

- Note. N = 46. All ts significant at p < .01.
- Cohen’s d values of |0.2| = small effect, |0.5| = medium effect, and |0.8| = large effect.

### Pre-Post Intervention Performance of an Anger Management Treatment Group: BIMAS–Parent T-scores

- Statistically significant change in theoretically expected direction

<table>
<thead>
<tr>
<th>BIMAS-P Scale</th>
<th>Pre-Test</th>
<th>Post-Test</th>
<th>t</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conduct</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>66.6</td>
<td>53.5</td>
<td>12.7*</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>5.8</td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td>Negative Affect</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>60.8</td>
<td>47.1</td>
<td>10.4*</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>9.5</td>
<td>6.9</td>
<td></td>
</tr>
<tr>
<td>Cognitive/ Attention</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>59.4</td>
<td>49.5</td>
<td>10.3*</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>5.4</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>Social</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>31.7</td>
<td>37.5</td>
<td>-4.7*</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>4.9</td>
<td>6.9</td>
<td></td>
</tr>
<tr>
<td>Academic Functioning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>40.0</td>
<td>45.7</td>
<td>-7.3*</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>4.4</td>
<td>4.1</td>
<td></td>
</tr>
</tbody>
</table>

- Note. N = 46. All ts significant at p < .01.
- Cohen’s d values of |0.2| = small effect, |0.5| = medium effect, and |0.8| = large effect.
Pre-Post Intervention Performance of an Anger Management Treatment Group: BIMAS–Self-Report T-scores

- Statistically significant change in theoretically expected direction

<table>
<thead>
<tr>
<th>BIMAS-SR Scale</th>
<th>Pre-Test</th>
<th>Post-Test</th>
<th>t</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conduct</td>
<td>M</td>
<td>65.5</td>
<td>52.2</td>
<td>13.8*</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>5.4</td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td>Negative Affect</td>
<td>M</td>
<td>59.2</td>
<td>44.6</td>
<td>11.5*</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>9.8</td>
<td>6.5</td>
<td></td>
</tr>
<tr>
<td>Cognitive/ Attention</td>
<td>M</td>
<td>62.7</td>
<td>49.6</td>
<td>12.9*</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>6.6</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>Social</td>
<td>M</td>
<td>35.1</td>
<td>39.5</td>
<td>−4.5*</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>6.2</td>
<td>4.8</td>
<td></td>
</tr>
<tr>
<td>Academic Functioning</td>
<td>M</td>
<td>38.9</td>
<td>46.2</td>
<td>−10.1*</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>5.0</td>
<td>3.0</td>
<td></td>
</tr>
</tbody>
</table>

- Note. N = 46. All ts significant at p < .01.
- Cohen’s d values of |0.2| = small effect, |0.5| = medium effect, and |0.8| = large effect.

Intervention Study with ADHD children

Cherise Lerew, Ph.D.
Achilles N Bardos, Ph.D.
The Intervention study

- children with ADHD have poor executive functions.

- children with ADHD perform low in planning on the CAS (Naglieri and Reardon, 1991)

- Past intervention research with CAS used students with LD, NOT children with ADHD

- The Planning Facilitation Method has been used with math and reading comprehension, but never with behavior

Measures Used

**Math**
- Math worksheets
- Basic Achievement Skills Inventory (BASI)

**Reading**
- Qualitative Reading Inventory-Third (QRI-3)
- Basic Achievement Skills Inventory (BASI)

**Behavior**
- BIMAS early piloting of items
- Devereux Scales of Mental Disorders (DSMD)
Math Results

- Math scores improved for all students
- Percent change ranged from 13% to 185%
- All students displayed stable baselines (before intervention was introduced)
- 5 out of 6 students showed an upward trend in the intervention phase

Example:
Reading Comprehension Results

- All students displayed a positive gain in reading
- Percent change ranged from 1 to 10%
- All students displayed stable baselines (before intervention was introduced)
- 3 out of 6 students displayed a trend during the intervention phase

Example:

Weekly Behavior Scale (BIMAS) Results

- All students appeared to have a **decrease in behaviors** from baseline to intervention phases on a weekly rating scale (BIMAS)
- Example:
Pre- and Post Test Behavior Results

• 5 of the 6 students displayed a **significant decrease** in behaviors from pre- to post-testing on overall behavior (Total DSMD score)

Findings...

• The planning facilitation intervention **improved academic achievement** for children with ADHD

• The intervention **decreased overall behavior** for all students on weekly rating scales and 5 out of 6 children on pre- and posttest measures of behavior
Concurrent Validity - Conner (CBRS)

Table 11.5. Correlations Between the BIMAS Standard Conduct Scale T-scores and Relevant Conners CBRS Scales

<table>
<thead>
<tr>
<th>Conners CBRS Scale</th>
<th>Teacher</th>
<th>Parent</th>
<th>Self-Report</th>
<th>M</th>
<th>SD</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>N</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Defiant/Aggressive Behaviors</td>
<td>.52</td>
<td>.73</td>
<td>.04</td>
<td>51.1</td>
<td>12.0</td>
<td>51.7</td>
<td>13.7</td>
<td>53.3</td>
<td>14.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Violence Potential Indicator</td>
<td>.54</td>
<td>.74</td>
<td>.04</td>
<td>50.1</td>
<td>10.1</td>
<td>52.7</td>
<td>13.8</td>
<td>52.4</td>
<td>12.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSM-IV-TR Conduct Disorder</td>
<td>.50</td>
<td>.68</td>
<td>.02</td>
<td>50.0</td>
<td>10.3</td>
<td>51.1</td>
<td>13.9</td>
<td>53.2</td>
<td>14.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSM-IV-TR Oppositional Defiant</td>
<td>.49</td>
<td>.78</td>
<td>.02</td>
<td>13.7</td>
<td>13.9</td>
<td>50.6</td>
<td>12.5</td>
<td>51.0</td>
<td>11.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Teacher, N = 112, Parent, N = 117, Self-Report, N = 100. All correlations significant at p < .01 (2-tailed).

Table 11.7. Correlations Between the BIMAS Standard Negative Affect Scale T-scores and Relevant Conners CBRS Scales

<table>
<thead>
<tr>
<th>Conners CBRS Scale</th>
<th>Teacher</th>
<th>Parent</th>
<th>Self-Report</th>
<th>M</th>
<th>SD</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>N</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional Distress</td>
<td>.47</td>
<td>.70</td>
<td>.04</td>
<td>48.7</td>
<td>8.4</td>
<td>50.6</td>
<td>13.7</td>
<td>51.0</td>
<td>13.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSM-IV-TR Major Depressive Episode</td>
<td>.30</td>
<td>.62</td>
<td>.06</td>
<td>49.3</td>
<td>9.1</td>
<td>50.0</td>
<td>12.1</td>
<td>52.0</td>
<td>11.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIMAS Negative Affect Scale M</td>
<td>48.3</td>
<td>49.2</td>
<td>.50.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIMAS Negative Affect Scale SD</td>
<td>7.9</td>
<td>10.1</td>
<td>10.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Teacher, N = 112, Parent, N = 117, Self-Report, N = 100. All correlations significant at p < .01 (2-tailed).

Coming up next... the BIMAS Online!!!

www.achillesbardos.com
Summary/Strengths of BIMAS

• BIMAS: empirically-based; sensitive to change (excellent for RtI) √
• Standard & Flex √
• Good Normative data & Good Psychometric Properties √
• Powerful Web-based Interface √
  – Web-based administration and scoring options √
  – Wide Selection of Informative Web-based Reports
• But before we close!!!!! √

Academic-Behavior Connection

• “Viewed as outcomes, achievement and behavior are related; viewed as causes of each other, achievement and behavior are unrelated. In this context, teaching behavior as relentlessly as we teach reading or other academic content is the ultimate act of prevention, promise, and power underlying PBS and other preventive interventions in America’s schools.”
