IS IT WORKING? THE USE OF CHANGE-SENSITIVE MEASURES TO EVALUATE SCHOOL-BASED MODELS OF BEHAVIORAL SUPPORT

Scott Meier, PhD, SUNY Buffalo

James McDougal, PsyD, SUNY Oswego

Jill Snyder, PhD, Boston Public Schools

SYMPOSIUM AGENDA

- Scott Meier, Assessing school based mental health services: Progress monitoring and outcome assessment with child and adolescents
- Jim McDougal, The use of integrated data in comprehensive school-based models of RtI/ MTSS
- Jill Snyder, Data-based decision making within a Comprehensive Behavioral Health Model

Treatment Failure and Clinical Feedback

- High rates of treatment failure reported in the literature (e.g., between 10% and 50% of all adult clients, Persons & Mikami, 2002).
- Studies show that providing clinicians with regular feedback about client progress reduces failure rates

Progress monitoring and outcome assessment…

- …refer to using tests to measure the effects of counseling and psychotherapy
  - Outcome assessment: Pre and post completion of therapy
  - Progress monitoring: Measurement during course of therapy
- But which measures to use?
- Traditional measures such as the MMPI are too long to be used repeatedly to assess change

Another option: Change-sensitive tests

- Construct validity of outcome assessments depends upon their sensitivity to change
- Research suggests that items and tests vary in their ability to detect treatment effects
- Research finds that items selected by factor analysis are different from items shown to be treatment sensitive

Weinstock and Meier (2003), for example…

- Examined change in a client database of 615 students at a university counseling center who completed a 56-item symptom checklist at intake and termination
- They subjected the intake responses to a factor analysis and found a single factor, assessed by 8 items, that they labeled “Social Anxiety”

Weinstock & Meier (cont.)

- They also subjected the intake-termination scores to change analyses and found 25 change-sensitive items
- While the coefficient alphas were comparable for the 8-item factor analytic scale and the 25-item change-sensitive scale (about .90), the Effect Size (ES) for the change sensitive scale equaled .74, compared to .47 for the factor analytic scale.

Steps to create a change-sensitive test:

1. Select items on basis of theory & research
2. Aggregate at appropriate levels
3. Assess range of pretest item scores
4. Detect change in items scores post intervention
5. Change occurs in expected direction?

Steps (cont.)

6. Change in item scores between intervention and comparison groups after intervention?
7. Intake differences between intervention and comparison groups?
8. Relations between item scores and systematic errors?
9. Cross-validate to minimize chance effects

Change-Sensitive Test Scores May Be Used to Increase Power

- Power refers to the ability to detect effects produced by an intervention, such as psychotherapy, education, and employee training.
- Many of the effects produced are small, so we need measures with sufficient power to detect them if they occur
- You can increase power by creating measures designed to detect these effects

Then, you can use valid test scores for clinical feedback…

- Provide feedback to therapist or client to adjust intervention
- Considerable research suggests that clinicians are less skilled at judging clients' response to interventions
- This may be particularly true with failing clients
- Persons and Mikami (2002) suggested that research indicates that as many as 50% of clients fail to improve
- Other research indicates that many clinicians fail to adjust their interventions even when they know clients are failing to improve

Lambert’s approach to feedback

- Lambert and colleagues have conducted a series of studies that indicate that you can reduce the failure rate of clients by providing clinicians with basic information about lack of progress
- Effect not as strong for persons making progress

Determining the effects of psychotherapy

- Long-standing question is <what> to assess in psychotherapy outcome (content validity)
- Research suggests that depression and anxiety items show the largest change on comprehensive outcome measures
- This implies that all outcome measures should assess depression and anxiety or they will miss detecting the largest effects (decreased power)

Program Evaluation Implications

- Many agencies and organizations have demands to show their work is effective
- Program evaluations need measures that are powerful enough to detect these effects
- So the use of change-sensitive measures would seem necessary to provide a fairer test of finding any effects

Summary and Implications

- So we have a method for increasing our power to detect “treatment” (broadly defined) effects
- With this method, we are looking for items that show change in treatment conditions but stability in the absence of a treatment
- But we still have the issue of error involved with how individuals make sense of and provide data related to test questions, instructions, format, and so forth

What is the BIMAS?

A brief behavior rating scale designed for :

Screening-

- detect students in need of further assessment
- identify areas of behavior concerns and adaptive skills

Progress Monitoring of:

- System-wide interventions (Tier I- PBIS; SEL)
- Small groups interventions (Tier II)
- Interventions for individuals (Tier III)

A multi-informant web-based delivered assessment system

- RATINGS are offered by:
  - Parents
  - Teacher
  - Self (12 -18 yrs old)
  - Clinician

BIMAS theoretical foundation

- Utilized Meier's approach to construct the scale using his Intervention Item Selection Rules (IISR) procedures
- Data from a variety of clinical and school settings (e.g., Meier, 2004, 2000, 1998).
- IISR procedures lead to scales with
  - demonstrated larger treatment effect sizes
  - adequate reliability estimates.

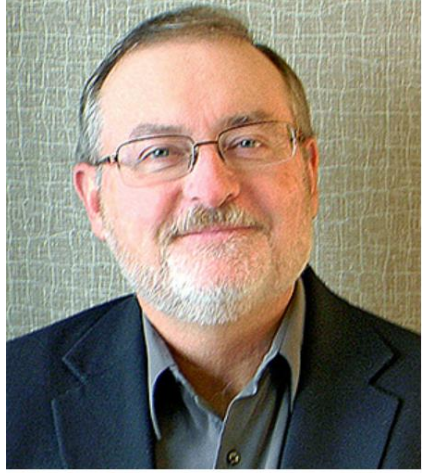
Intervention Item Selection Rules (**IISRs**; Meier 1997, 1998, 2000, 2004)

- BIMAS: developed using empirically derived model for designing change-sensitive measures to assess RTI: IIRSs
- "State" scale as opposed to a "Trait" scale
- BIMAS developed using clinical & school samples in field settings (rare combo in psychotherapy research)
- Identified constructs that change as a result of emotional and behavioral intervention
- Edison metaphor of developing light bulb

IISRs (cont'd)

1. Be grounded in theory- to provide for interpretation and minimize the effects of chance;
2. Be aggregated across individuals- to reduce the effects of random error;
3. Not evidence ceiling and floor effect;
4. Evidence change after a psychosocial intervention;
5. Change in the theoretically expected direction;
6. Evidence change relative to control and comparison groups;
7. Show no difference at pre-intervention;
8. Have no relation to relevant systematic errors; and
9. Be subject to cross-validation studies.

(Meier 1997, 1998, 2000, 2004)



Dr. Scott Meier
Intervention Item Selection Rules

The central philosophy of the IISRs is that **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).

The BIMAS Scales -Standard Form

Behavioral Concern Scales:

- **Conduct**
- **Negative Affect**
- **Cognitive/ Attention**

Adaptive Scales:

- **Social**
- **Academic Functioning**

Application of the BIMAS within the RTI Framework

1. **Screening-** *To detect students in need of further assessment and to identify their respective areas of need.*
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.*

Progress Monitoring with the BIMAS-2

How to build small group and/or Individual student progress monitoring plans

PROGRESS MONITORING

INDIVIDUAL or small GROUP

- WE NEED TO:
  - Identify the Intervention needed
  - Determine the goals and behaviors that will be Monitored?
  - How often should they be monitored
  - How long should the monitoring be?
  - Who is going to provide us with feedback?

Progress Monitoring with BIMAS

- Documenting and Measuring Change/Progress
  - BIMAS Standard
  - BIMAS Flex

- A Case Study
  - Anger management group

Anger Management Treatment Study

N = 46 (ages 12 to 18 years)

Gender: 32 males and 14 females.

Race/Ethnicity:

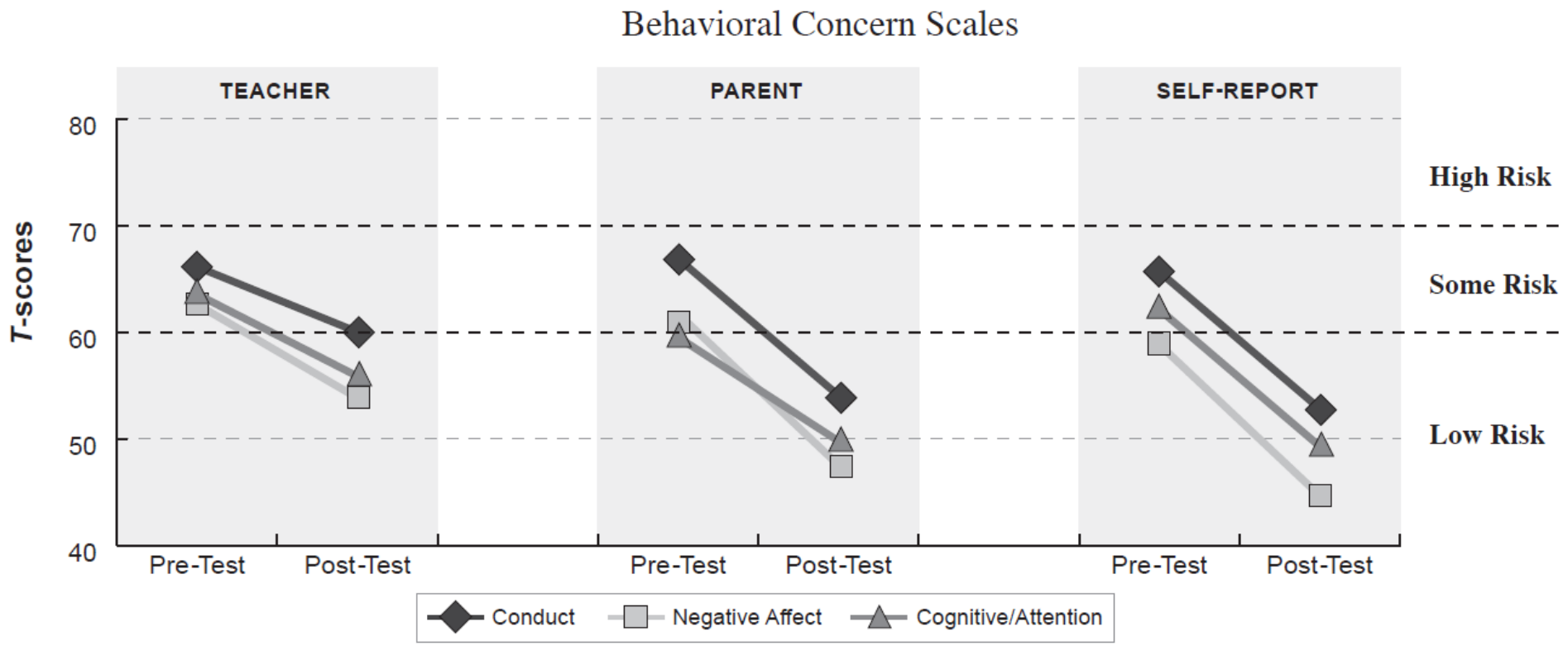
30 African American,

2 Hispanic &

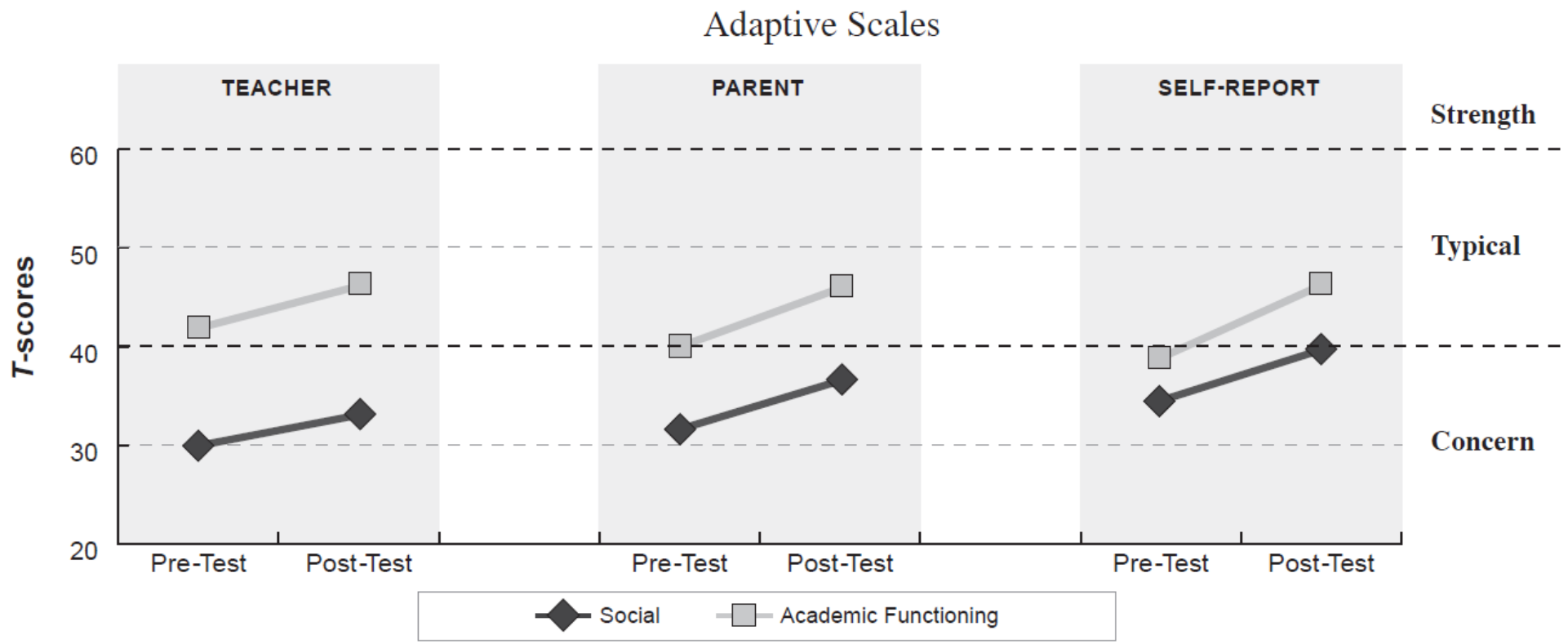
14 Caucasian students

- BIMAS scores showed **good sensitivity to change** in response to intervention in theoretically expected direction

Anger Management Treatment Group: Pre- to Post-Treatment BIMAS *T*-scores



Anger Management Treatment Group: Pre- to Post-Treatment BIMAS *T*-scores



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

Statistically significant change in theoretically expected direction

BIMAS-T Scale		Pre-Test	Post-Test	*t*	Cohen's *d*
Conduct	*M*	65.9	59.3	**9.2***	**1.5**
	SD	4.8	3.7		
Negative Affect	*M*	63.0	53.9	**6.6***	**1.0**
	SD	10.7	7.7		
Cognitive/ Attention	*M*	63.3	55.3	**7.3***	**1.2**
	SD	6.6	6.9		
Social	*M*	30.0	34.4	**−3.4***	**−0.7**
	SD	5.5	7.2		
Academic Functioning	*M*	41.9	45.7	**−5.2***	**−0.8**
	SD	4.9	4.1		

Note. $N = 46$. All *t*s 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

BIMAS-P Scale		Pre-Test	Post-Test	*t*	Cohen's *d*
Conduct	*M*	66.6	53.5	**12.7***	**2.6**
	SD	5.8	4.3		
Negative Affect	*M*	60.8	47.1	**10.4***	**1.7**
	SD	9.5	6.9		
Cognitive/ Attention	*M*	59.4	49.5	**10.3***	**2.0**
	SD	5.4	4.6		
Social	*M*	31.7	37.5	**−4.7***	**−1.0**
	SD	4.9	6.9		
Academic Functioning	*M*	40.0	45.7	**−7.3***	**−1.3**
	SD	4.4	4.1		

Note. $N = 46$. All *t*s 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

BIMAS-SR Scale		Pre-Test	Post-Test	*t*	Cohen's *d*
Conduct	*M*	65.5	52.2	**13.8***	**2.8**
	SD	5.4	3.8		
Negative Affect	*M*	59.2	44.6	**11.5***	**1.8**
	SD	9.8	6.5		
Cognitive/ Attention	*M*	62.7	49.6	**12.9***	**2.4**
	SD	6.6	4.2		
Social	*M*	35.1	39.5	**−4.5***	**−0.8**
	SD	6.2	4.8		
Academic Functioning	*M*	38.9	46.2	**−10.1***	**−1.8**
	SD	5.0	3.0		

Note. $N = 46$. All *t*s significant at $p < .01$.

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

THE USE OF INTEGRATED DATA IN COMPREHENSIVE MODELS OF RTI/ MTSS IN THE SCHOOLS

James McDougal, PsyD, SUNY Oswego

EMOTIONAL AND BEHAVIORAL DISORDERS

- About 20% of children present themselves with diagnosable disorders (i.e., U.S. Department of Health and Human Services, 1999)
- 3–6% of children with serious and chronic disorders (Kauffman, 1997)
- Progression of disorders is very predictable
  - Externalizing behaviors (severe tantrums, disobedience)
  - Internalizing difficulties (anxiety, depression, suicide)

Negative Long Term Outcomes

- EBD students have the poorest outcomes of the "high incident" disability groups.
- Drop-out rate over 50%
- After school 40% are unemployed with no additional training/ education
- 50% are arrested within 5 years of leaving school
- Of EBD drop-outs this figure exceeds 70% !!

(e.g., see Quinn & McDougal, 1998)

Negative Long Term Outcomes

- ¾ of students with EBD have been suspended or expelled from school (National Longitudinal Transition Study 2 [NLTS2], 2004).

- rates increase from elementary to middle to high school transition.

Negative Long Term Outcomes

- 75% of children with significant externalizing behaviors (severe tantrums, disobedience) eventually engage in predictable and serious law breaking and antisocial behavior (e.g., Reid, 1993).
- Internalizing disorders (anxiety, depression) result in increased rates of pathology and lower rates of socialization and academic attainment (Hops, Walker, & Greenwood, 1988).
- Internalizing & externalizing difficulties lead to increased rates of substance abuse and lowered academic attainment.

SUICIDE

- Nearly 30,000 Americans commit suicide every year.
- In the U.S., suicide rates are highest during the spring.
- Suicide is the 3rd leading cause of death for 15 to 24-year-olds and 2nd for 24 to 35-year-olds.
- On average, 1 person commits suicide every 16.2 minutes.
- Each suicide intimately affects at least 6 other people.

Early Identification

- early identification and intervention with children who are at risk for EBD appear to be the “most powerful course of action for ameliorating life-long problems associated with children at risk for EBD” (Hester et al., 2004)

- Younger children are more likely to be responsive to and maintain the positive outcomes from early prevention and intervention programs (Bailey, Aytch, Odom, Symons, & Wolery, 1999

RTI AND BEHAVIOR

RATIONALE AND NEED FOR UNIVERSAL BEHAVIOR SCREENING

- There is a strong link between behavior/emotions and academic performance

- How do we define health?
  - Most schools screen for vision, hearing, speech, and academic achievement
  - Behavior/Emotional screening occurs in less than 2% of districts across the U.S.

RTI AND BEHAVIOR

RATIONALE AND NEED FOR UNIVERSAL BEHAVIOR SCREENING

- Screeners for children are mostly done in primary care settings; we miss many children for early identification and intervention (Pagano et al., 2000).
- Screening measures for children are frequently focused on one disorder to the exclusion of others, an approach that neglects large numbers of children who have problems other than the target condition (August et al., 1992; Taylor et al., 2000; Matthey & Petrovski, 2002).

RTI AND BEHAVIOR

RATIONALE AND NEED FOR UNIVERSAL BEHAVIOR SCREENING

- Teachers accurately identify young children at high risk of academic and behavioral problems related to school adjustment with a great deal of accuracy (Taylor et al., 2000).

- Schools are the ideal setting for large-scale, broad based mental health screening of children and adolescents (Wu et al., 1999).

SEARCH FOR THE BEHAVIORAL MEASURE FOR SCREENING AND PM

DENO, MIRKIN, & CHAING (1982)

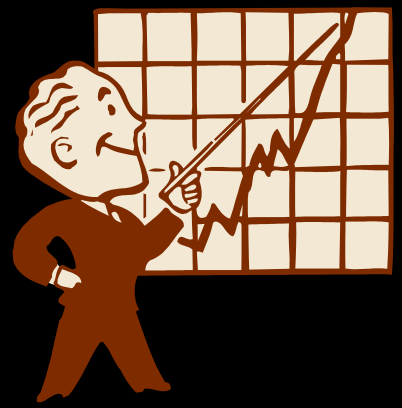
- Technically adequate (reliability, validity) for Screening and PM
- Sensitive to change
- Suitable for frequent administration (brief)
- Reflects overall improvement
- Easy to administer
- Useful across a range of interventions/ populations

BEHAVIOR SCREENING AND PROGRESS MONITORING TECHNIQUES

- Behavior Screening techniques allow for quick identification of students for additional assessment

- A monitoring technique allows you to measure change over time- broad band and narrow band

- Baseline levels, intervention goals, and the monitoring technique should all be consistent!



Rating Scales

- ✓ Able to measure low frequency behavior
- ✓ Less time consuming than Direct observation
- ✓ Require less training to implement
- ✓ Can be used to assess multiple settings and informants
- ✓ Can provide broad and narrow band assessments

Shortcomings of Traditional Behavior Rating Scales

- Behavior Rating Scales were develop for diagnostic purposes- identifying individuals in different groupings
- Diagnostic tools are developed to capitalize on discrimination of individual differences
- These differences are usually “trait-related” and not likely to evidence short term change
- Most diagnostic scales are time consuming – meet with resistance

Shortcomings of Traditional Behavior Rating Scales

- Impractical when a number of data collection points are needed.
- Not designed to be sensitive to change
- Behavior monitoring parallels the evolution of CBM within RTI Emphasis on reliable and valid procedures for screening and progress monitoring.

Recent Developments: Creating Comprehensive Screening and Progress Monitoring Systems

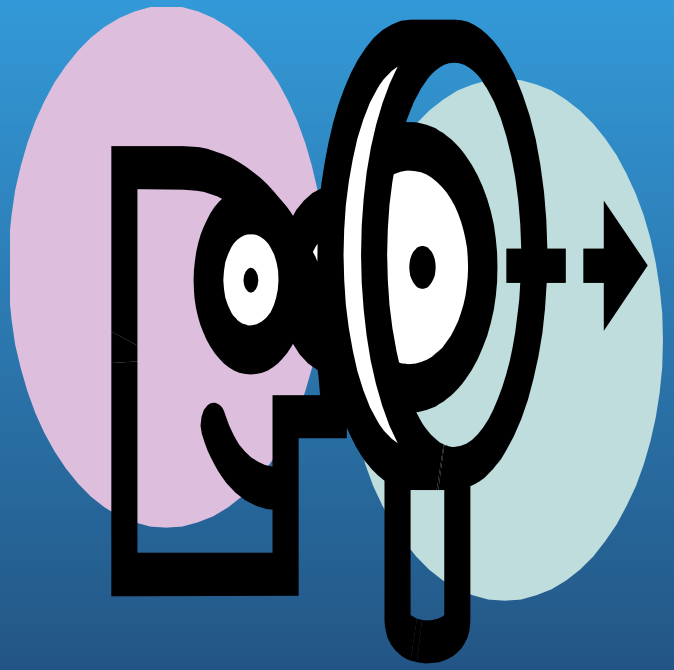
- Need to measure:
  - Broad domains of emotional and behavioral disorders
  - Related areas of adaptation or impairment, and
  - Specific targets for intervention (narrow band)

Volpe & Gadow, 2010

Evolution of change sensitive measures geared toward 3 tier models

Purpose

- Screening,
- Progress Monitoring,
- Program Evaluation



Behavior Intervention Monitoring Assessment System

By James L. McDougal, Psy. D., Achilles N. Bardos, Ph.D., & Scott T. Meier, Ph.D.

Creating “Change Sensitive” Measures

Based on the Work of
Dr. Scott Meier

Intervention Item Selection Rules:

A model for change sensitive scale development

Intervention Item Selection Rules (IIRS): Model Overview

1. Based on Theory
2. Aggregate Items
3. Avoid Ceiling Effect
4. Detect Change
5. Expected Direction?
6. Relative to Comparison?
7. No Pre-Test Difference
8. Systematic Errors dropped
9. Cross-Validate

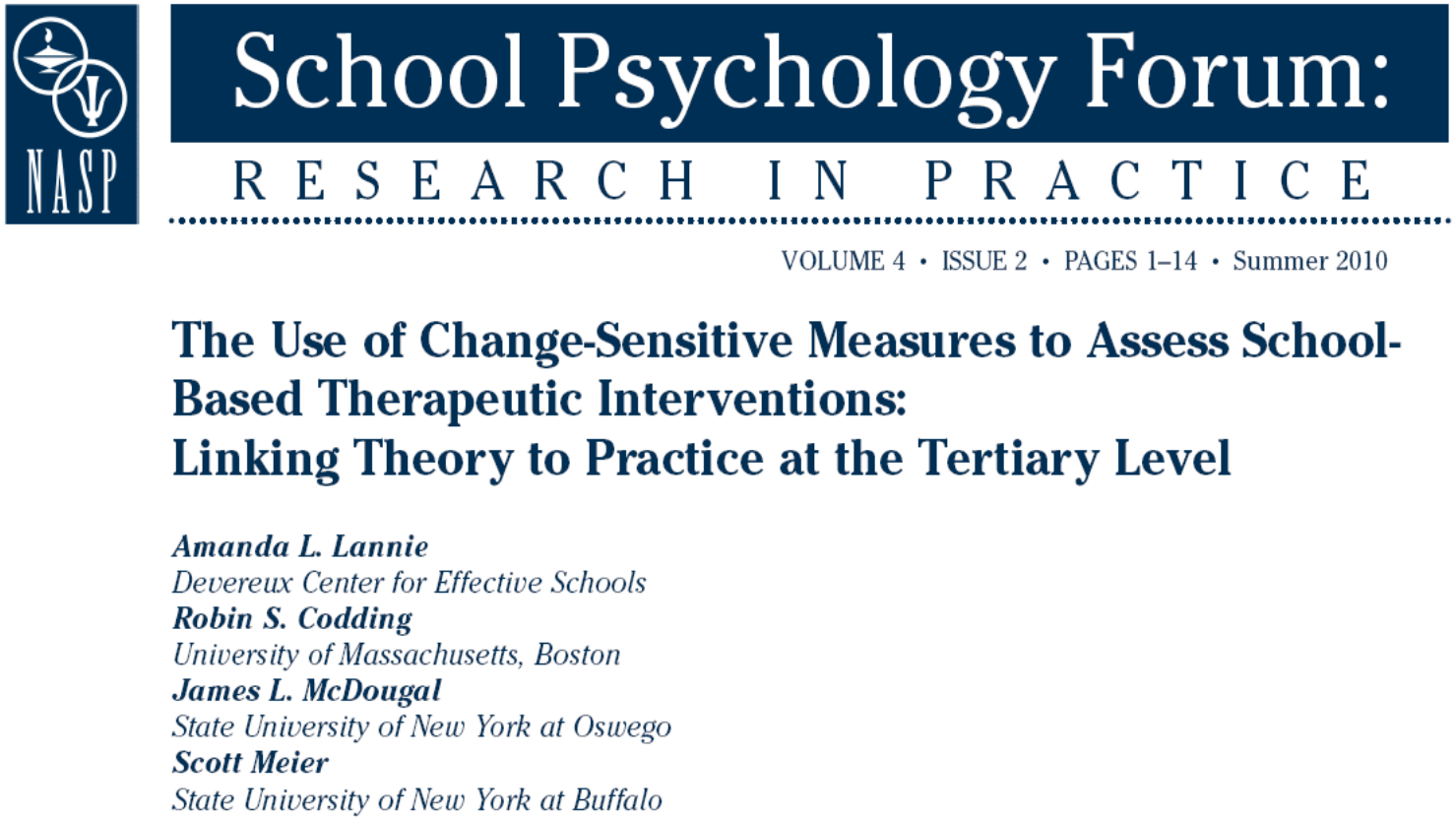
Canadian Journal of
School Psychology
Volume XX Number X
Month XXXX XX-XX

10.1177/0829573507307693
http://cjsp.sagepub.com
hosted at
http://online.sagepub.com

Development of a Change-Sensitive Outcome Measure for Children Receiving Counseling

Scott T. Meier
University at Buffalo
James L. McDougal
State University of New York at Oswego
Achilles Bardos
University of Northern Colorado

RtI & Behavior



School Psychology Forum:
RESEARCH IN PRACTICE

VOLUME 4 • ISSUE 2 • PAGES 1–14 • Summer 2010

The Use of Change-Sensitive Measures to Assess School-Based Therapeutic Interventions: Linking Theory to Practice at the Tertiary Level

Amanda L. Lannie
Devereux Center for Effective Schools
Robin S. Codding
University of Massachusetts, Boston
James L. McDougal
State University of New York at Oswego
Scott Meier
State University of New York at Buffalo

What is the BIMAS?

1. **Screening-** *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

The BIMAS can be used by school-based mental health providers

- Public/private organizations providing school or community-based intervention programs
- community mental health agencies
- managed care agencies (HMOs)
- other providers who require an outcome measure sensitive to short term therapeutic gains

Format of the BIMAS

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

The BIMAS Scale Structure

BIMAS Standard

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:

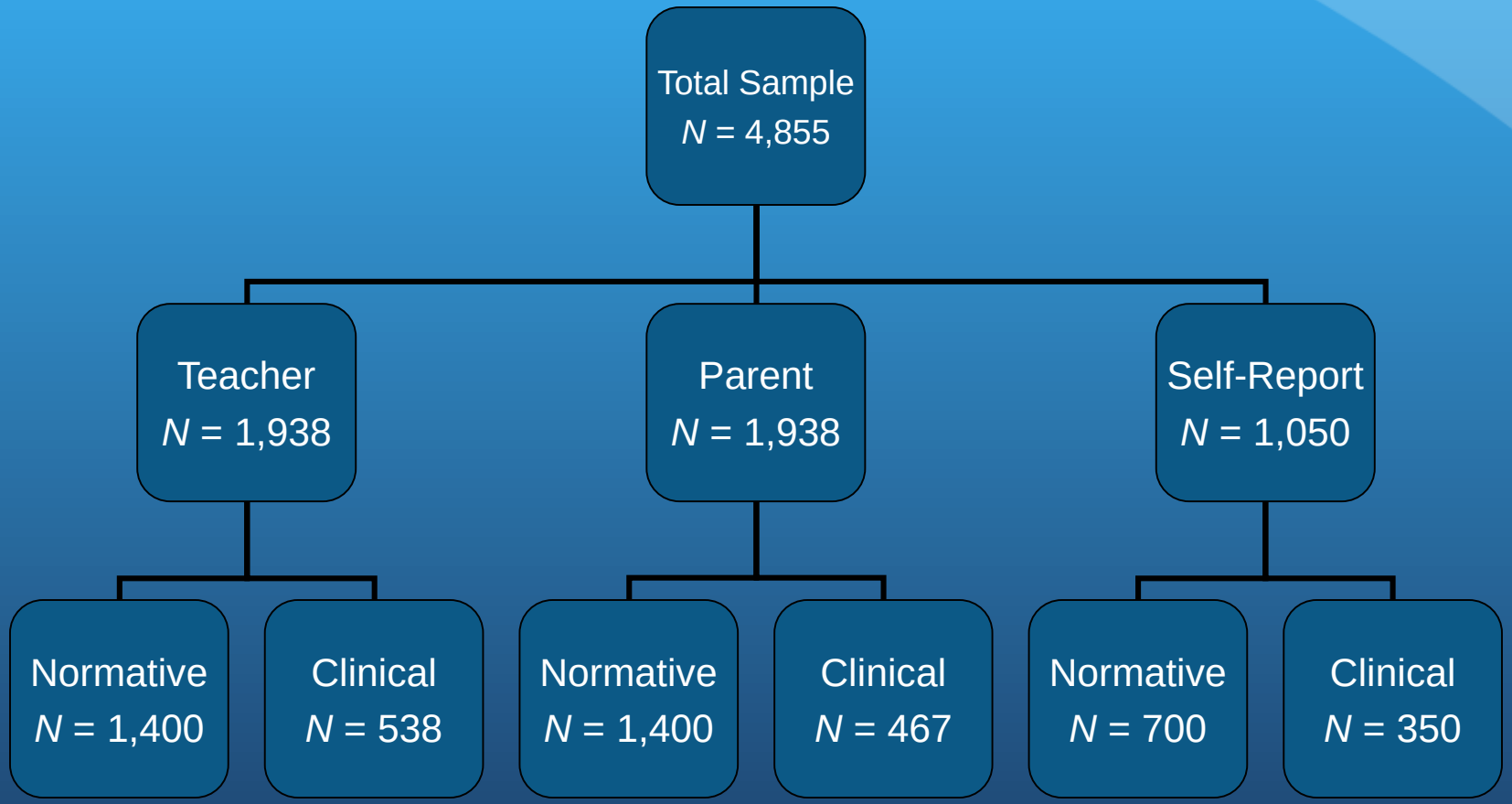
- **Social** — social functioning, friendship maintenance, communication
- **Academic Functioning** — academic performance, attendance, ability to follow directions

BIMAS Flex features

- List of specific behavioral items corresponding to each Standard item for progress monitoring
- provide frequent narrow band assessments that can be validated with the Standard
- User can select items based on elevated Standard scale score for an individual student
    - — customized treatment goals
- Ability to make notes to describe specific behaviors, response to services, or to add other comments
- Teacher, Parent, Self and Clinician forms

Psychometric Properties

Large Normative Sample



Age x Gender Distribution: Normative Sample

Age Group	Teacher Ratings			Parent Rating			Self-Reports		
	Male (*N*)	Female (*N*)	Total (*N*)	Male (*N*)	Female (*N*)	Total (*N*)	Male (*N*)	Female (*N*)	Total (*N*)
5-6	100	100	200	100	100	200			
7-9	150	150	300	150	150	300			
10-11	100	100	200	100	100	200			
12-13	100	100	200	100	100	200	100	100	200
14-16	150	150	300	150	150	300	150	150	300
17-18	100	100	200	100	100	200	100	100	200
Total	**700**	**700**	**1400**	**700**	**700**	**1400**	**350**	**350**	**700**

Race/Ethnicity Distribution

- Highly comparable to the most recent U.S. Census

Form			Asian	African American	Hispanic	White	Other	Total
Teacher	**Total**	***N***	55	218	203	836	50	1361
		%	4.0	16.0	14.9	61.4	3.7	
	Census	**%**	3.8	15.7	15.1	61.9	3.5	
	Difference	**%**	**0.22**	**0.29**	**- 0.22**	**-0.47**	**0.18**	
Parent	**Total**	***N***	30	214	207	873	75	1400
		%	2.2	15.3	14.8	62.4	5.4	
	Census	**%**	3.8	15.7	15.1	61.9	3.5	
	Difference	**%**	**- 1.65**	**- 0.39**	**- 0.33**	**0.47**	**1.89**	
Self-Report	**Total**	***N***	28	110	107	433	25	703
		%	4.0	15.6	15.2	61.6	3.5	
	Census	**%**	3.8	15.7	15.1	61.9	3.5	
	Difference	**%**	**0.23**	**- 0.07**	**0.09**	**- 0.29**	**0.03**	

(Weighted *N*'s)

Geographic Region Distribution

- Highly comparable to the most recent U.S. Census

Form			Northeast	Midwest	South	West	Total
Teacher	**Total**	***N***	251	299	486	325	1361
		%	18.4	22.0	35.7	23.9	
	Census	**%**	18.1	21.9	36.7	23.3	
	Difference	**%**	**0.35**	**0.08**	**-1.03**	**0.61**	
Parent	**Total**	***N***	272	265	530	333	1400
		%	19.4	18.9	37.9	23.8	
	Census	**%**	18.1	21.9	36.7	23.3	
	Difference	**%**	**1.39**	**-2.97**	**1.13**	**0.47**	
Self-Report	**Total**	***N***	128	159	259	157	703
		%	18.3	22.6	36.8	22.4	
	Census	**%**	18.1	21.9	36.7	23.3	
	Difference	**%**	**0.21**	**0.70**	**0.03**	**-0.93**	

(Weighted *N*'s)

Parental Education Level

- Highly comparable to the most recent U.S. Census

Parent Education Level		High school or Lower	Apprenticeship/ 2-year College	University or higher	Total
Total	***N***	646	385	369	1400
	%	46.2	27.5	26.4	
Census	**%**	46.6	27.2	26.2	
Difference	**%**	**- 0.43**	**0.28**	**0.16**	

(Weighted *N*'s)

BIMAS Standard scale descriptor cut-offs

BIMAS Scales	***T*-score**	**Scale Descriptors**
Behavioral Concern Scales	*T* = 70+	**High Risk**
	T = 60-69	**Some Risk**
	T = 60 or less	**Low Risk**
Adaptive Scales	*T* = 40 or less	**Concern**
	T = 41-59	**Typical**
	T = 60+	**Strength**

Internal Consistency
Cronbach's Alpha

Form	Behavioral Concern Scales			Adaptive Scales	
	Conduct	Negative Affect	Cognitive/ Attention	Social	Academic Functioning
Teacher	.91	.85	.91	.85	.81
Parent	.87	.82	.90	.84	.77
Self-Report	.88	.85	.87	.83	.75

Test-Retest Reliability Coefficients

Form	Behavioral Concern Scales			Adaptive Scales	
	Conduct	Negative Affect	Cognitive/ Attention	Social	Academic Functioning
Teacher (*N* = 112)	.89	.85	.91	.91	.91
Parent (*N* = 83)	.79	.91	.84	.96	.80
Self-Report (*N* = 53)	.81	.87	.82	.90	.85

All *r*s significant, $p < .001$.; A 2-4 week interval (non-clinical sample; no intervention in between)

Across-Informant Correlations

- Correlation between parent & teacher ratings
- Correlation between self-report & parent/teacher
- Are the behaviors assessed by the BIMAS consistently detected by raters in different settings?
- (Diff informant: Diff observation context)

- Parent to Teacher *r*: range = .79 - .86
- Parent to Self *r*: range = .59 - .69
- Teacher to Self *r*: range = .54 - .59

BIMAS Validity

A multi-informant screening tool to identify emotional and behavior concerns.

How is the BIMAS for Screening Purposes?

- Parent
- Self
- Teacher

The BIMAS Clinical Samples

- Large clinical samples; different diagnostic groups

Clinical Diagnoses of the samples rated by teachers, parents and students themselves.							
Clinical Group	**Teacher**		**Parent**		**Self-Report**		**Total**
	N	**%**	***N***	**%**	***N***	**%**	***N***
DBD	123	22.9	70	15.0	65	18.6	258
ADHD	109	20.3	117	25.1	89	25.4	315
Anxiety	55	10.2	67	14.3	56	16.0	178
Depression	60	11.2	73	15.6	62	17.7	195
PDD	95	17.7	86	18.4	65	18.6	246
LD	45	8.4	--	--	--	--	45
DD	30	5.6	--	--	--	--	30
Other	21	3.9	54	11.6	13	3.7	88
Total	538	100.0	467	100.0	350	100.0	1355

BIMAS–Parent scores can differentiate between Clinical vs. Non-Clinical

BIMAS-P Standard Scales	Clinical Sample			Cohen's *d*
	N	*M*	*SD*	
Conduct	467	60.3	10.5	1.0
Negative Affect	467	61.5	10.3	1.1
Cognitive/Attention	467	60.7	9.9	1.1
Social	467	38.4	9.9	−1.2
Academic Functioning	467	40.4	7.9	−1.0

Note. Clinical *M*s (*SD*s) 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
(All satisfactory)

Classification Accuracy Statistic	Full Range of Scores	Cut-Scores
Overall Correct Classification	**78.3%**	**78.6%**
Sensitivity	**80.1%**	**73.4%**
Specificity	**77.7%**	**80.3%**
Positive Predictive Power	**54.6%**	**55.4%**
Negative Predictive Power	**92.1%**	**90.1%**

BIMAS–Self Report scores can differentiate between Clinical vs. Non-Clinical

BIMAS-P Standard Scales	Clinical Sample			Cohen's *d*
	N	*M*	*SD*	
Conduct	350	57.3	9.7	0.7
Negative Affect	350	59.2	9.7	0.9
Cognitive/Attention	350	57.3	8.2	0.8
Social	350	41.4	9.7	−0.9
Academic Functioning	350	42.3	8.3	−0.8

Note. Clinical *M*s (*SD*s) 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(All satisfactory)

Classification Accuracy Statistic	Full Range of Scores	Cut-Scores
Overall Correct Classification	71.5%	71.8%
Sensitivity	76.3%	67.1%
Specificity	69.1%	74.1%
Positive Predictive Power	55.3%	56.5%
Negative Predictive Power	85.3%	81.9%

BIMAS–T scores can differentiate between Clinical vs. Non-Clinical

BIMAS-T Standard Scales	Clinical Sample			Cohen's *d*
	N	*M*	*SD*	
Conduct	516	63.5	10.9	1.3
Negative Affect	537	66.4	10.4	1.6
Cognitive/Attention	538	66.6	9.8	1.7
Social	538	35.6	10.3	−1.4
Academic Functioning	538	40.2	9.8	−1.0

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| = small effect, |0.5| = medium effect, and |0.8| = large effect.

Classification Accuracy of BIMAS–Teacher Scales (All satisfactory)

Classification Accuracy Statistic	Full Range of Scores	Cut-Scores
Overall Correct Classification	85.2%	82.5%
Sensitivity	83.5%	80.1%
Specificity	85.8%	83.4%
Positive Predictive Power	68.4%	64.9%
Negative Predictive Power	93.4%	91.6%

Psychometric Properties

- Large normative sample closely matching U.S. Census
- Reliability (internal consistency, test-retest reliability & interrater reliability)
- Validity - content based on IISRs & scale developed based on EFA & CFA

  - converged with another behavioral assessment (Conners CBRS)

  - showed good ability to screen

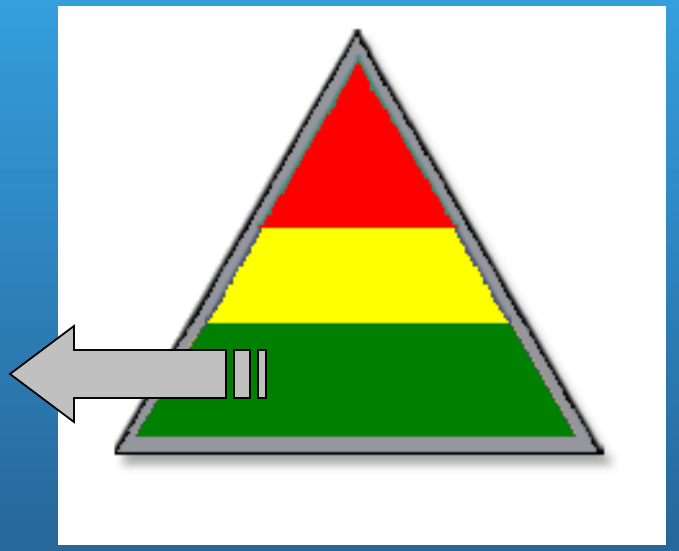
  - showed good ability to detect change post intervention

Sreening and Progress Monitoring within the RTI Framework

How to use BIMAS within the RtI Framework

Tier 1

- Universal Level
- Students without serious problem behaviors (80-90%)
- Use BIMAS to Screen



PBIS Screening: Lanigan School

- Elementary school

  approximately 400 students

- Grades Pre-K to 6

ODRs- office discipline referrals

Most commonly used data

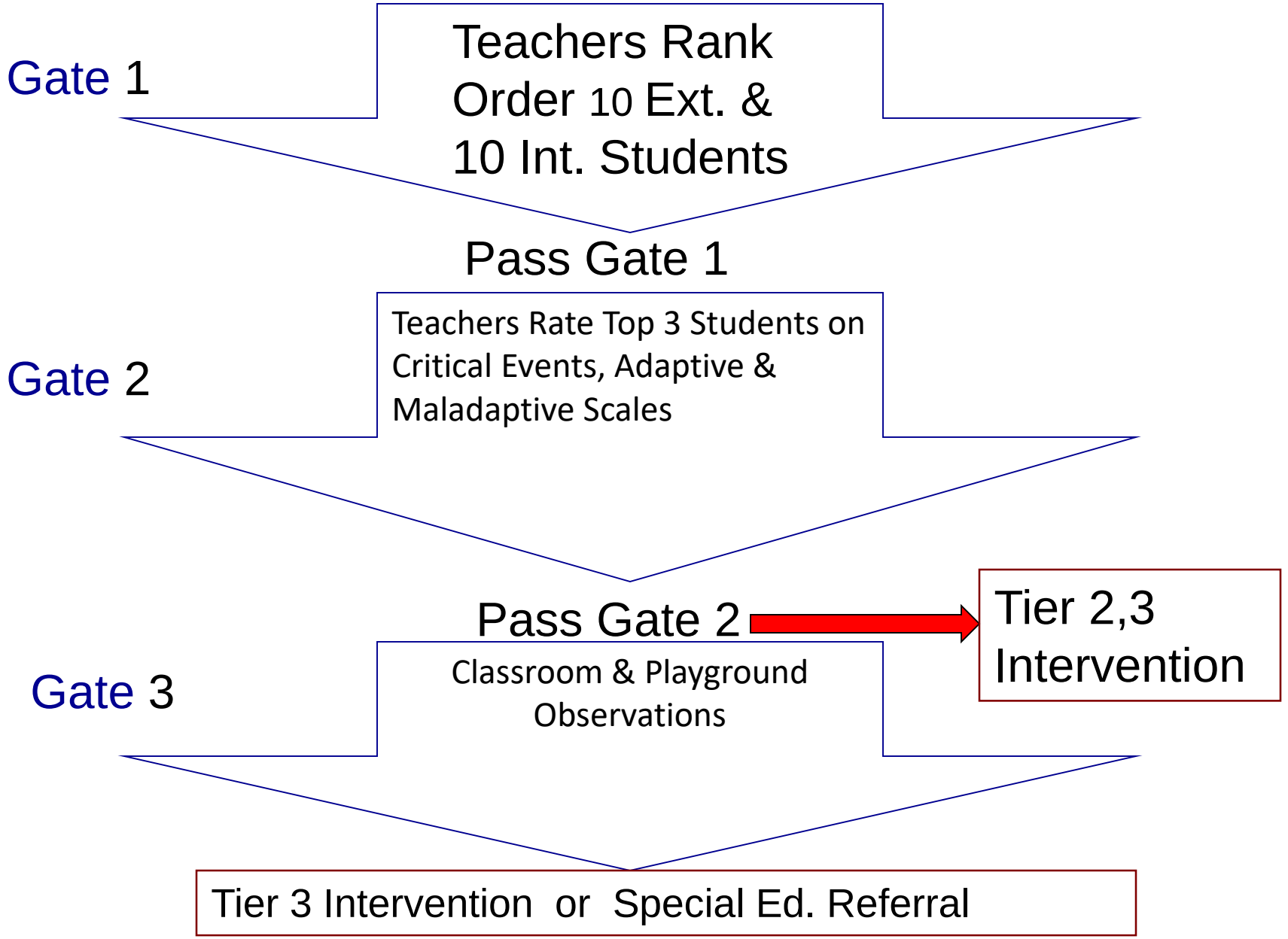
- Pros-
  - Easy to collect
  - Of interest to schools
  - Helps to identify areas, times, places and students in need of improvement
- Cons-
  - Lack of validity and reliability for screening and PM
  - Under-identify non-externalizing students

The Systematic Screening for Behavior Disorders (SSBD) (Walker and Severson, 1992)

Developed as a school-wide (Universal) screening tool for children in grades 1-6

- Provides systematic screening of ALL students in grades 1-6 based on teacher nomination from class lists
- Screens for externalizing (e.g. “acting out”) AND internalizing (e.g. introverted) behaviors

Multiple Gating Procedure (Severson et al. 2007)



SSBD- Referred to as the gold standard of screening in the schools

- Pros-

  SSBD does have demonstrated validity (and to a lesser extent reliability) especially for externalizing behaviors

  Better sensitivity than ODRs for proactively identifying externalizing students

  Feasible for teacher and schools to use- though playground observations are not likely typical

- Cons-

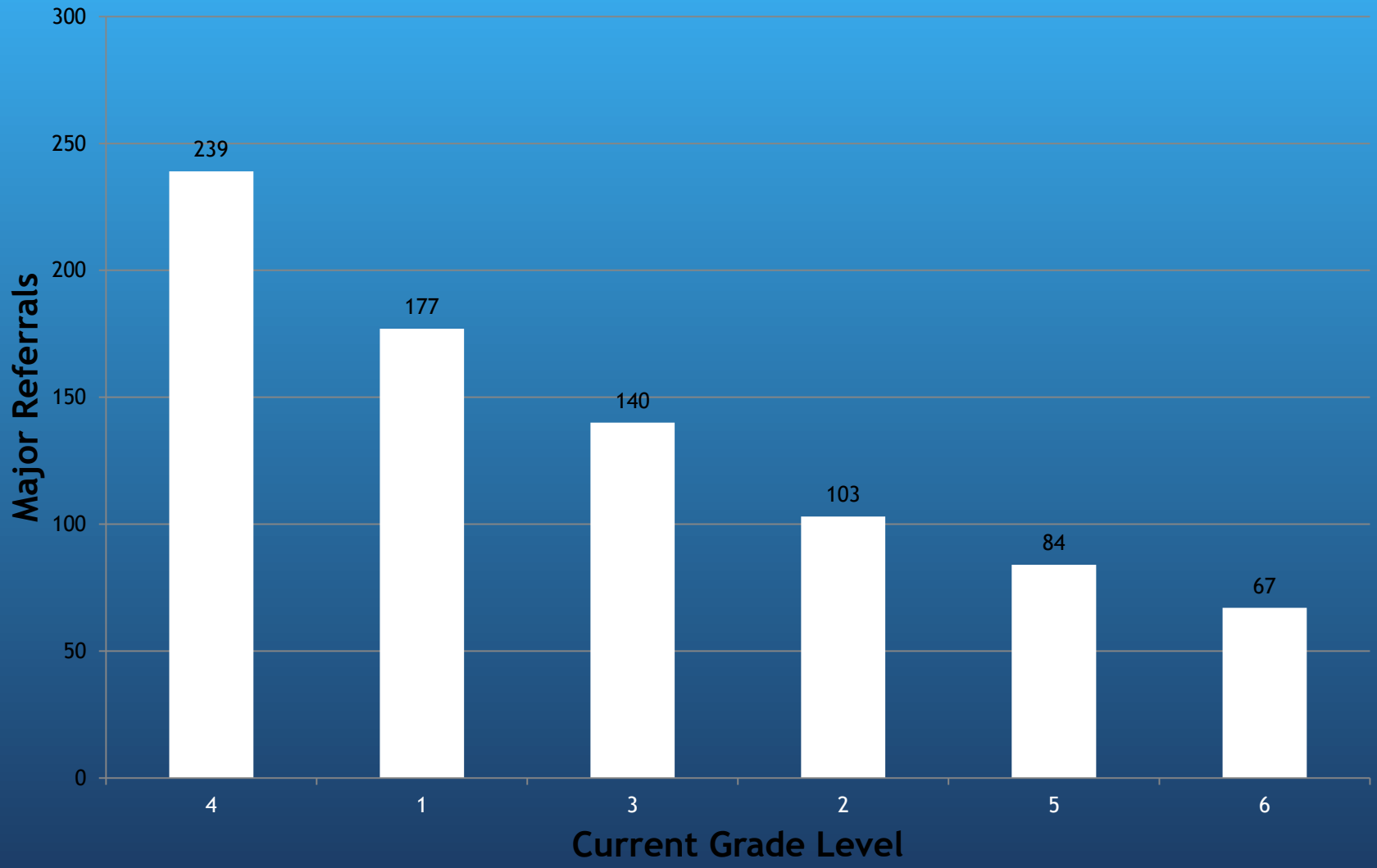
  Forced nomination of 3 students per category per class (maybe too many/few)

  Observations are time consuming

  Better sensitivity for externalizing than internalizing

  Limited usefulness for progress monitoring and program evaluation

ODRs 2011-2012. Data used to target 4th Grade



SSBD/ ODR Information 2012-2013

Externalizing

SSBD Concern Level	2012-2013 Major Referrals
1	21
1	6
1	19
2	4
2	5
2	6
3	0
3	7
3	23
4	0
4	8
4	0
5	0
5	6
6	0

Internalizing

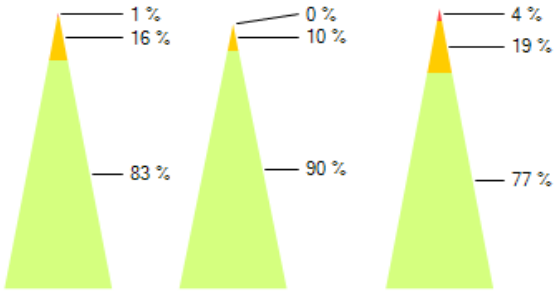
SSBD Concern Level	2012-2013 Major Referrals
1	0
1	2
1	0
2	0
2	0
3	0
4	0
5	0

4th Grade Screening Results - BIMAS

Total For Grade 4
70 Students

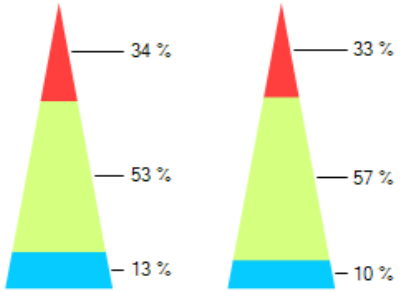
Behavioral Concern Scales

Percent of Students



Adaptive Scales

Percent of Students



Levels Of Risk	Conduct	Negative Affect	Cognitive/ Attention	Levels Of Functioning	Social	Academic Functioning
High Risk	1 (1 %)	0 (0 %)	3 (4 %)	Concern	24 (34 %)	23 (33 %)
Some Risk	11 (16 %)	7 (10 %)	13 (19 %)	Typical	37 (53 %)	40 (57 %)
Low Risk	58 (83 %)	63 (90 %)	54 (77 %)	Strength	9 (13 %)	7 (10 %)
Total	70 (100%)	70 (100%)	70 (100%)	**Total**	70 (100%)	70 (100%)

Note: Total percentage may not always add up to 100% due to rounding.

SSBD Screening Externalizing Behaviors

		BIMAS				
		Externalizing	Not identified			
SSBD	Externalizing	10	5	15	Sensitivity	0.83
	Not identified	2	11	13	Specificity	0.69
		12	16	28	Efficiency	0.75

SSBD Screening Internalizing Behaviors

		BIMAS				
		Internalizing	Not identified			
SSBD	Internalizing	2	6	8	Sensitivity	0.40
	Not identified	3	17	20	Specificity	0.74
		5	23	28	Efficiency	0.68

ODRs Screening Externalizing Behaviors

		BIMAS				
		Externalizing	Not identified			
2012-2013 ODR	identified	9	2	11	Sensitivity	0.75
	Not identified	3	14	17	Specificity	0.88
		12	16	28	Efficiency	0.82

ODRs Screening Internalizing Behaviors

		BIMAS				
		Internalizing	Not identified			
2012-2013 ODR	Identified	0	11	11	Sensitivity	0.00
	Not identified	5	12	17	Specificity	0.52
		5	23	28	Efficiency	0.43

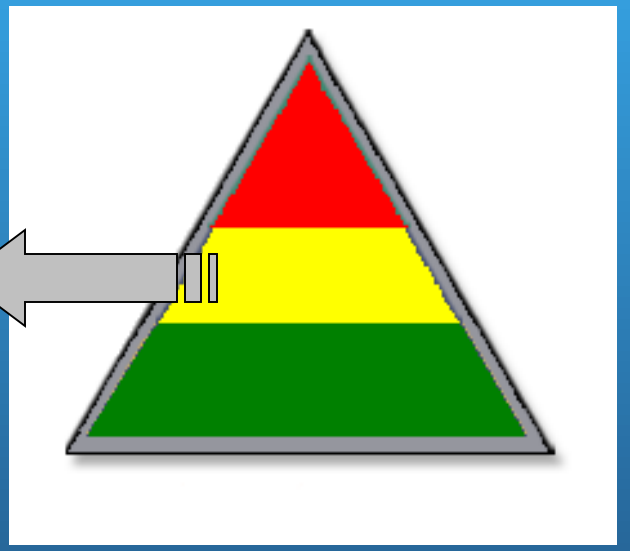
Implications

- SSBD & ODRs demonstrate moderate to strong classification rates for externalizing behaviors
- SSBD & ODRs demonstrate low classification rates for internalizing behaviors
- Neither approach is ideal for progress monitoring after screening

How to use BIMAS within the RtI Framework

Tier 2

- Targeted Level
- Students at risk for problem behaviors (5-15%)
- Use BIMAS to monitor & assess response to intervention/treatment



Integrated RTI Approach: Tier 2

- 8 year old student receiving academic (reading), behavioral, and psychosocial interventions in school

- Measures:
  - Direct observations
  - CBM
  - BIMAS- teacher and parent

Direct Observation: Donovan

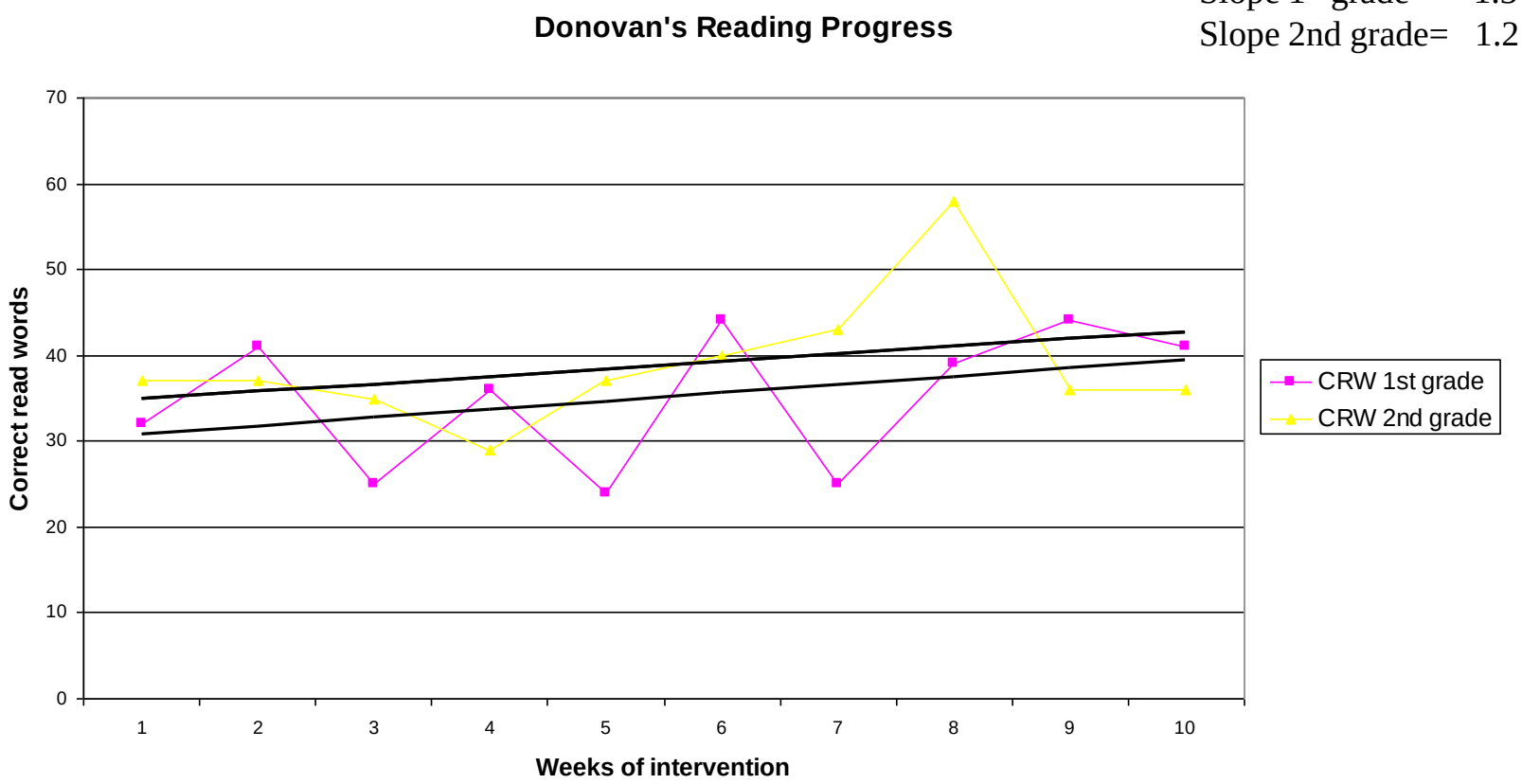
Percent of Time On Task in Class

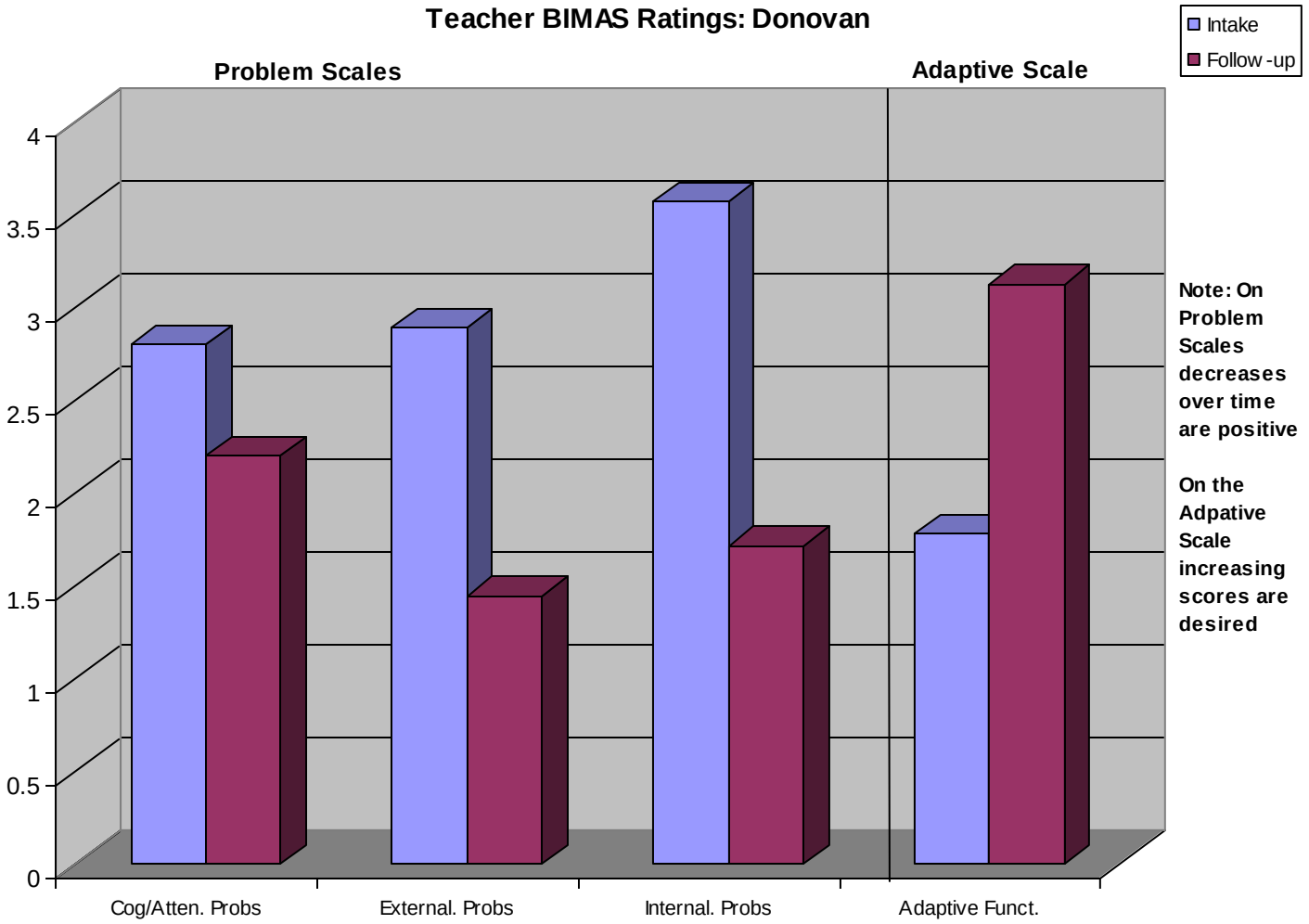
Student	Pre Intervention	Post Intervention	Change
Donovan	69.99	94.45	24.46

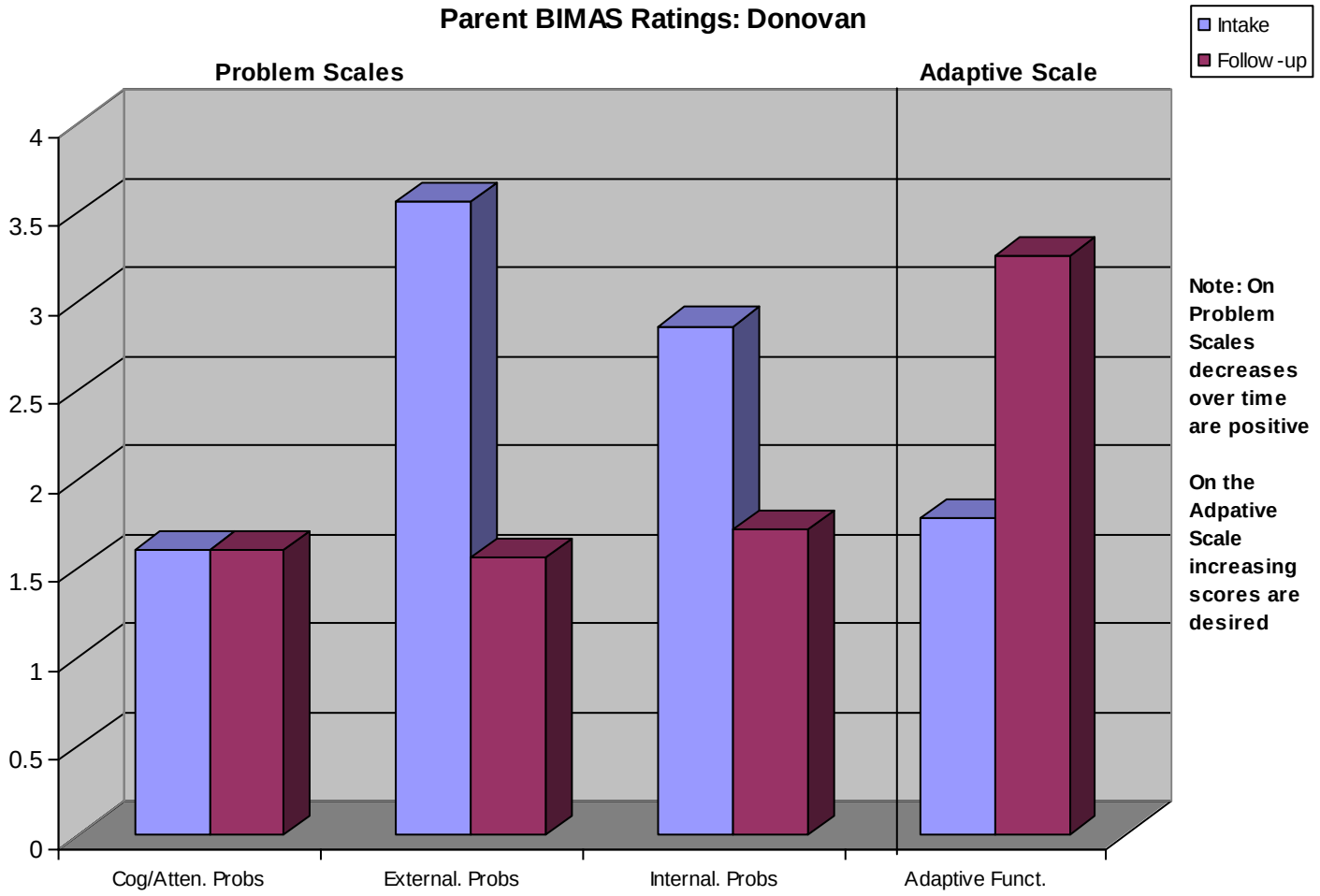
Numbers presented represent the median of three 10 minute observations

Donovan's Progress in Reading

Slope 1st grade= 1.3
Slope 2nd grade= 1.2





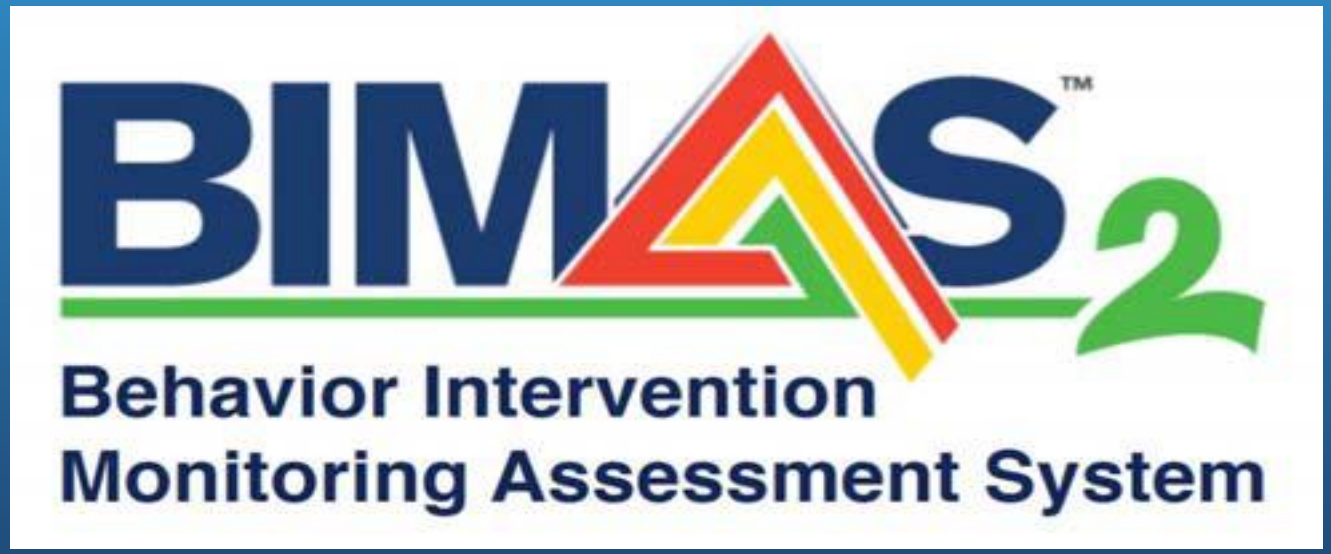


Note: On Problem Scales decreases over time are positive

On the Adpative Scale increasing scores are desired

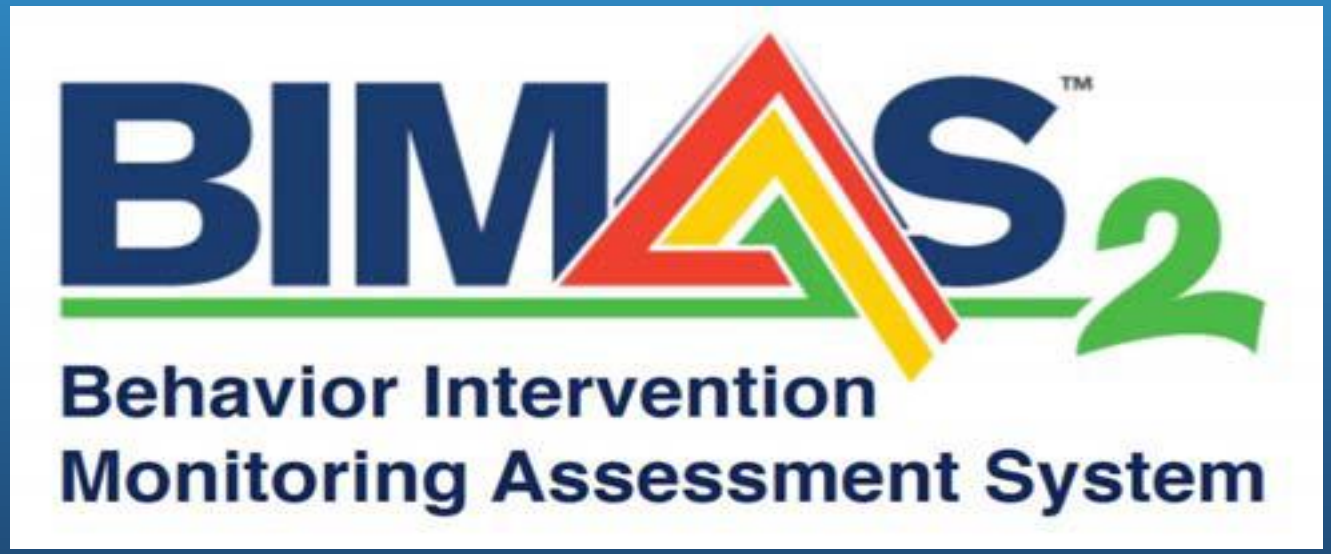
The BIMAS-2

- https://app.edumetrisis.com/login



The BIMAS-2

- Teacher's Experience
- https://www.youtube.com/watch?v=6piA2E2f15o



COMPREHENSIVE BEAVIORAL HEALTH MODEL (CBHM):

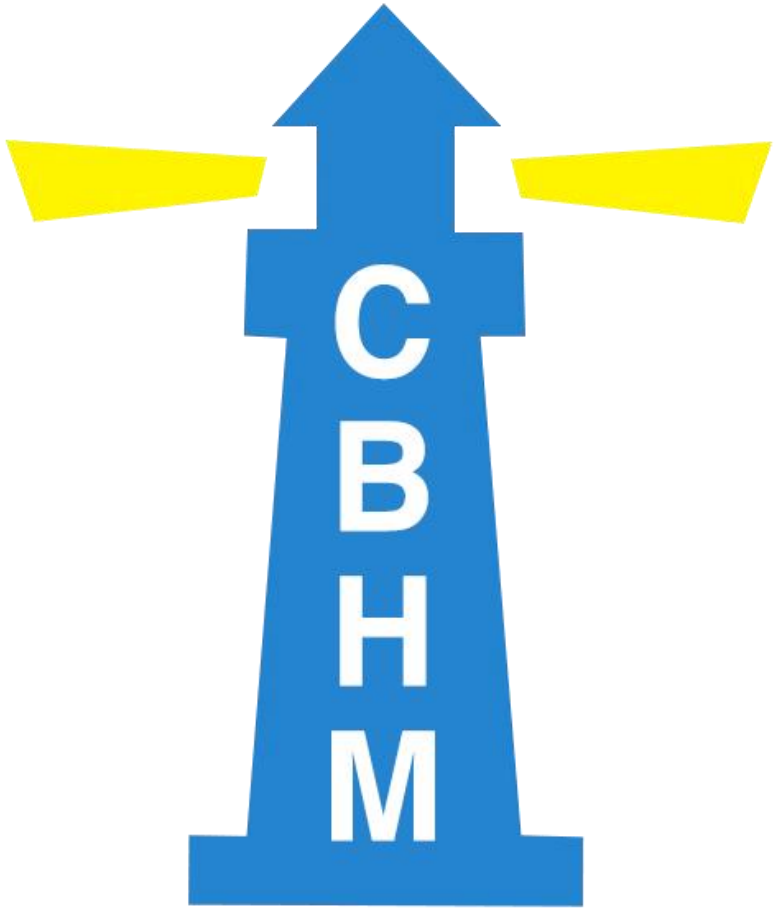
IMPACT ON STUDENT OUTCOMES OVER TIME

Jill Snyder, Ph.D., NCSP

CBHM Data & Research Coordinator

COMPREHENSIVE BEHAVIORAL HEALTH MODEL (CBHM)

A PARTNERSHIP…



BOSTON Public Schools
Focus on Children

Boston Children's Hospital
Until every child is well℠

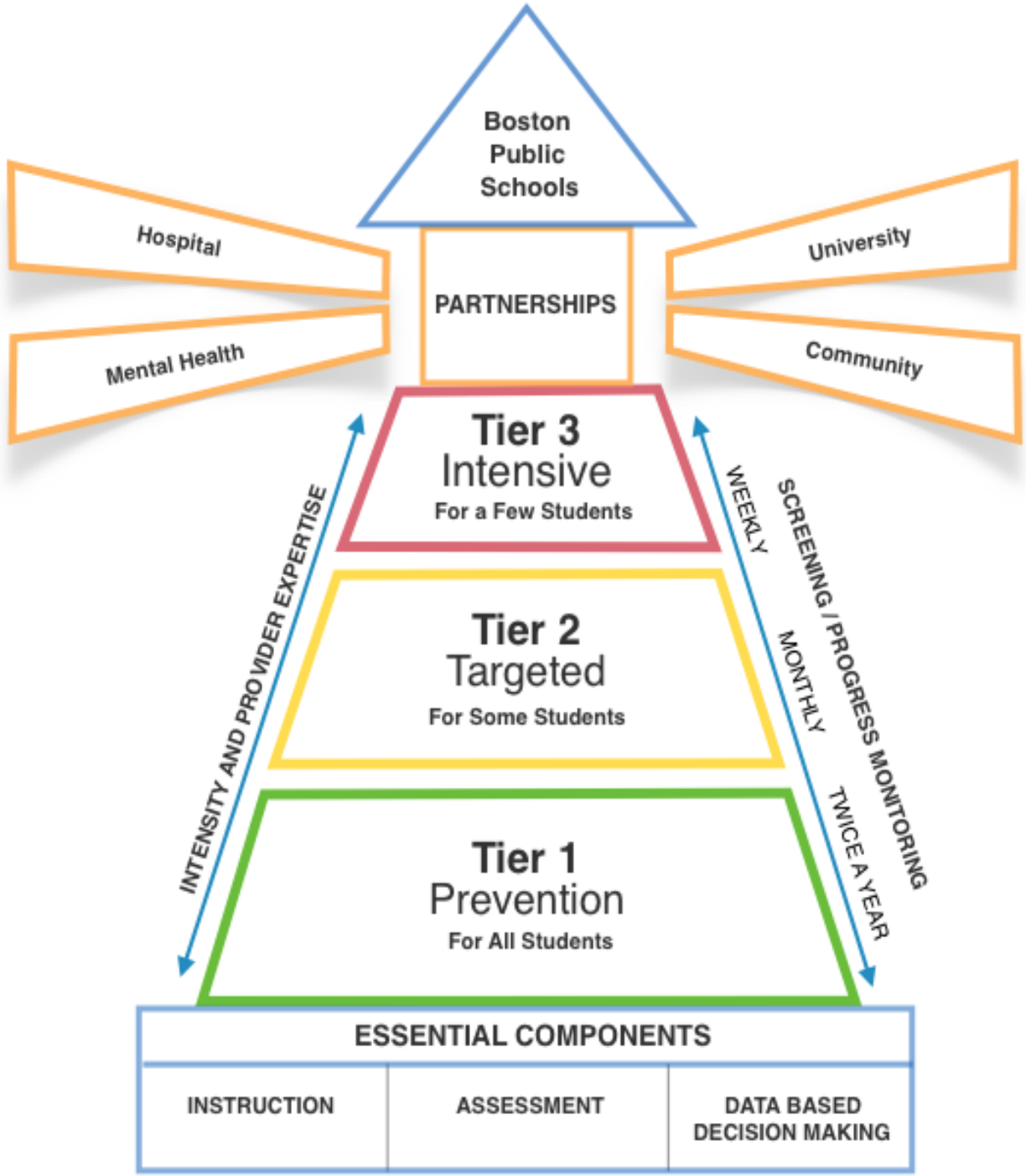
UMASS BOSTON

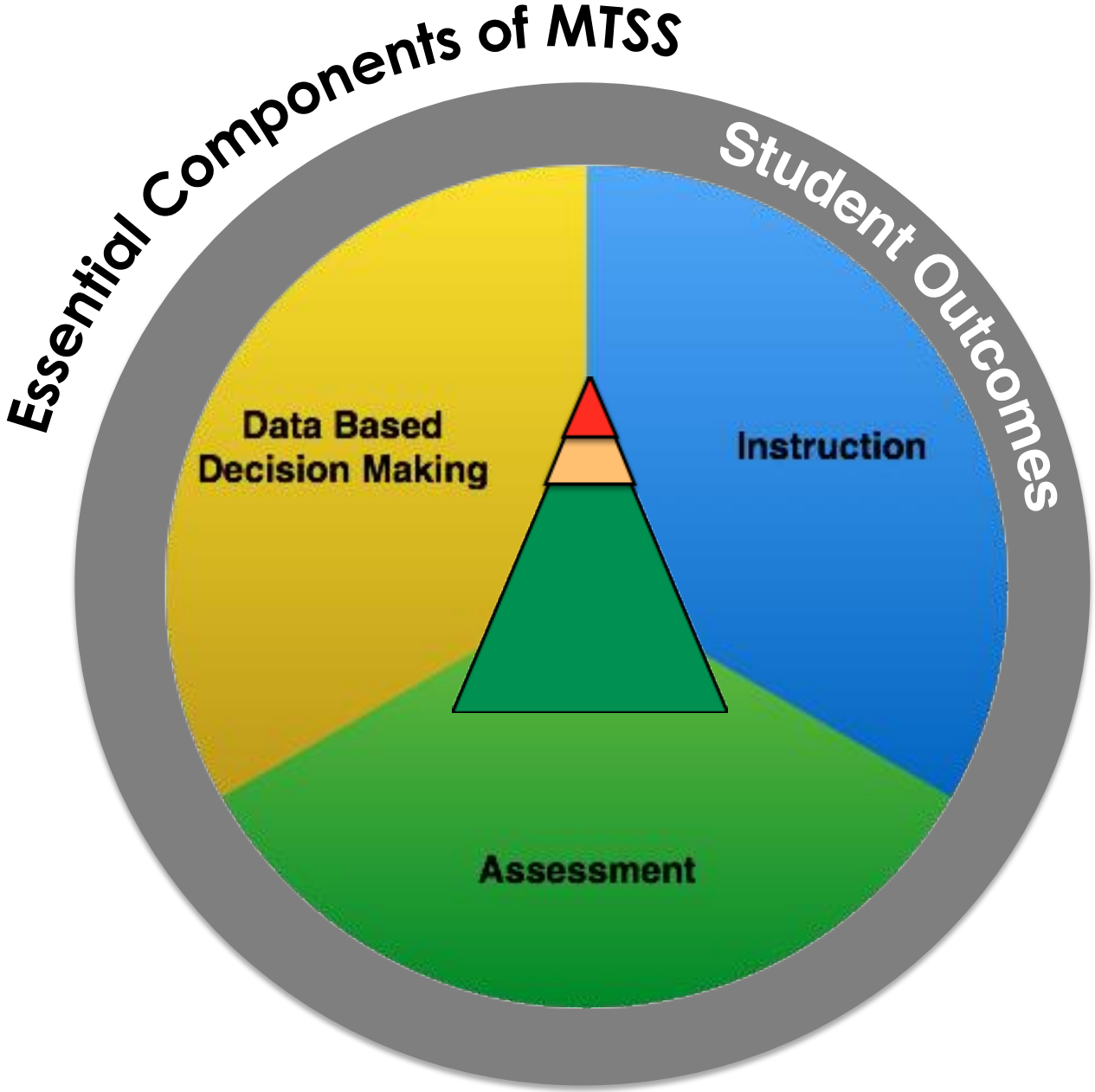
Mission:
Ensuring that all students have a safe and supportive school where they can be successful

EVERY STUDENT DESERVES A SAFE AND SUPPORTIVE SCHOOL

- Preventative model to build capacity within BPS to meet the **behavioral health & social emotional** needs of all students.
- Builds capacity within BPS schools to **provide instruction and intervention supports** along a continuum of student need (e.g. universal, targeted, intensive).
- Incorporates use of a **universal screener** to identify students at risk for social, emotional and/or behavioral health concerns early, and monitor student progress throughout intervention services.
- Implementation began in 10 schools during the 2012-13 School Year.
- Currently being implemented in **60 BPS schools**, serving **over 24,000** students.

www.cbhmboston.com


Boston Public Schools
Hospital
Mental Health
PARTNERSHIPS
University
Community
Tier 3
Intensive
For a Few Students
Tier 2
Targeted
For Some Students
Tier 1
Prevention
For All Students
INTENSITY AND PROVIDER EXPERTISE
WEEKLY
MONTHLY
TWICE A YEAR
SCREENING / PROGRESS MONITORING
ESSENTIAL COMPONENTS
INSTRUCTION
ASSESSMENT
DATA BASED DECISION MAKING

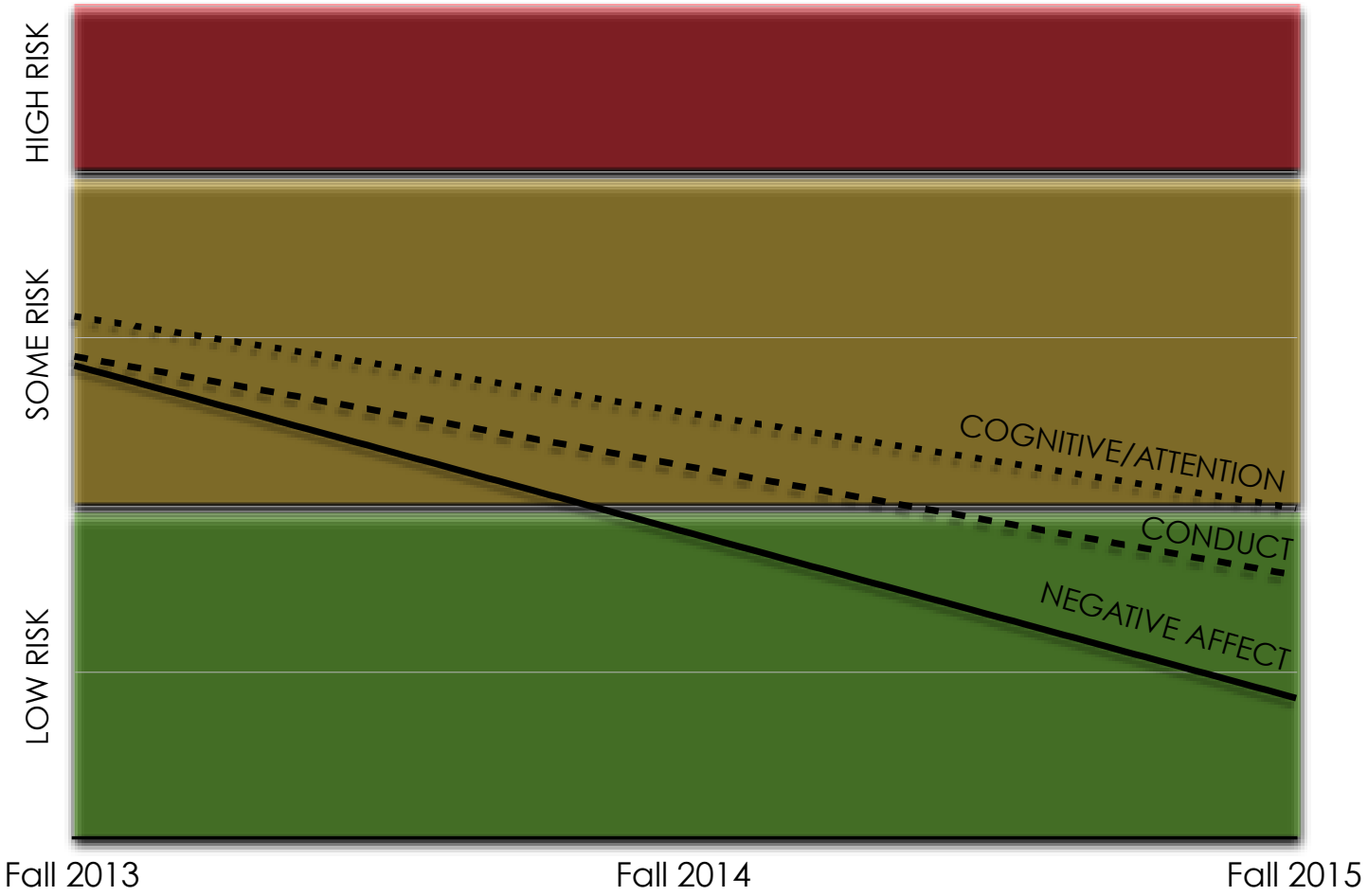

Essential Components of MTSS
Student Outcomes
Data Based
Decision Making
Instruction
Assessment

CBHM AT TIER 1

	WHAT	WHY	HOW
INSTRUCTION	• School Wide Positive Behavioral Interventions and Supports (SWPBIS)	Students need to know behavioral expectations throughout the school building in order to be successful in the school environment	Organize the school environment to prevent problem behaviors and reinforce positive behaviors
	• Social Emotional Learning (SEL) Curricula	Students need social and emotional skills to successfully navigate interactions with peers and adults	Instruction in fundamental social skills, such as empathy, relationship building, and conflict management
ASSESSMENT	• Universal Screening	Schools need universal data from all students to understand the strengths of instructional programming, as well as areas of need.	Collect objective information that can be used to guide instruction at multiple levels (e.g. school, grade, class, and individual student)
DATA BASED DECISION MAKING	• Problem Solving Teams & Data Based Decision Making	School teams need to understand how to use universal assessment data to make systemic decisions about instruction	School teams are effectively organized to promote efficient data-based decision making.

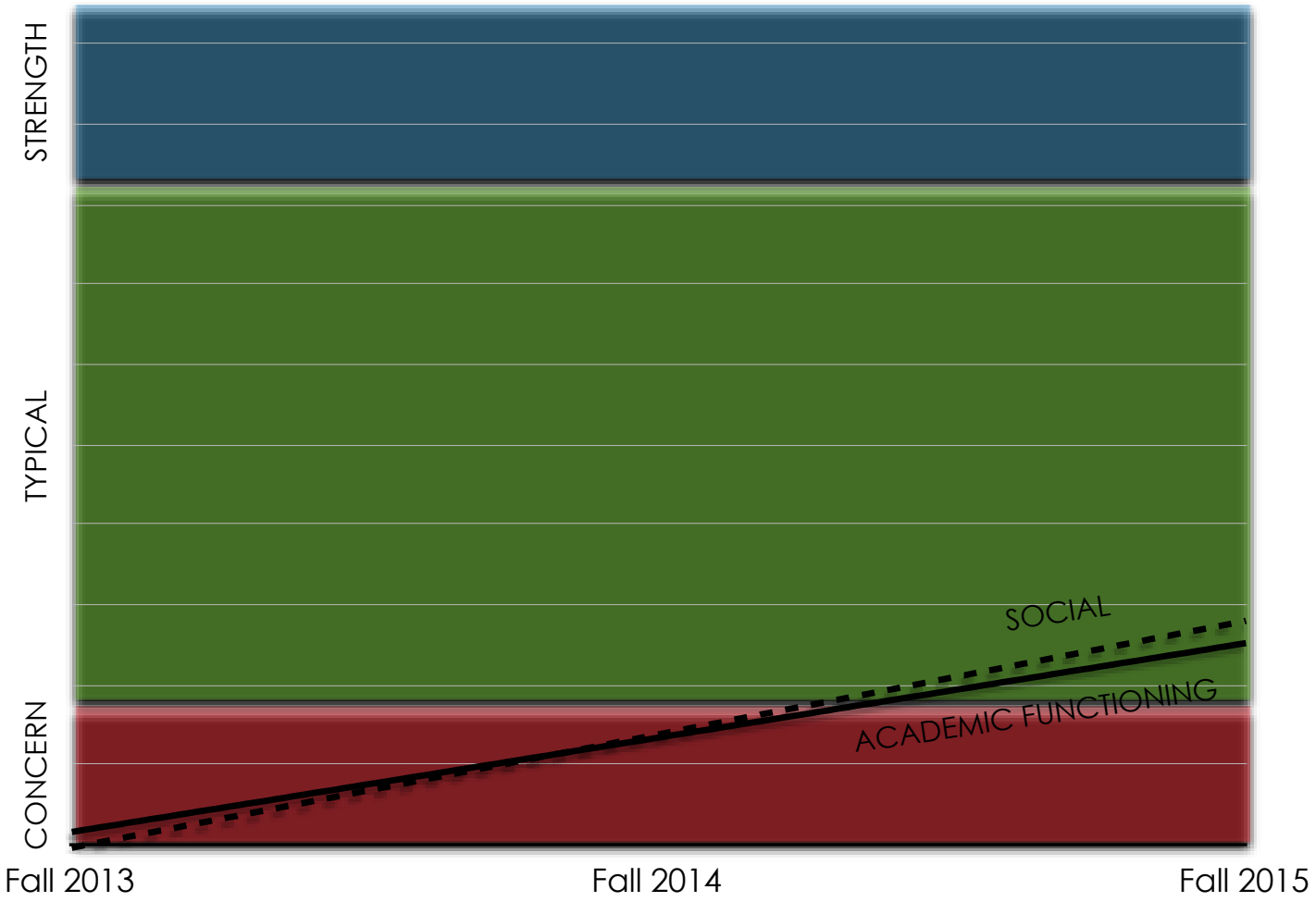
BIMAS Outcomes Over Time

Trend in Student Outcomes Among Students
At-Risk For Behavioral Concerns at Onset (Fall 2013)

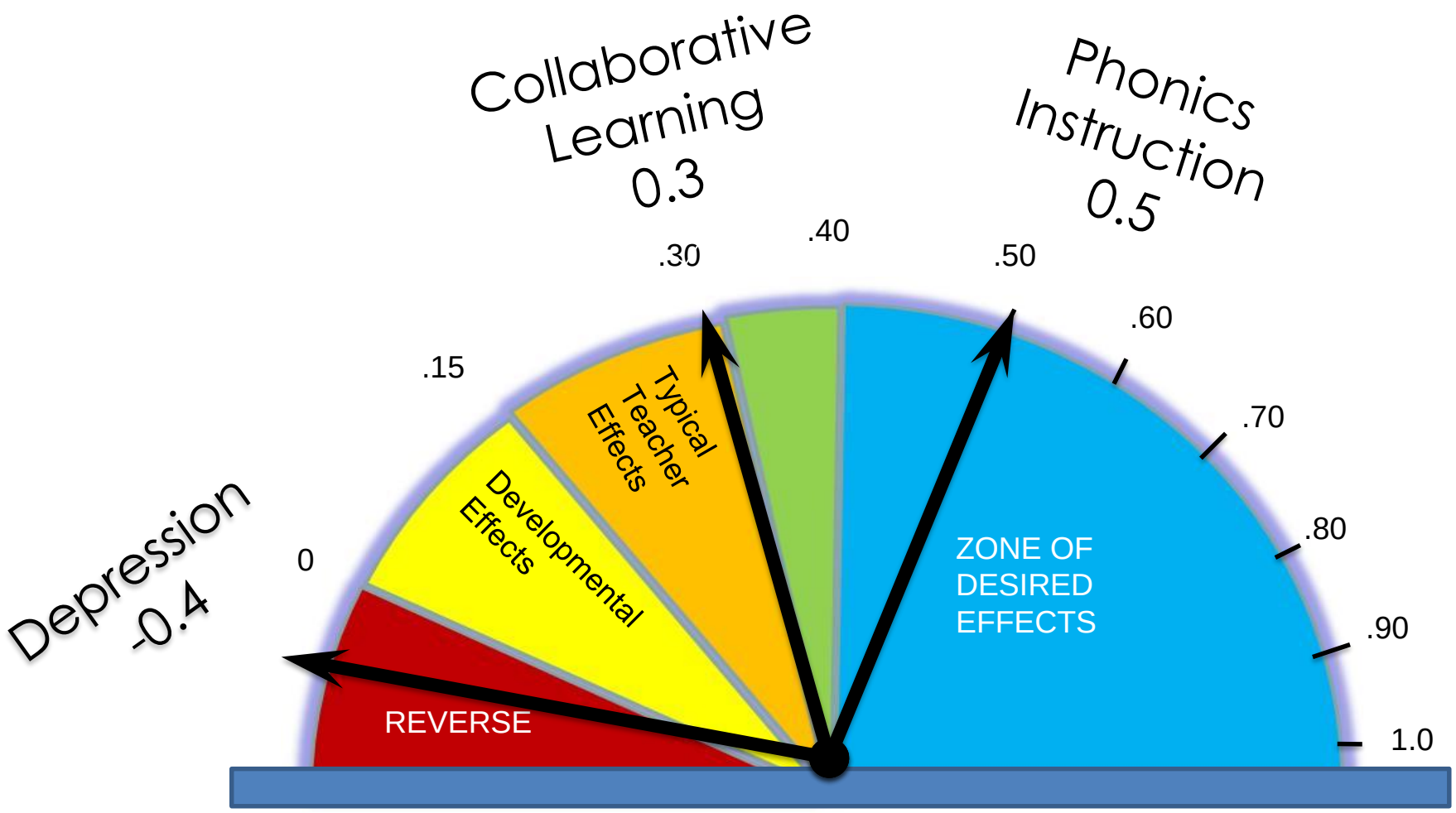


BIMAS Outcomes Over Time

Trend in Student Outcomes Among Students
At-Risk For Adaptive Concerns at Onset (Fall 2013)

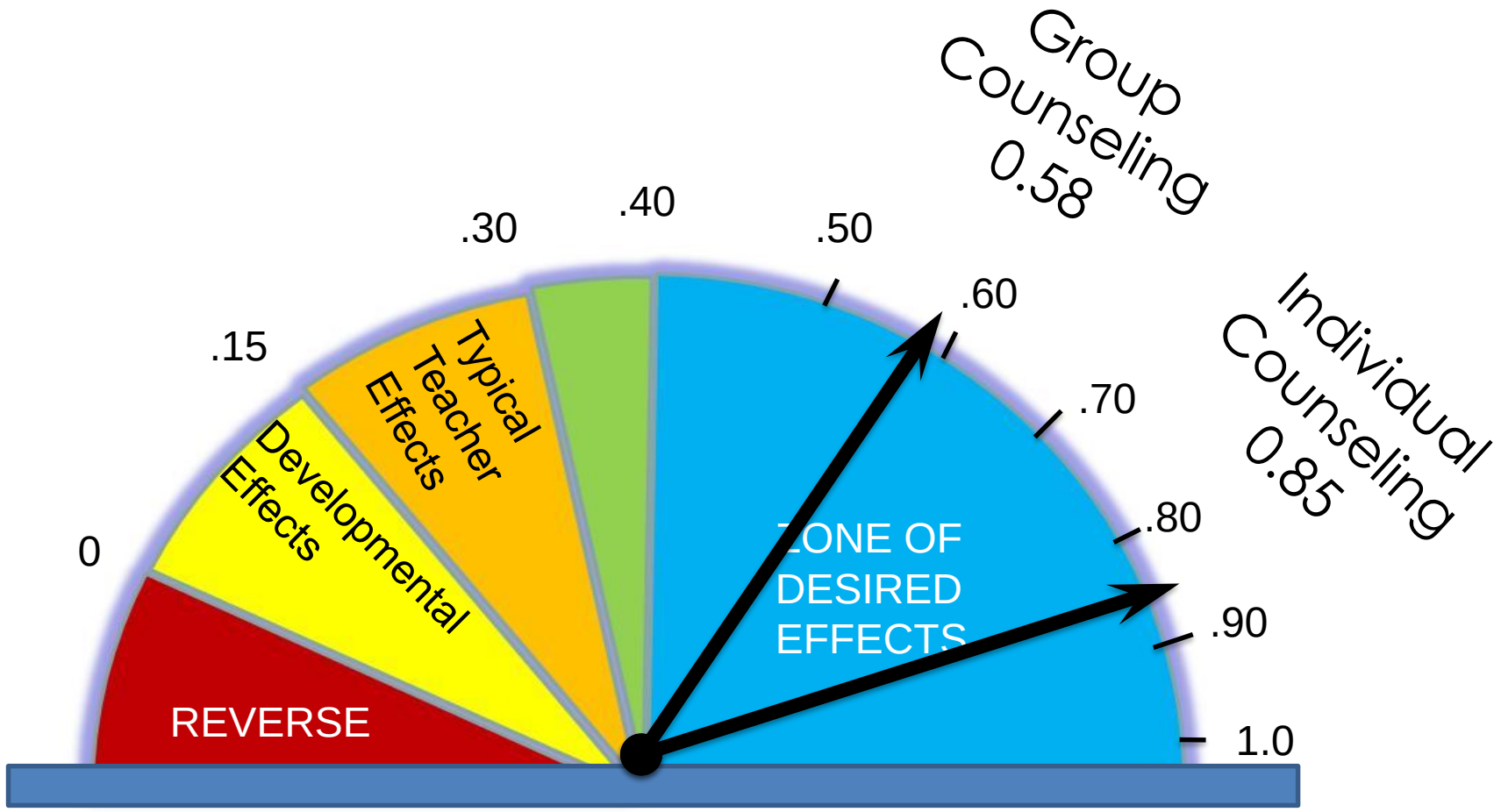


EFFECT SIZES: Impact on Learning



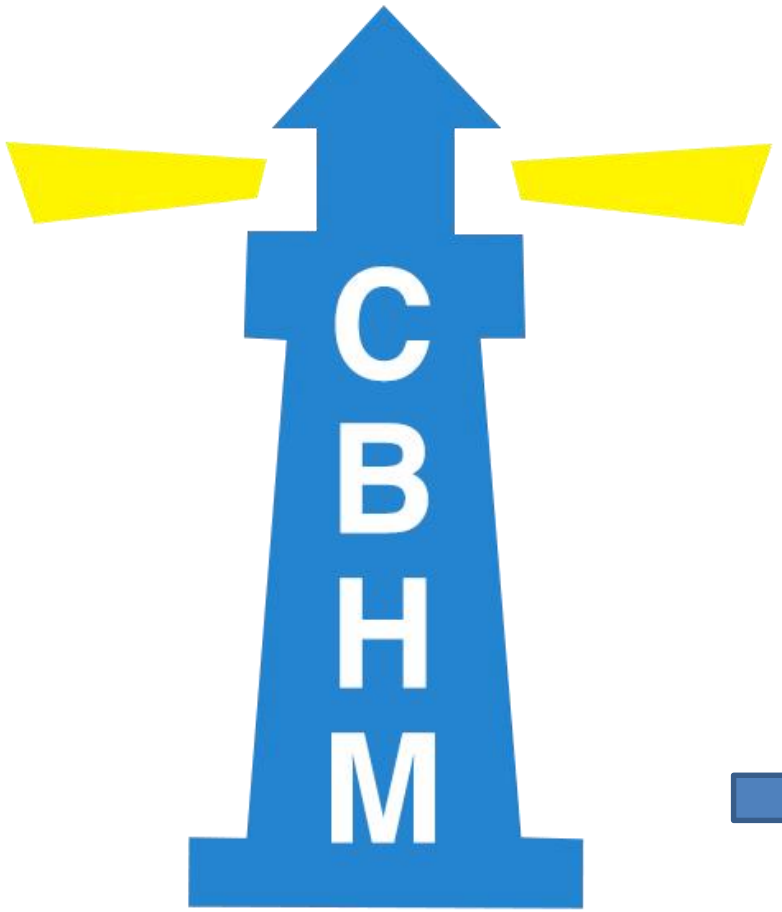
John Hattie, Visible Learning
http://visible-learning.org/

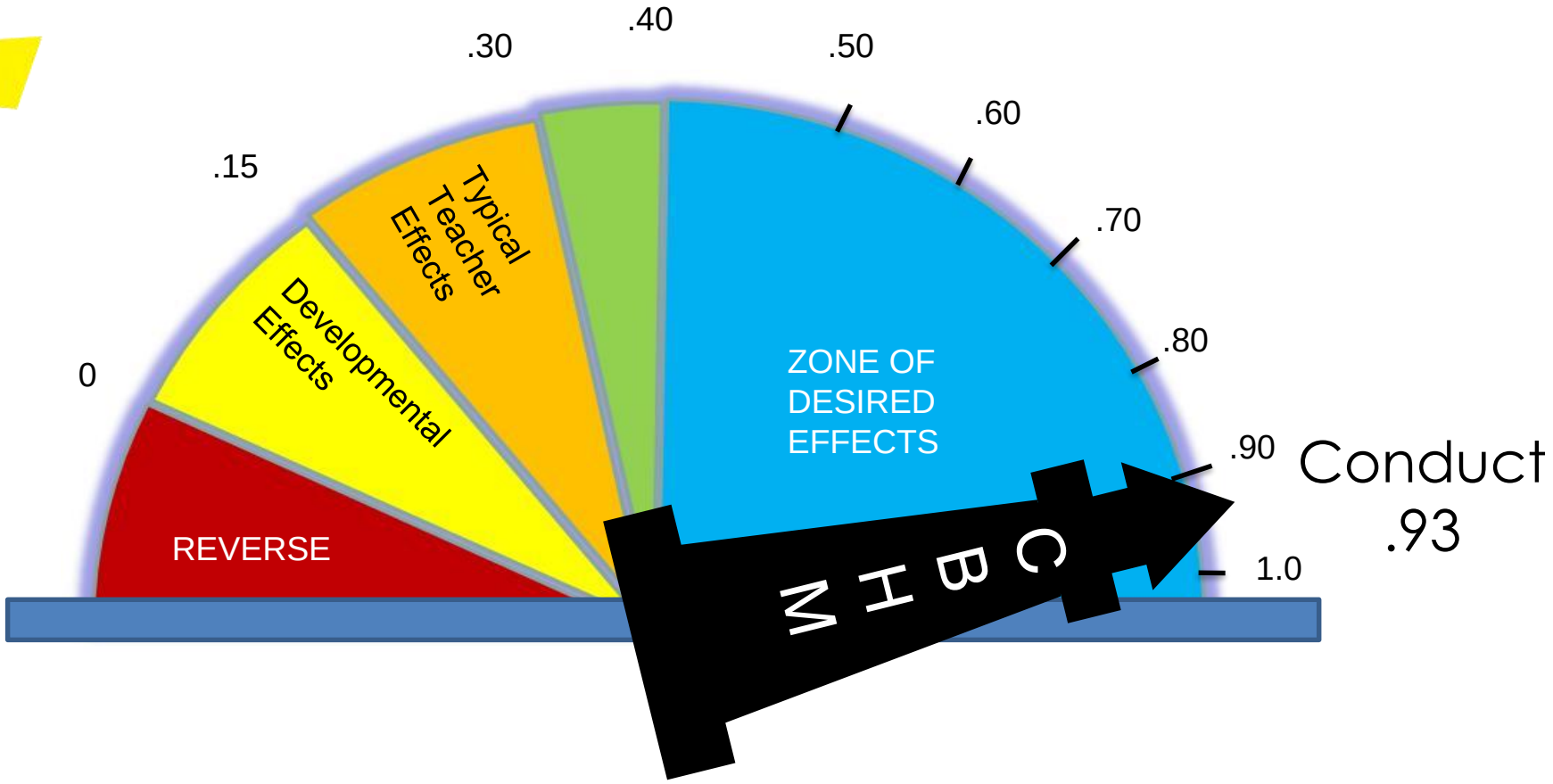
EFFECT SIZES: Impact on Learning



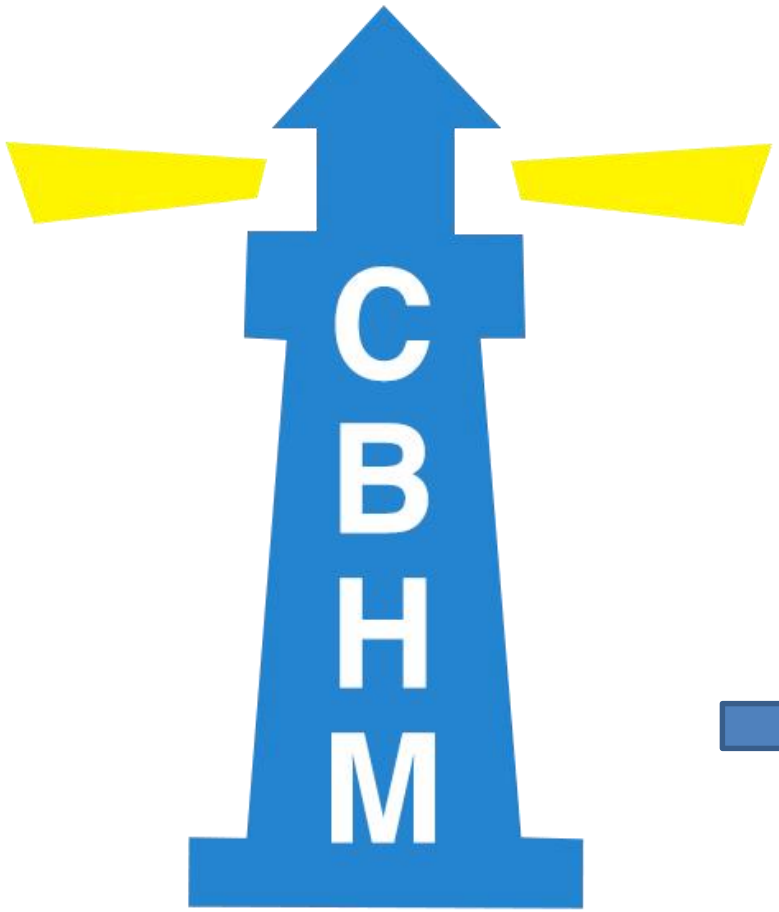
Reynolds, Wilson, & Hooper (2012)

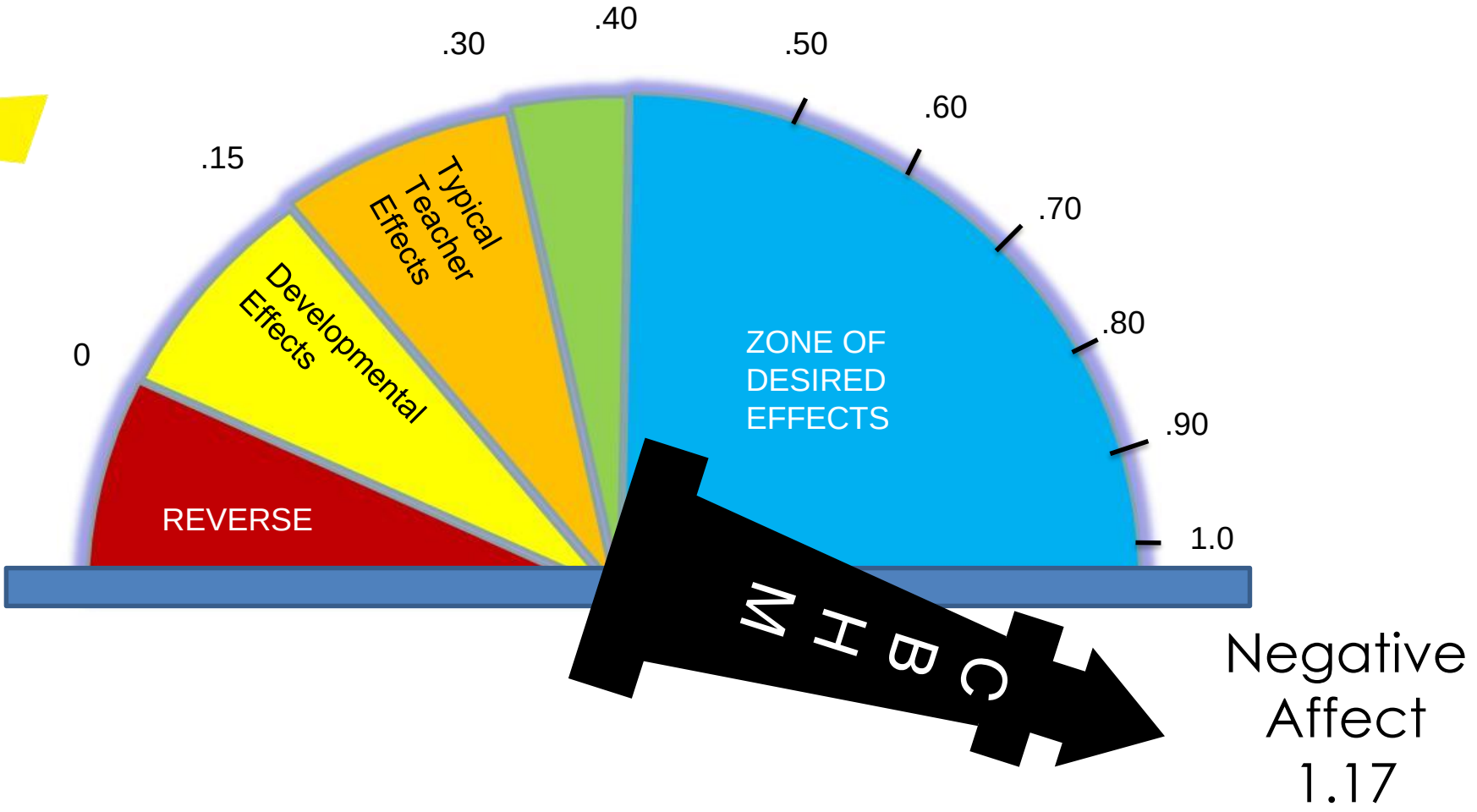
EFFECT SIZES: Impact on Learning



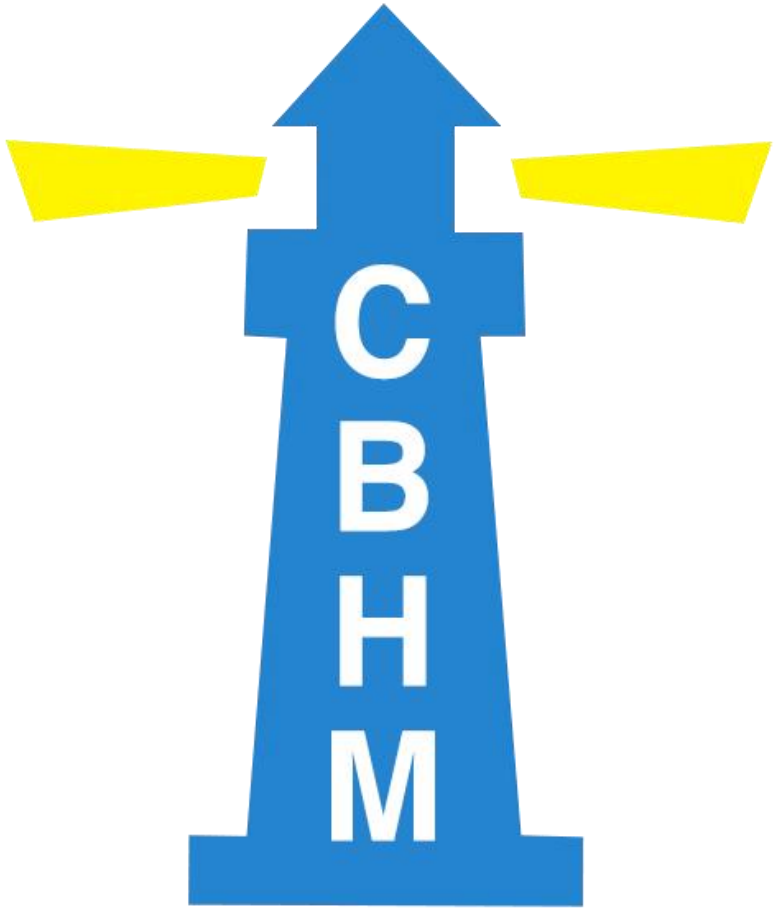


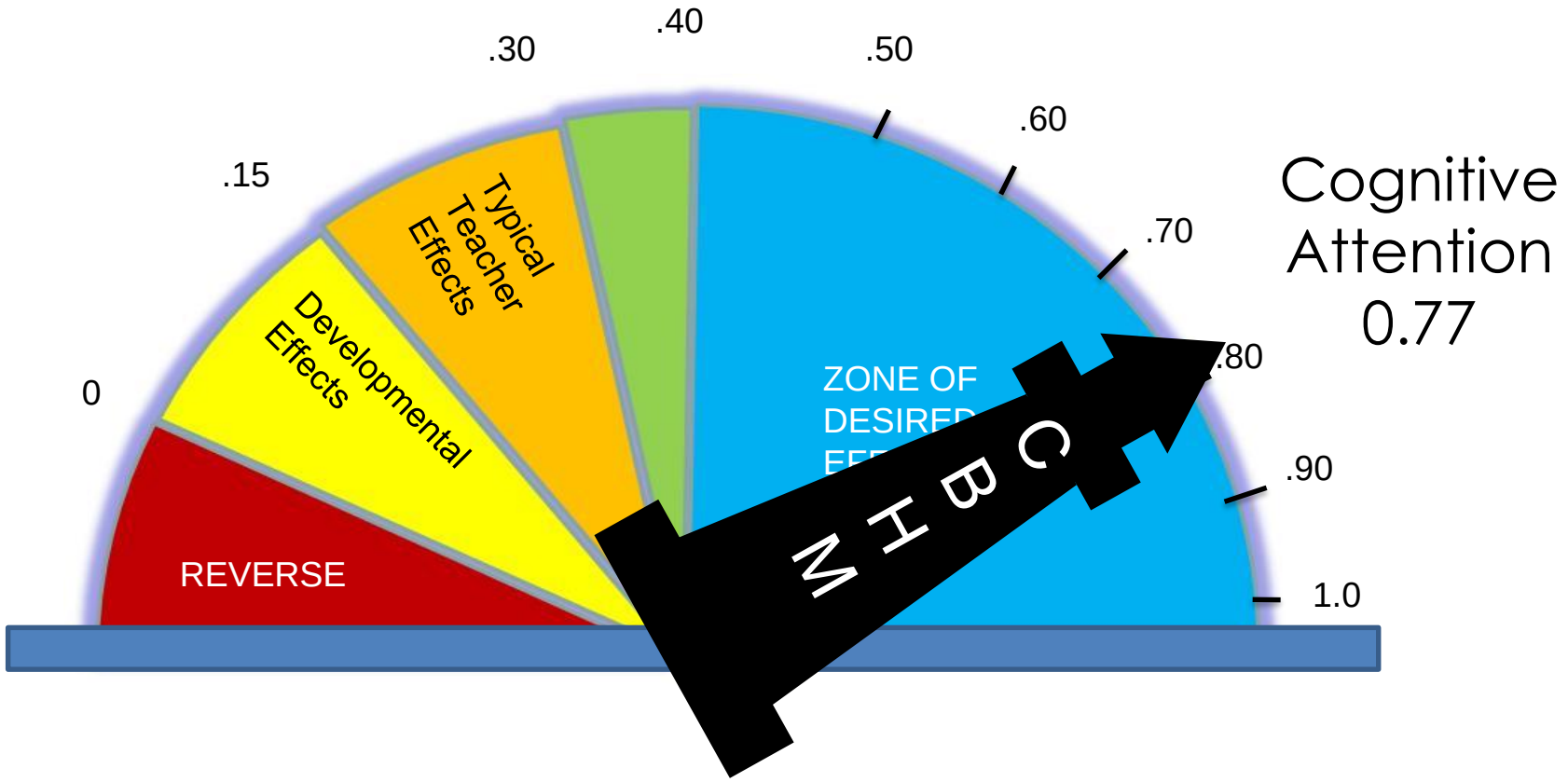
EFFECT SIZES: Impact on Learning



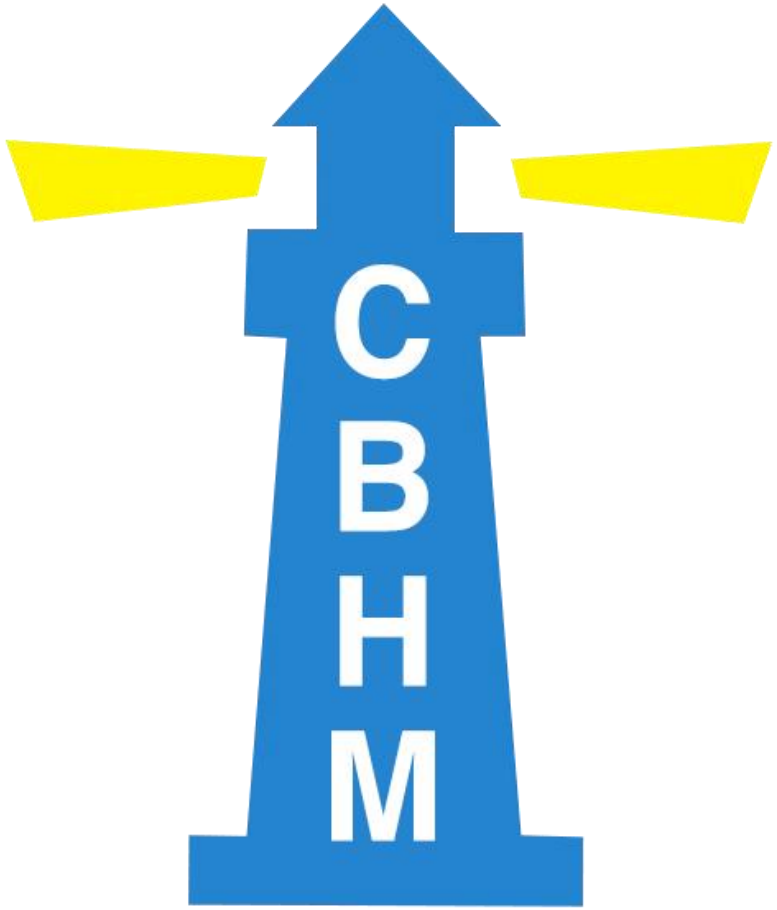


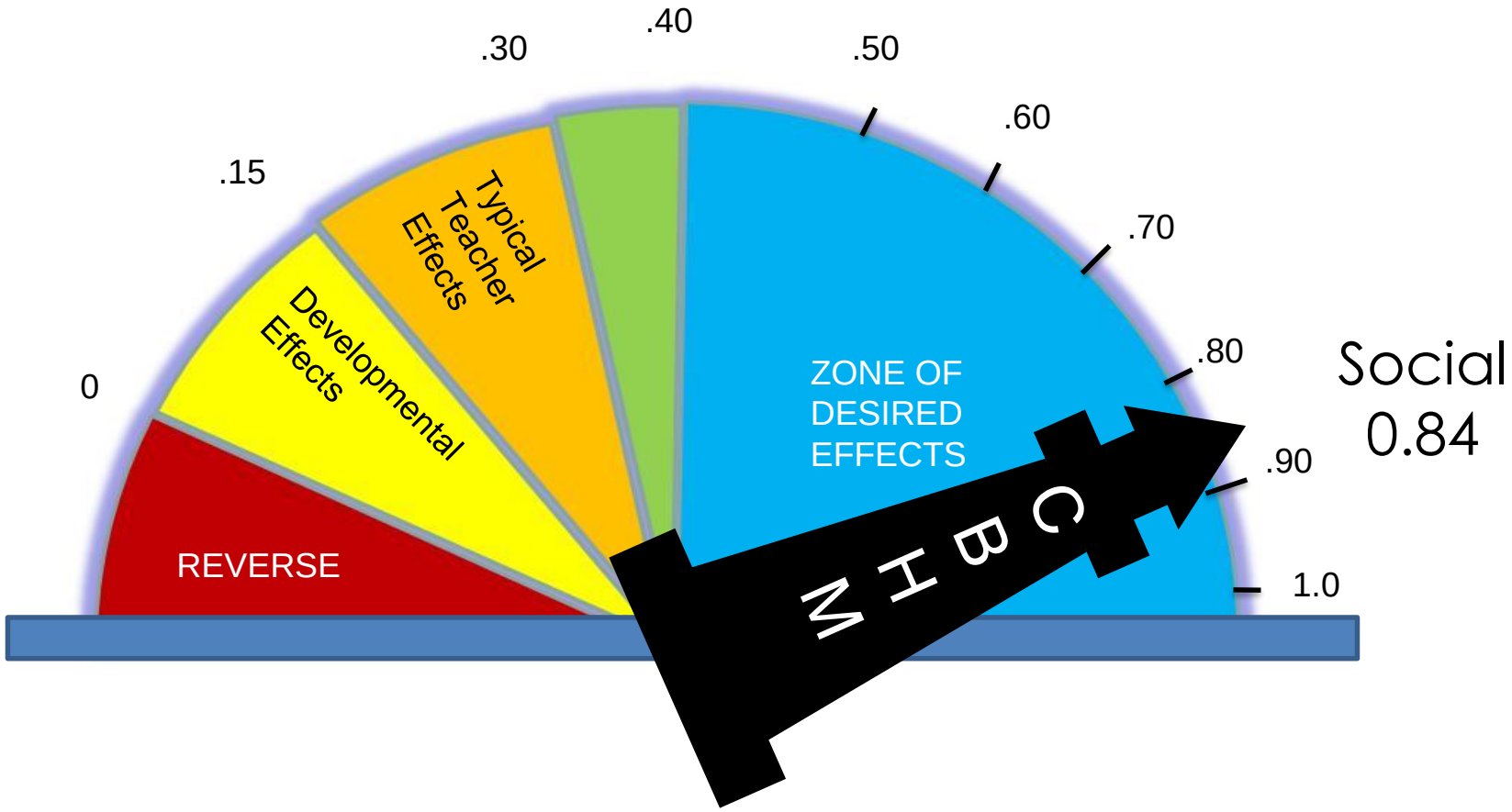
EFFECT SIZES: Impact on Learning



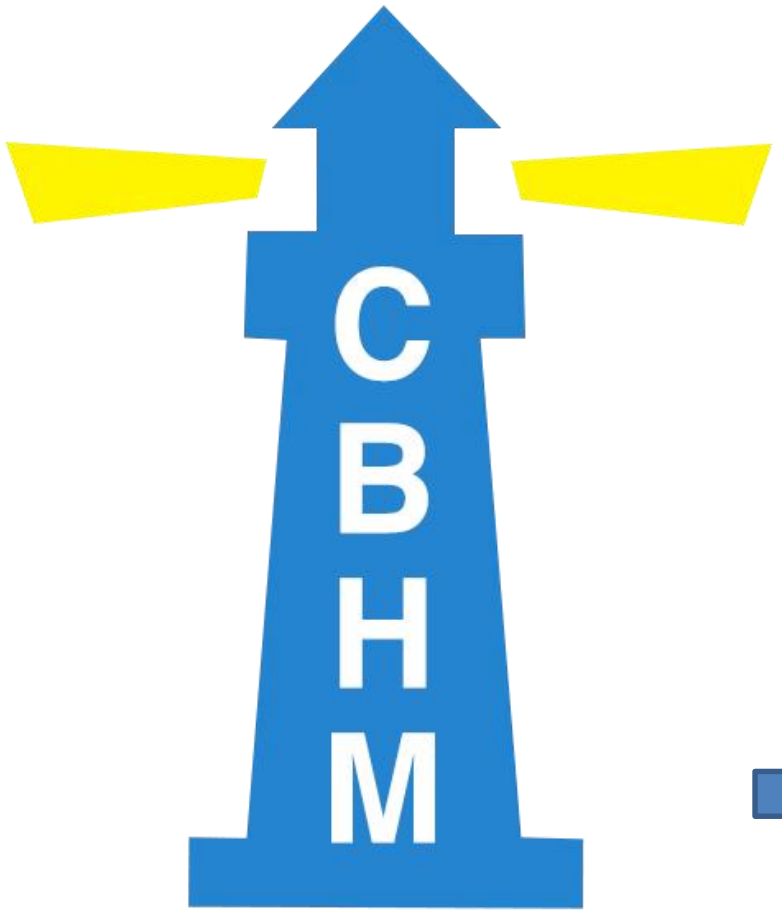


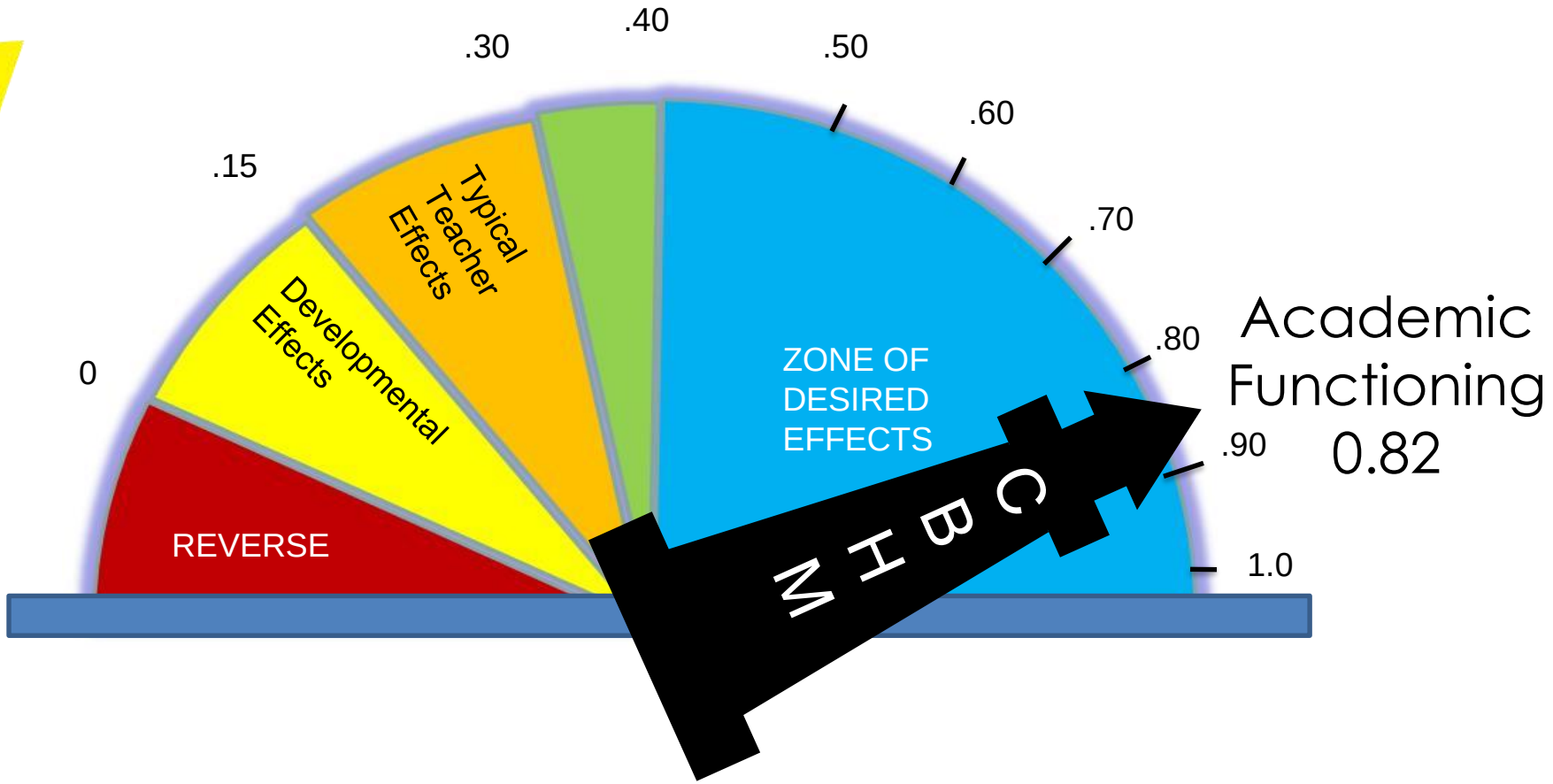
EFFECT SIZES: Impact on Learning





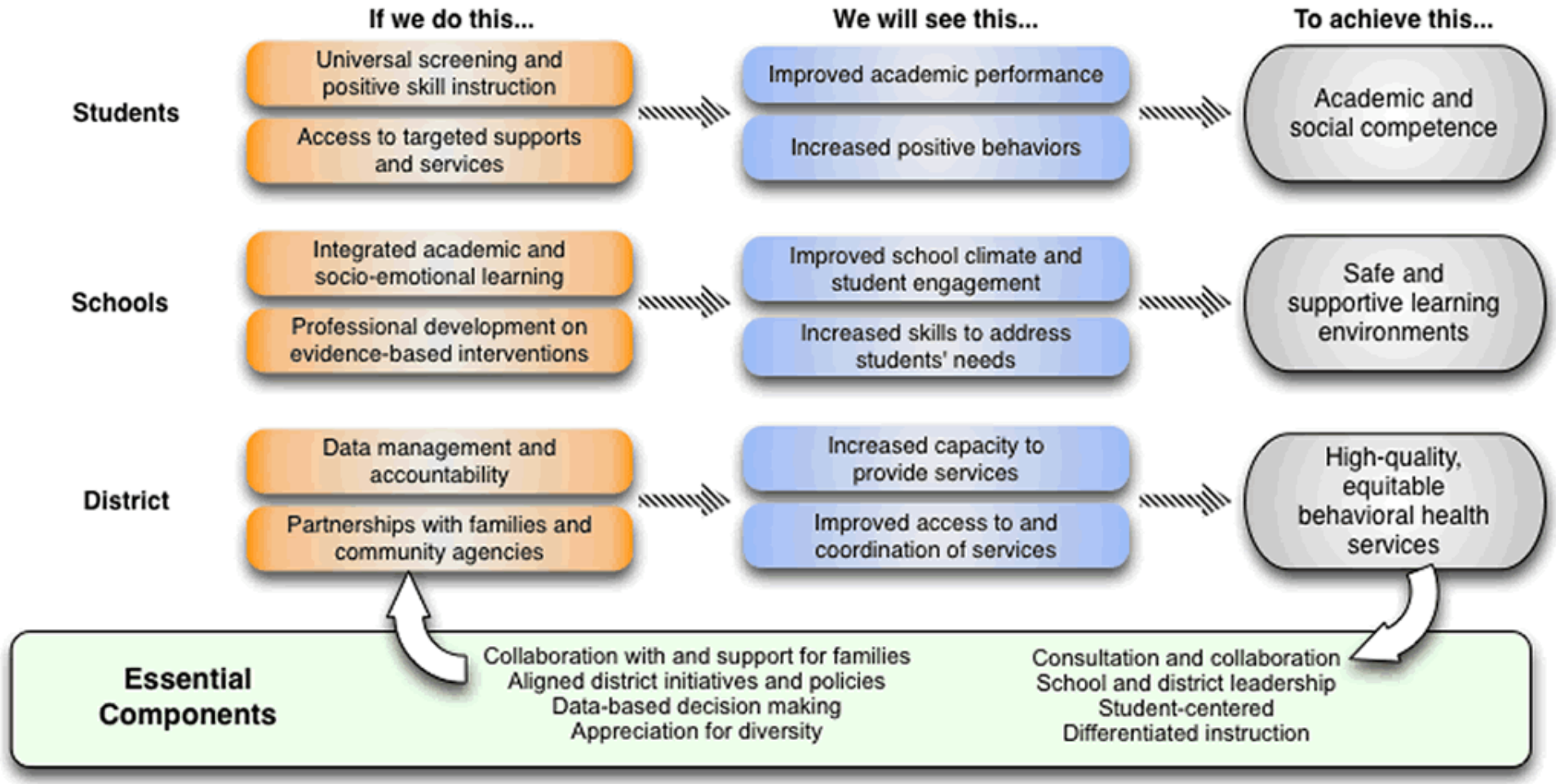
EFFECT SIZES: Impact on Learning





BPS Comprehensive Behavioral Health Model

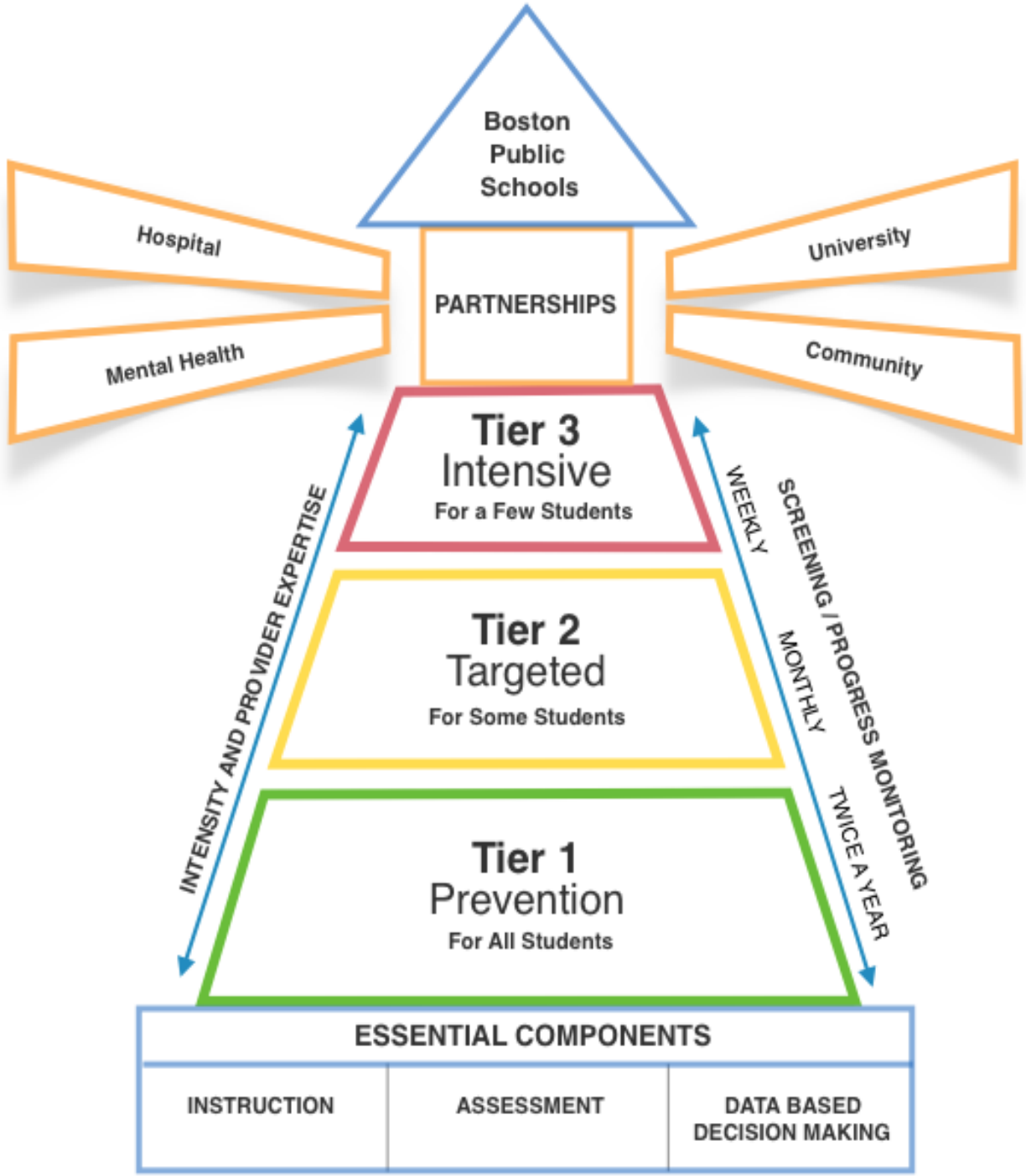
Mission: Ensuring that all students have a safe and supportive school where they can be successful



Guided by Massachusetts Department of Elementary and Secondary Education's Behavioral Health Framework

Theory of Change: Integrating behavioral health services into schools will create safe and supportive learning environments that optimize academic outcomes for all students.

	WHAT	WHY	HOW
INSTRUCTION	• School Wide Positive Behavioral Interventions and Supports (SWPBIS)	Students need to know behavioral expectations throughout the school building in order to be successful in the school environment	Organize the school environment to prevent problem behaviors and reinforce positive behaviors
	• Social Emotional Learning (SEL) Curricula	Students need social and emotional skills to successfully navigate interactions with peers and adults	Instruction in fundamental social skills, such as empathy, relationship building, and conflict management
ASSESSMENT	• Universal Screening	Schools need universal data from all students to understand the strengths of instructional programming, as well as areas of need.	Collect objective information that can be used to guide instruction at multiple levels (e.g. school, grade, class, and individual student)
DATA BASED DECISION MAKING	• Problem Solving Teams & Data Based Decision Making	School teams need to understand how to use universal assessment data to make systemic decisions about instruction	School teams are effectively organized to promote efficient data-based decision making.

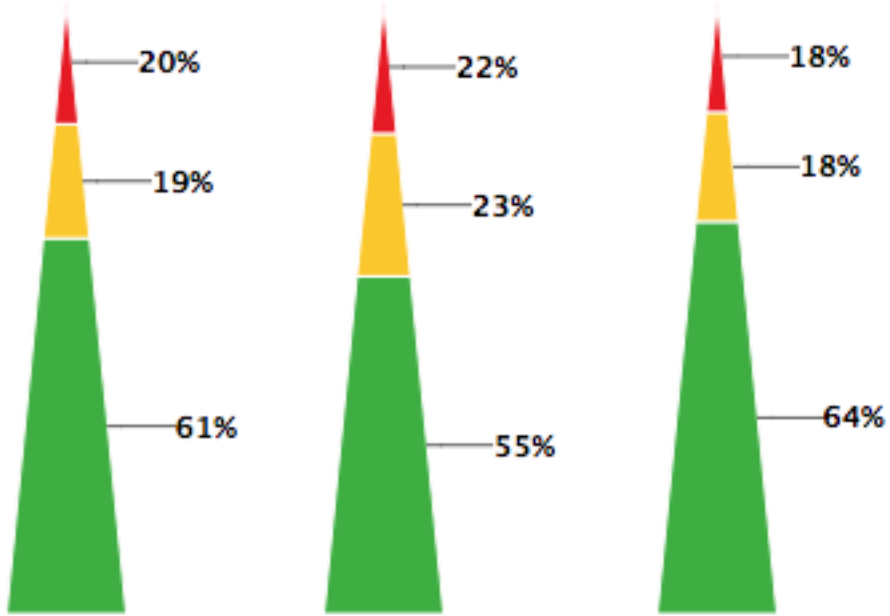

Boston Public Schools
Hospital
Mental Health
PARTNERSHIPS
University
Community
Tier 3
Intensive
For a Few Students
Tier 2
Targeted
For Some Students
Tier 1
Prevention
For All Students
INTENSITY AND PROVIDER EXPERTISE
WEEKLY
MONTHLY
TWICE A YEAR
SCREENING / PROGRESS MONITORING
ESSENTIAL COMPONENTS
INSTRUCTION
ASSESSMENT
DATA BASED DECISION MAKING

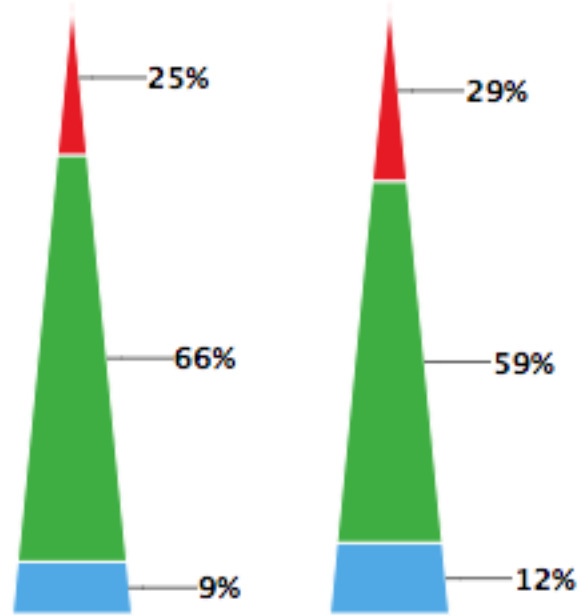
EXPORT AS .PDF

CLASS/GROUP STUDENT SCORES

UA Period: Spring 2017 | School: Sample School | Grade: All | Teachers: All

Student Name	MTSS	Conduct	Negative Affect	Cognitive/ Attention	Social	Academic Functioning
Dory, Finding	1	76	78	65	37	40

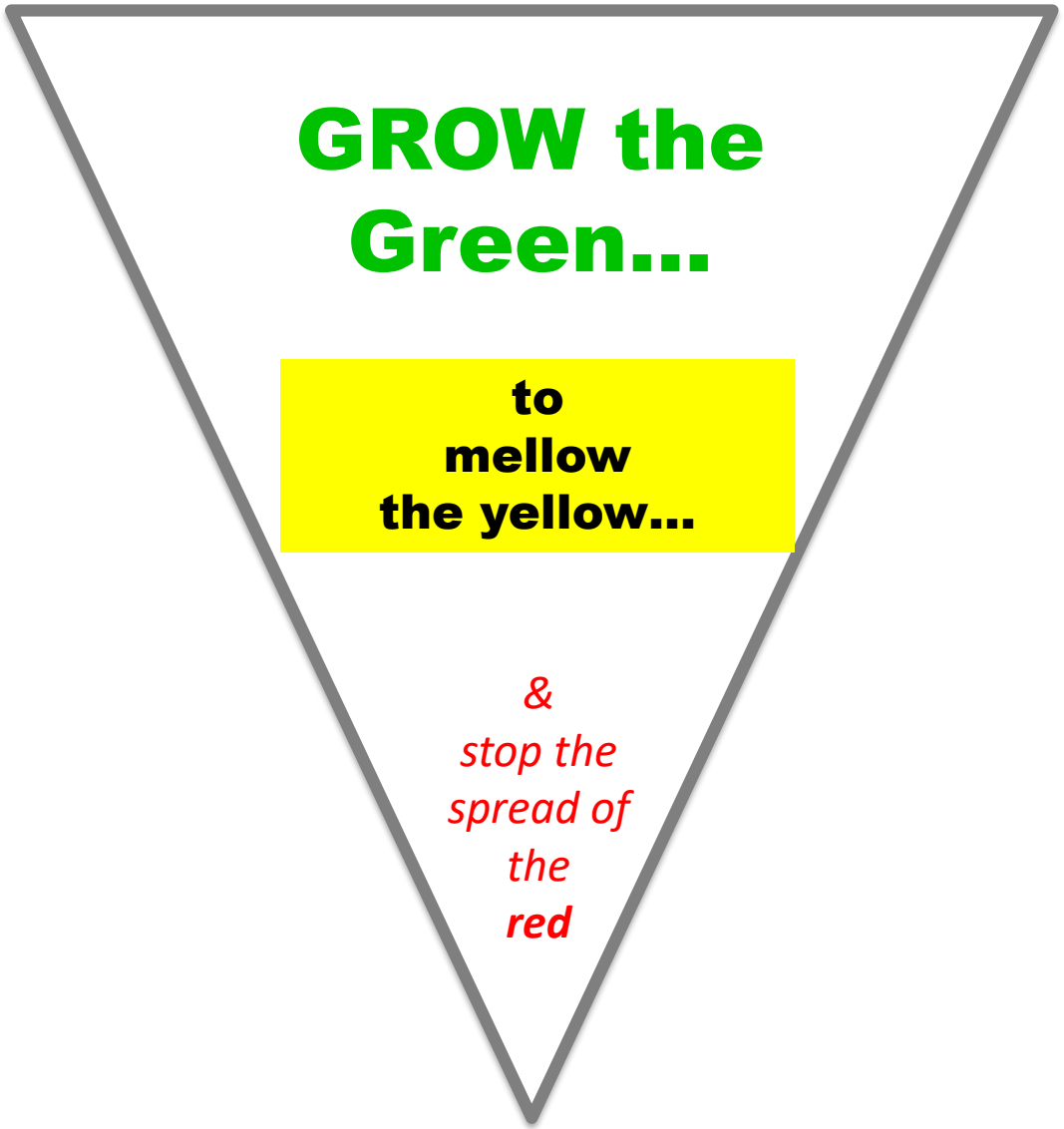




Levels Of Risk	Conduct	Negative Affect	Cognitive/ Attention	Levels Of Functioning	Social	Academic Functioning
High Risk	52 (20%)	56 (22%)	47 (18%)	Concern	65 (25%)	76 (29%)
Some Risk	48 (19%)	60 (23%)	46 (18%)	Typical	171 (66%)	152 (59%)
Low Risk	158 (61%)	142 (55%)	165 (64%)	Strength	22 (9%)	30 (12%)
Total	**258 (100%)**	**258 (100%)**	**258 (100%)**	**Total**	**258 (100%)**	**258 (100%)**

GROW THE GREEN!

- Continuous Improvement
- Developed by schools
  - Grade Level Teams
  - Tier 1 Team
- Goals:
  - Assessment Literacy & Inquiry Cycle
  - Avoid using data as a hammer
  - Help educators/teams reach consensus and action.



GROW THE GREEN!

1. Define the Problem
2. Develop a Plan
3. Implement the Plan
4. Evaluate the Progress

GROW THE GREEN!

- **Behavioral Concern Scales:**
  - **Strength:** Put a checkmark next to the scale that has the greatest percentage of student in green (low risk).
  - **Concern**: Circle the scale that has the lowest percentage of students in the green (low risk)

Conduct	Negative Affect	Cognitive/Attention
Externalizing concerns such as anger management, bullying, substance abuse, deviance	*Internalizing* concerns such as anxiety, depression, withdrawal	Concerns around attention, focus, organization, planning & memory

GROW THE GREEN!

- **Adaptive Scales** - Academic Functioning, Social Functioning
  - **Strength:** Put a checkmark next to the scale that has the greatest percentage of student in green & blue (typical & strength combined).
  - **Concern**: Circle the scale that has the lowest percentage of students in the green & blue (typical & strength combined)

Social	Academic Functioning
Maintaining friendships, communication skills	Academic performance, attendance, following directions