# Different Ways to Measure Fidelity of Implementation of PBIS

"Brown Bag" Presentation

March 5, 2012

Tary J. Tobin, Ph.D.

ttobin@uoregon.edu

Educational and Community Supports

University of Oregon

Eugene, OR

# Acknowledgements

The primary resources for this presentation can be found on the following web sites:

OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports: http://pbis.org

Educational and Community Supports (ECS):

https://www.pbisassessment.org

http://www.pbiseval.org/Resources.aspx

http://www.uoecs.org

# Advance Organizer: 5 Sections

I.     What is fidelity of implementation (FOI)?

II.    What is Positive Behavior Interventions and Supports (PBIS)?

III.   Why take time to measure fidelity?

**IV.  Examples of PBIS measures of fidelity (With some new data about the Team Implementation Checklist)**

V.     Your experiences with this? Your thoughts?

# I.    What is fidelity of implementation (FOI)?

- *Example*: Crash diet plan: Cabbage Soup!
- *Definition*: "the extent to which the delivery of an intervention adheres to the protocol or program model originally developed"

  (Mowbray et al. 2003, p. 315).

# Traditional use of FOI is to decide if the independent variable was well implemented.

**Dependent Variable (DV)**

- *Outcome:* Did we get the results we wanted?

**Independent Variable (IV)**

- *Intervention:* What was the intervention?

# In the past, FOI often was just assumed instead of being measured.

- When it was measured in school research, it often was used by researchers to explain why some schools got good results and some got bad results.

"Analyses of the implementation data revealed considerable variability . . . Meaningful progress toward implementing the CDP [Child Development Project] program could only be said to have been made at five of the 12 program schools" (Battistich, Schops, Watson, Solomon, D., & Lewis, 2000, p. 87).

*Out of 12 treatment schools, only the 5 with good implementation were used to provide evidence of the effect of the intervention; the 7 with poor implementation were not included in the analysis.*

# A different perspective on FOI

- Fidelity measurement is becoming *part* of the independent variable with respect to program implementation in schools.

- "Through the use of self-assessment, teachers themselves become the originators of change and professional learning" (Powell 2000, p. 47).

# II.    What is Positive Behavior Interventions and Supports (PBIS)?

- You probably know!
- Several definitions exist
- Here is a "word bank" and a "fill in the blank" for one definition --

# Word Bank for "Fill in the Blank"

A.   all

B.   diminishing problem behavior

C.   intensive individual

D.   framework

E.   social culture

F.   data-based decision-making

G.   social and academic gains

H.   applied behavior analysis

I.   evidence-based practices

# "Fill in the Blank" for:
## *What is Positive Behavior Interventions and Supports (PBIS)?*

PBIS is a school-wide systems [1]_____for incorporating [2]_____*practices,* including [3]_____, and [4]_____ to establish the [5]_____and [6]_____ behavior **supports** needed for schools to achieve [7]_____while [8]_____for [9]___ students.

# Answers from "Word Bank" for "Fill in the Blank"

## *What is Positive Behavior Interventions and Supports (PBIS)?*

PBIS is a school-wide systems
1 __D_____for incorporating
2 __I_____*practices,* including
3 __H_____, and
4 __F_____ to establish the
5 __E_____and 6 __C_____
behavior **supports** needed for schools to
achieve 7 ___G_____while
8 __B_____for 9 _A__ students.

# What is Positive Behavior Interventions and Supports (PBIS)?

PBIS is a school-wide systems framework for incorporating evidence-based *practices,* including applied behavior analysis, and **data-based** ***decision-making*** to establish the social culture and ***intensive individual behavior*** supports needed for schools to achieve **social and academic gains** while ***diminishing problem behavior*** for **all** students.

# *Important Aspects of Well Implemented PBIS*

- A multi-component process
- Primary, secondary, and tertiary levels of prevention (Tiers 1, 2, & 3)
- Team work
- Uses data for problem solving
- **Includes the use of fidelity measures <u>as an essential element</u>**

# III.   Why take time to measure fidelity?

- Research / Summative Evaluation
- Action Planning / Formative Evaluation

# Research / Summative

- Compare outcome measures, before and after the intervention – did it work? Examples:
  - ◆Fewer suspensions
  - ◆Academic achievement improved
- Quit or continue? – Wait, did we really do it?
- Is PBIS itself – the whole approach -- an "evidence-based intervention?"

# Action Planning / Formative

- Figure out how to improve
- Like self-management
- PBIS is more like a "recipe" than a medication
- Which ingredient is missing or which substitution didn't work?

# IV. Several examples of PBIS measures of fidelity, 2 types:

- Research Tools / External Evaluation
- Self-Assessment Tools / Internal Evaluation

# About this set of examples --

- Available online
- Focused on core elements of PBIS
- Have been used for some time by many schools with good results
- Assessments of their psychometric properties and/or descriptions of the development and use have been published.

*Stay tuned\* – other measures are currently being developed or are "in press" or can be found in connection with specific projects.*

*\*By checking web sites such as the ones mentioned on the 2nd slide (Acknowledgments) or found in the references.*

# Research Tools / External Evaluation

- Direct observations with inter-observer agreement reported

- Interviews

- Review of archival records and permanent products

# Formative, internal evaluation:

*Frequent progress monitoring of how well we are doing with the implementation*

- Are we really using PBIS?

- Using the data to make action plans?

- Checking on how well our own local action plans are followed?

# Research Tools / External Evaluation

- Schoolwide Evaluation Tool (SET)
- Individual Student System Evaluation Tool (ISSET)

# Schoolwide Evaluation Tool (SET)

Sugai, G., Lewis-Palmer, T., Todd, A. W., & Horner, R. H. (2001). *School-wide Evaluation Tool*. Eugene: University of Oregon, Educational and Community Supports. Retrieved from http://www.pbis.org/evaluation/evaluation_tools.aspx

# 7 Constructs in SET

- Behavioral expectations defined
- Behavioral expectations taught
- Consistent reward system
- Consistent violation system
- Continuous monitoring and evaluation
- Management
- District support

# Tier 1: Primary Prevention / Universal Intervention

- To assess the impact of PBIS as a three tiered approach, the SET should be used in combination with measures that address secondary and tertiary levels of prevention:

- "The SET does not offer information about secondary and tertiary prevention efforts in schools" (Horner et al. 2004, p. 11).

# Individual Student System Evaluation Tool (ISSET)

Anderson, C. M., Lewis-Palmer, T., Todd, A. W., Horner, R. H., Sugai, G., and Sampson, N. K. (2011). *Individual Student Systems Evaluation Tool, Version 2.8*. Eugene: University of Oregon, Educational and Community Supports. Retrieved from http://www.pbis.org/common/pbisresources/tools/ISSET_TOOL_v2.8_February_2011.pdf

# The ISSET has 3 Scales:

- **Foundations:** Basic attitudes indicating staff interest in trying to use **positive rather than punitive and exclusionary methods**

- **Targeted:** Interventions that can be **efficiently** provided to students who need extra support; focus on **Tier 2**, Secondary Prevention, Selective Interventions

- **Intensive**: Function-based, **individualized** interventions; focus on **Tier 3**, Tertiary Prevention, Indicated Interventions

# In the works: MATT

- **M**onitoring **A**dvanced **T**ier **T**ool
- Similar to ISSET but does not require an external evaluator –
- Instead, a **PBIS coach** works with the school team to use this for action planning
- Currently being pilot tested at the University of Oregon -- for up-dates in near future, see

https://www.pbisassessment.org and
http://www.pbiseval.org/Resources.aspx

# The MATT has three parts:

- The Coaches' Interview Guide
- The Team Scoring Guide
- The Action Plan

# Self-Assessment Tools / Internal Evaluation & Action Planning

- Benchmarks of Advanced Tiers (**BAT**)

- Team Implementation Checklist (**TIC**)

- Benchmarks of Quality (**BoQ**)

- Staff Self-Assessment Survey (**SAS**) -- also known as Effective Behavior Support Survey (EBS Survey)

- Implementation Phases Inventory (**IPI**)

# Benchmarks for Advanced Tiers (BAT)

Anderson, C., Childs, K., Kincaid, D., Horner, R., George, H., Todd, A. . . . Spaulding, S. (2009). *Benchmarks for Advanced Tiers.* Unpublished instrument. Educational and Community Supports, University of Oregon & Department of Child and Family Studies, University of South Florida. Retrieved from http://pbis.org/common/pbisresources/tools/BAT_v2.5.pdf

# BAT was derived from the ISSET.

- Does not depend on external evaluators or on a PBIS coach

- School teams self-assess Tiers 2 and 3 implementation status

- Then plan for next steps

# Benchmarks of Quality (BoQ)

Kincaid, D., Childs, K., and George, H. (2010). *School-wide Benchmarks of Quality (Revised).* Unpublished instrument. Tampa: University of South Florida. Retrieved from https://www.pbisassessment.org/Evaluation/Surveys

BoQ recently has been revised (Kincaid, Childs, & George, 2010) and assessed when used in different ways (Childs, George, & Kincaid, 2011).

"The Benchmarks of Quality was developed and validated to be administered by the Coach . . . utilizing the Scoring Guide to complete the Scoring Form with feedback from members through the Team Member Rating . . . However, many PBIS Teams and some statewide PBIS Projects desired access and input using the more comprehensive Scoring Guide/Scoring Form"

Retrieved from
http://www.pbis.org/evaluation/evaluation_briefs/mar_11_(2).aspx

# A Tier 1 measure, the BoQ groups 53-items into 10 subscales:

- PBIS team
- Faculty commitment
- Effective discipline procedures
- Data entry
- Expectations and rules

- Reward system
- Lesson plans for teaching behavioral expectations
- Implementation plans
- Crisis plans
- Evaluation

# Team Implementation Checklist (TIC)

Sugai, G, Horner, R.H., Lewis-Palmer, T.,& Rossetto Dickey, C. (2011). *Team Implementation Checklist, Version 3.1*, University of Oregon. Retrieved from https://www.pbisassessment.org/Evaluation/Surveys.

- Focus on the universal level of prevention (Tier 1) although it has 3 items to assess progress toward implementing Tiers 2 and 3.
- For example, item 20 asks if "Personnel are able to provide behavior expertise for students needing Tier II and Tier III support."

# TIC, Version 3.1, has 22 numbered items, organized into 7 features:

- Establish Commitment
- Establish and Maintain Team
- Self-Assessment
- Prevention Systems *(defining and teaching expectations, rewarding appropriate behavior, and responding to violations)*
- Classroom System *(a new category)*
- Information System
- Build Capacity for Function-based Support

# New Research: 2010-2011

- Studied TIC and BoQ concurrent validity
- 893 schools with both TIC and BoQ
- Total scores for each tool were correlated. The Pearson Correlation was statistically significant ($p < .01$), positive, and fairly high (.59)
- Used the last TIC completed during the academic year, which typically would have been completed shortly before the BoQ.

(Tobin, Vincent, Horner, Rossetto Dickey, & May, in review)

# Also studied internal consistency of the third version of the TIC

- To compare results for the current version with results found by Barrett, Bradshaw, and Lewis-Palmer (2008) for an earlier version
- Barrett et al. used a 26-item version with 1,633 completed survey forms and assessed internal consistency for the total scale, finding a Cronbach's alpha of .93.
- Cronbach's alpha for the total scale for the 3rd version of the TIC, based on 22 standardized items, was .91, with 3,408 schools providing data on all items.
- This indicates high internal consistency (Nunnally, 1978; Rosenthal & Rosnow, 1991)

# Also assessed internal consistency for the 7 features of the current TIC

- Keep in mind that scales with **few** items, unless the items refer to **very** similar competencies, tend to have low alphas, and should be interpreted within the context of the overall measure (Achenbach & Rescorla, 2001).

- In addition, the following table compares total possible score with average score for all 3,408 schools completing the online TIC with consent for use in research.

**Internal Consistency of Third Version of the Team Implementation Checklist and Mean Scores**

| Scale | Cronbach's Alpha, Items Standardized | Number of Items | Maximum Possible Score | Average Scale Score for 3,408 Schools |
|---|---|---|---|---|
| Total | .91 | 22 | 44 | 32.39 ($SD = 8.38$) |
| Commitment | .56 | 2 | 4 | 3.36 ($SD = .81$) |
| Team | .57 | 3 | 6 | 4.84 ($SD = 1.17$) |
| Self-assessment | .71 | 3 | 6 | 4.13 ($SD = 1.66$) |
| Prevention | .86 | 6 | 12 | 9.68 ($SD = 2.74$) |
| Classroom | .72 | 3 | 6 | 3.21 ($SD = 1.81$) |
| Information | .87 | 2 | 4 | 2.49 ($SD = 1.46$) |
| Function-based | .72 | 3 | 6 | 4.68 ($SD = 1.52$) |

# 3 Point VS. 2 Point Rating

- Many implementation measures ask teachers to rate specific PBIS elements using a 3-point rating:

  2 = "achieved"

  1 = "in progress"

  0 = "not yet started"

- We conducted an analysis with TIC data comparing that to a 2 point rating, either "achieved" or not.

# Raw Points or Achieved Items?

- A count of raw points includes points for "in progress" (the 3 point rating)

- Because an item rated "in progress" is not yet fully implemented, an alternative calculation of the TIC criterion could be based on the number of items rated "achieved" (2 point rating).

- Like: *Are you doing this? Yes or no?*

# Maximum Points Possible and % Needed to Score as "Implementer"

- TIC        44
- BoQ       107
- 70% of the BoQ total score is considered representative of PBIS implementation.
- Schools scoring 70% on the BoQ typically would score 80% on the TIC.

*Percent of raw points or percent of items fully implemented? Evaluation brief.*

Vincent & Tobin (in review)

- Using BoQ score as the criteria,

- We compared 2 groups:

→ Schools that obtained 80% or more of all TIC **raw points** (using the 3 point rating)

→ Schools that rated 80% or more of all TIC **items** as "achieved" (using the 2 point rating).

*N* = 448 schools

2010-2011

| School Type | Number | Mean enrollment (*SD*) | Mean TIC Total Points (*SD*) | Mean BoQ Total Points (*SD*) |
|---|---|---|---|---|
| Pre-K | 1 | 220 | 29 | 86 |
| Elementary | 282 | 469 (206) | 36.48 (5.73) | 75.15 (13.15) |
| Middle | 89 | 687 (403) | 35.31 (6.38) | 74.11 (13.29) |
| High | 38 | 1008 (741) | 35.89 (6.15) | 69.82 (12.79) |
| K8-12 | 38 | 324 (142) | 36.76 (4.92) | 72.08 (11.97) |
| Alternative | 22 | 375 (243) | 36.55 (4.48) | 71.77 (14.41) |

# Results

- 272 schools met the traditional TIC criterion (80% or more of TIC raw points, using 3 point rating)
- 180 schools met the alternative TIC criterion (80% or more of TIC items rated "achieved," using 2 point rating).
- The 2 point rating system sets a higher standard.
- However, the groups did not differ substantively on the BoQ measure.

| TIC Criterion Met/Not Met | Mean Percent of BoQ points (SD) |
|---|---|
| ≥80% of TIC raw points (n = 272) | 84.17 (10.28) |
| < 80% of TIC raw points (n = 176) | 71.10 (14.10) |
| ≥80% of TIC items "achieved" (n = 180) | 86.74  (9.56) |
| < 80% of TIC items "achieved" (n = 268) | 73.86 (13.34) |

- Of the 180 schools with 80% or more of the TIC items achieved, 93% (n = 167) also met the BoQ implementation criterion of 70% or more BoQ points.

- Of the 272 schools that met the traditional TIC criterion (80% or more TIC raw points), 88% (n = 238) also met the BoQ implementation criterion of 70% or more BoQ points

# Schools were doing well – either way!

- Schools that met the TIC criterion calculated with either method had similar BoQ scores far exceeding the BoQ implementation criterion.

- Schools that did not meet the TIC criterion calculated with either method also had similar BoQ scores close to the minimum BoQ implementation criterion of 70%.

# Staff Self-Assessment Survey (SAS) -- also known as Effective Behavior Support Survey (EBS Survey)

Sugai, G., Horner, R.H., & Todd, A. (2003). *Self Assessment Survey*. University of Oregon, Retrieved from https://www.pbisassessment.org/Evaluation/Surveys.

Todd, A.W., Sugai, G., and Horner, R.H. (2003). *Effective Behavior Support (EBS) Survey.* University of Oregon, Retrieved February 2, 2012, from http://pbis.org/pbis_resource_detail_page.aspx?Type=4andPBIS_ResourceID=219.

# SAS

- Measures implementation level and priority for improvement using 46 items across four systems:

- School-wide (18 items)

- Specific Setting (9 items)

- Classroom (11 items)

- Individual Student (8 items)

# Often is used as part of staff development with initial PBIS training.

- Measures "Priority for Improvement" as well as staff perceptions of level of implementation ("in place" or "partially in place" or "not in place")

-  Many schools continue to use this as a measure of progress over time.

- Data from the online version of SAS was sensitive to the value of using School Wide Information System (SWIS, May et al. 2006) when implementing PBIS (Tobin 2006a).

- Also, SAS data made it possible to identify **specific strategies** associated with reductions in disproportionate suspensions of African American students (Tobin & Vincent, 2011).

# Implementation Phases Inventory (IPI)

Bradshaw, C. P., Barrett, S., & McKenna, M. (2008). *Implementation Phases Inventory* (IPI), Baltimore: PBIS Maryland. Retrieved from http://www.pbismaryland.org/forms.htm.

# IPI

- Measures all 3 tiers.
- 44 items clustered into 4 stages of implementation:

  Preparation (10 items)

  Initiation (13 items)

  Implementation (11 items)

  Maintenance (10 items)

# Concurrent validation of the IPI

- Occurred through correlation of IPI scores with scores on the SET and TIC.

- SET and TIC both focus on Tier 1 but IPI also covers Tier 2 and Tier 3.

- Results indicated moderate correlations for total scale scores:

  IPI-SET ($r = .46$, $p < .01$)

  IPI-TIC ($r = .58$, $p < .01$)

# Summary of Our Sample of <u>Currently Available Online</u> Tools

- External Evaluation / Research:

  SET for Tier 1

  ISSET for Tier 2 and Tier 3

- Internal Evaluation / Action Planning:

  BAT for Tier 2 and Tier 3

  TIC mostly for Tier 1, does have some on Tiers 2 & 3

  BoQ for Tier 1

  SAS mostly for Tier 1, does have some on Tiers 2 & 3

  IPS for all Tiers + looks at phases in the process

# Considerations in Selecting a Tool

- What decisions need to be made? Why do you want to measure implementation fidelity?
- Which of the 3 tiers is of most concern?
- Can you afford external evaluators?
- Does the school staff have time to do internal evaluations?
- Does the school have a PBIS coach who can help?

For more information, see:
[http://pbis.org](http://pbis.org)
[https://www.pbisassessment.org](https://www.pbisassessment.org)
and the separate hand out on references

# Big Picture Questions

- Do you also have outcome data on social behaviors, discipline, academic achievement?

  Can this be related to the PBIS implementation?

- Are you interested in specific groups of students?

  Consider disaggregating data by ethnicity, race, gender, grade, etc.

- Besides PBIS, what other interventions are going on that may have affected outcomes – or staff willingness to participate?

V. Your experiences with this?
Your thoughts?