



CENTER ON

PBIS

Positive Behavioral
Interventions & Supports



PBIS EVALUATION BLUEPRINT

Version 2 | December 2020



Positive Behavioral Interventions and Supports (PBIS) Evaluation Blueprint

Center on Positive Behavioral Interventions and Supports
U. S. Department of Education, Office of Special Education Programs
and Office of Elementary and Secondary Education

Version 2.0 | December 2020

www.pbis.org

This document was supported from funds provided by the Center on Positive Behavioral Interventions and Supports cooperative grant supported by the Office of Special Education Programs (OSEP) and Office of Elementary and Secondary Education (OESE) of the U.S. Department of Education (H326S180001). Dr. Renee Bradley serves as the project officer. The views expressed herein do not necessarily represent the positions or policies of the U.S. Department of Education. No official endorsement by the U.S. Department of Education of any product, commodity, or enterprise mentioned in this document is intended or should be inferred..

Suggested Citation for this Publication

Center on Positive Behavioral Interventions and Supports (December 2020). *Positive Behavioral Interventions and Supports (PBIS) Evaluation Blueprint*. University of Oregon. www.pbis.org.



Table of Contents

Introduction	5
Purpose	5
Intended Users	5
Using this Blueprint	5
<i>Engage Stakeholders</i>	6
<i>“Do Less Better”</i>	6
<i>Utilize Existing Data When Feasible</i>	6
<i>Evaluate Effectiveness Only When Implemented with Fidelity</i>	6
<i>Use Evaluation Plans for Continuous Improvement, not to Punish.</i>	6
Planning Evaluations	7
1. Identify the Purpose	7
2. Identify Key Stakeholders	7
3. Select the Evaluation Team	8
4. Define the Timeline	9
5. Define Core Evaluation Questions	9
Table 1. Common Evaluation Questions by Area and Measures	10
<i>Reach: “Who is participating in PBIS?”</i>	11
<i>Process: “What is happening with the PBIS initiative?”</i>	11
<i>Capacity: “What is the ability of the organization to implement and sustain PBIS?”</i>	12
<i>Fidelity: “Are the core features of PBIS being implemented?”</i>	12
<i>Outcomes: “Is the initiative achieving valued outcomes and worth sustaining?”</i>	12
6. Select Evaluation Measures	14
Measures for Reach Evaluation	15
Measures for Process Evaluation	15
Measures for Capacity Evaluation	16
Measures for Fidelity Evaluation	17
Measures for Outcomes Evaluation	19
Conducting Evaluations	22
Reporting Evaluation Results	24



1. Introduction. 24

2. Process 25

3. Results 25

Reach: "Who is participating in PBIS?" 25

Process: "What is happening with the PBIS initiative?" 28

Capacity: "What is the capacity of the organization to implement and sustain PBIS?" 31

Fidelity: "Are the core features of PBIS being implemented?" 33

Outcomes: "Is the initiative achieving valued outcomes and worth sustaining?" 38

4. Summary and Recommendations 46

Using Results for Continuous Improvement 47

1. Share Results with Stakeholders 47

2. Use Evaluation Results to Improve the Initiative 49

3. Identify Improvements in the Evaluation Process for the Next Evaluation Cycle . . . 50

Summary 51

References. 52

Appendix A: Sample Professional Development Fidelity Assessment Form .54

Appendix B: Sample Professional Development Evaluation Form 56

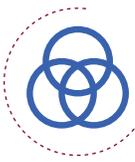
Appendix C: Additional Measures for Evaluation Plans. 57

Fidelity of Implementation Measures 57

Less Common Outcome Measures 60

Appendix D: Sample Evaluation Plan 61

Additional Resources 64



Introduction

Purpose

The PBIS Evaluation Blueprint provides guidance regarding the evaluation process, tools, and outcomes that guide both initial adoption and sustained use of PBIS. Evaluation is the process of gathering and reporting information to address questions and inform decisions. Evaluation is a core function of PBIS implementation, and evaluation reports from schools, districts, and states have been essential for guiding and improving PBIS practices, systems, and data. Evaluation often sounds straightforward until it is attempted. A goal of this Blueprint is to provide a suggested format, structure, and materials to support a wide range of evaluation plans.

The PBIS Evaluation Blueprint functions as a companion and supplement to the PBIS Implementation Blueprint and the PBIS Professional Development Blueprint (also available at www.pbis.org).

Intended Users

The PBIS Evaluation Blueprint was developed for PBIS evaluators, implementers, and researchers supporting organizations, primarily for states (and possibly large districts). It is not intended for school teams or those seeking to evaluate teacher or school effectiveness (e.g., identifying or evaluating turnaround schools).

Using this Blueprint

Effective evaluations vary in scope and complexity, yet most include an iterative process of collecting, compiling, and using information for decision-making. Evaluation is seldom a one-time event—it is most

helpful when data are collected repeatedly (if not continuously) and summarized at regular intervals and analyzed to inform decisions. This blueprint provides a model for identifying and answering key evaluation questions using an iterative approach to document the **reach, process, capacity, fidelity, and outcomes** of a PBIS initiative.

The sections below start with a set of guiding principles and provide details and options for the common steps of (1) planning evaluations, (2) conducting evaluations, (3) reporting evaluation results, and (4) using results for continuous improvement.

The figure below illustrates how each of these steps are combined into iterative cycles to assess and act on evaluation results to continuously improve PBIS implementation efforts.

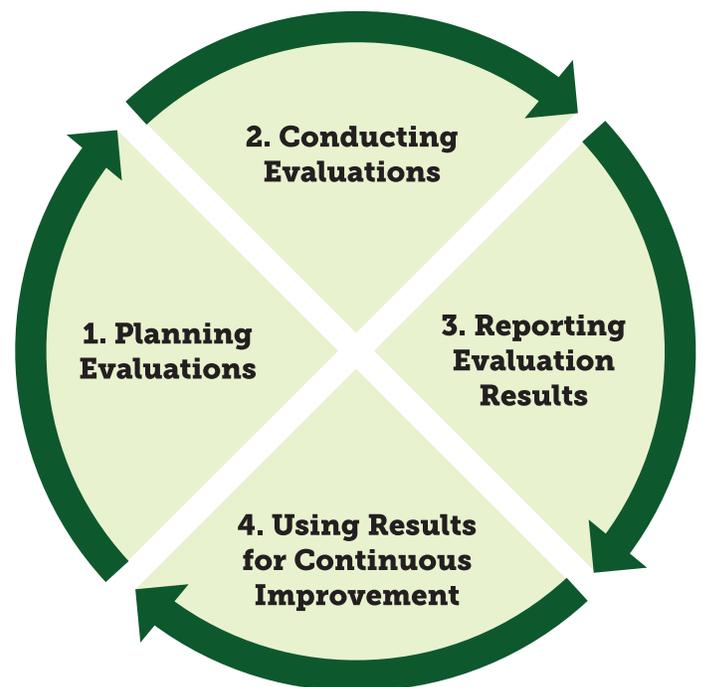
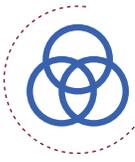


Figure 1. The Full Evaluation Cycle.



Guiding Principles

In our work with schools, districts, and states, there are a number of lessons learned that have shaped our recommendations for evaluations. Although each evaluation has its own nuances, here are some common themes that can guide teams.

ENGAGE STAKEHOLDERS

Creating an evaluation plan without considering the report's audience and who needs what information decreases the likelihood that it will be effective in improving implementation or sustaining the initiative.

"DO LESS BETTER"

A plan that effectively answers fewer questions is far superior to a plan that ineffectively answers more questions. This blueprint provides a wider range of options for questions, measures, and reporting that most evaluations will cover. It is important to be judicious, especially in the organization's first evaluation plan. Consider adding additional questions only when the context demands it and the team has built experience and capacity for a more comprehensive evaluation plan.

UTILIZE EXISTING DATA WHEN FEASIBLE

All else being equal, it is preferential to answer questions with existing data rather than collect new data. Adding measures to a school system should be done with caution, given that the addition of each measure typically requires added effort and cost. However, it is often necessary to collect some new data to answer questions that are important to stakeholders. The desired balance is to select only the additional measures needed to answer the core questions.

EVALUATE EFFECTIVENESS ONLY WHEN IMPLEMENTED WITH FIDELITY

Evaluation plans should examine student outcomes only when PBIS has been implemented to criterion. It is important to identify first whether training and technical assistance resulted in improved social-emotional-behavior support practices in schools (as measured by increased scores on fidelity measures). An important next question is whether those schools that implemented to criterion saw improved student outcomes. If schools are not implementing PBIS with fidelity, we do not expect to see changes in student outcomes.

USE EVALUATION PLANS FOR CONTINUOUS IMPROVEMENT, NOT TO PUNISH

Many educators are used to evaluations that identify deficiencies, whether in schools or individual teachers. In contrast, evaluations are most effective when they identify current strengths and what the organization can do to improve its support for districts, schools, and educators. We use the evaluation process to promote continuous improvement through regular cycles of data review and iterative action planning that builds on program strengths while addressing areas of need. This approach to using evaluation data focuses on identifying the conditions necessary to improve outcomes and supports needed to improve these conditions.



Planning Evaluations

There are many approaches and procedures for conducting effective evaluations. We have learned to focus first on how an evaluation is planned before starting data collection. Taking the time to plan in detail can substantially improve the effectiveness and efficiency of evaluation. Successful evaluations follow a clear process, such as: (1) Identify the Purpose, (2) Identify Key Stakeholders, (3) Select the Evaluation Team, (4) Define the Timeline, (5) Define Core Evaluation Questions, and (6) Select Evaluation Measures.

1. Identify the Purpose

In healthy organizations, evaluations are ongoing and often serve multiple purposes. Everyone involved in the evaluation should know its purpose (or purposes). The following are possible options:

- **Summative** evaluations provide information that decision makers use to document the overall implementation and value of interventions (or initiatives) based on the extent to which outcomes were improved or goals were met at a particular point in time (e.g., annually). Summative evaluations inform decisions about whether to continue, abandon, or modify an initiative.
- **Formative** (i.e., progress monitoring) evaluations provide information that decision makers use to review interventions as they are implemented, with an intent on improving fidelity of implementation or assessing short-term improvement in outcomes. They often occur more frequently and are



more limited in scope than summative evaluations.

- **Accountability** (i.e., compliance-driven) evaluations focus on documenting the presence or use of components of targeted procedures or practices (e.g., number of individualized education programs that include social-emotional-behavior goals). They typically reflect adherence to policy or regulatory requirements and indicators of the quality of the procedures or practices, although they can (and often should) also include elements of summative or formative evaluations.

This blueprint primarily covers summative assessments, although the same processes and tools can be used for a range of purposes. In addition to identifying a general purpose, it is useful to identify how specifically the evaluation report will be used. What are the decisions that will be made based on the results? Is it to identify the extent to which the PBIS initiative is meeting organizational goals and objectives? Is it how implementation can be optimized?

2. Identify Key Stakeholders

An evaluation is typically done at the request or direction of stakeholders. Stakeholders are individuals who will be affected by or make decisions based on the evaluation results. Unfortunately, stakeholders are often treated as passive observers in evaluation, which seriously limits its effectiveness. Effective evaluators identify key stakeholders early and work closely with them to define the purpose(s) and questions for an evaluation. Essentially, stakeholders will make



their own judgments about intervention importance and effectiveness, so involving them from the start provides significant benefit. Key stakeholders to include in evaluations are:

- Family members
- Students
- Educators (e.g., teachers or other direct implementers)
- Administrators and board members (school, district, region, state)
- Community partners and agency representatives and partners

Stakeholders can identify issues and develop strategies to address these issues. It is important to directly involve key stakeholders when their input is not readily accessible or if direct involvement contributes to credibility and the likelihood that evaluation results are used for program improvements (Bryson, 2018). Moreover, it is useful to consider strategies to obtain full participation by stakeholders, such as creating clearly defined roles and providing regular, explicit opportunities for each member to share their voices, including public and private communication channels.

3. Select the Evaluation Team

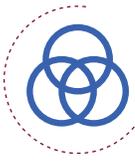
The evaluation team is the individual or group of individuals who will ensure that relevant data are collected, summarized, and reported to the stakeholders. The evaluation team is responsible for the integrity of the entire process. Members of the evaluation team play critical roles, including preparing the evaluation reports with an aim for straightforward

Guiding Questions for Identifying Key Stakeholders

- Does the potential stakeholder influence policy for PBIS implementation?
- Does the potential stakeholder influence or manage budgets that contribute to PBIS implementation?
- Does the potential stakeholder allocate resources for use to implement PBIS?
- Does the potential stakeholder provide public and/or visible support through communication that PBIS is a priority?
- Is the potential stakeholder a current collaborative partner in supporting PBIS?
- Is the potential stakeholder a direct or indirect recipient of PBIS?

communication and serving as a resource to stakeholders as decisions are made based on the evaluation results.

Evaluation teams exist on a spectrum, from internal evaluation teams to external evaluation teams. Internal evaluators typically are staff members who have a working knowledge of the intervention in the context in which it is being evaluated. External evaluators are non-staff members (e.g., contractors) who serve as more objective arms-reach observers to increase credibility of the evaluation. They may or may not be familiar with the specific intervention, yet have the skill sets necessary to complete independent evaluations. Considerations in determining where the evaluation team will exist in on the spectrum from internal



to external include accountability requirements, availability of expertise, and cost.

Once established, the team may revisit steps 1 (purpose) and 2 (stakeholders). The reason for establishing the team at step 3 is that the makeup of the team should be determined by the purpose and stakeholders. The team can then propose changes to the initial plan as needed.

4. Define the Timeline

The evaluation timeline identifies when data are to be collected, summarized, and reported to stakeholders. Although it is common for an evaluation effort to be an ongoing activity, at least annually the evaluation team should review and receive approval for a timeline that allows adequate opportunities for data collection, summary, and decision making based on formal evaluation reports. The timeline and progress toward completion can be communicated regularly.

Note that in development of the timeline, an evaluation team typically summarizes the costs and personnel commitments needed for an effective evaluation.

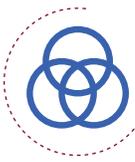
5. Define Core Evaluation Questions

Among the most important decisions in an evaluation is selection of the questions that will guide decision-making. Each evaluation is unique, and the demands of each situation will shape the evaluation questions needed by stakeholders. In addition, some questions will be valued more by some stakeholders than others. Although there are infinite variations, we have found it useful to consider evaluation questions within the following areas: reach, process, capacity, fidelity, and outcomes. Within each area, we include common evaluation questions that have been used. It is unlikely any one evaluation would address all these questions, but most effective evaluations will address some questions from each content area.



Table 1. Common Evaluation Questions by Area and Measures

Area	Common Questions	Common Measures
Reach. Who is participating in PBIS?	<ul style="list-style-type: none"> • How many students/schools/districts/community organizations are involved? • To what extent has PBIS implementation scaled across the district/region/state? • To what extent has PBIS implementation sustained? 	<ul style="list-style-type: none"> • Counts of schools • Counts of students
Process. What is happening with the PBIS initiative?	<ul style="list-style-type: none"> • What leadership team activities have been completed? • What PBIS professional development (i.e., training, coaching, evaluation) has been delivered? • To what extent was the professional development delivered with fidelity? 	<ul style="list-style-type: none"> • DSFI/TFI Action plan completion • PD Calendar • PD activity evaluations • SISS: Stakeholder Input and Satisfaction Survey
Capacity. What is the ability of the organization to implement and sustain PBIS?	<ul style="list-style-type: none"> • What is the organization's capacity to implement and sustain PBIS? • What resources are available to support PBIS adoption, installation, and sustained implementation? • To what extent has PBIS implementation improved capacity for the organization to replicate, sustain, or scale up PBIS? 	<ul style="list-style-type: none"> • DSFI: District Systems Fidelity Inventory • DCA/RCA/SCA: District/Regional/State Capacity Assessment • SSFI: State Systems Fidelity Inventory
Fidelity. Are the core features of PBIS being implemented?	<ul style="list-style-type: none"> • What percent of implementing schools measured fidelity of implementation? (Tier 1, Tier 2, Tier 3) • To what extent is PBIS implemented as intended in schools? (Tier 1, Tier 2, Tier 3) • Are the specific practices within PBIS implemented as intended? • What schools can serve as model schools for local implementation demonstrations? 	<ul style="list-style-type: none"> • TFI: Tiered Fidelity Inventory • Team Meeting Fidelity Checklist • SEB Lesson Plan calendar • CICO-FIM Tool • Wraparound Integrity Tool
Outcomes. Is the initiative achieving valued outcomes and worth sustaining?	<ul style="list-style-type: none"> • To what extent do schools implementing PBIS with fidelity show desired changes in student outcomes? (e.g., discipline rates, academic achievement, graduation, SEB outcomes, student satisfaction, equity) • To what extent do schools implementing PBIS with fidelity show desired changes in other areas of schooling? (e.g., adult perception, staffing, policies) 	<ul style="list-style-type: none"> • Office Discipline Referrals (ODRs) • Exclusionary Practices Data (isolated time-out, seclusion, detention, suspension, expulsion) • Restraint Data • Disproportionality Data • Academic Achievement • Attendance • Tardiness • SEB Competence Measures • School Climate Surveys



REACH: “WHO IS PARTICIPATING IN PBIS?”

Reach questions are intended to identify who is participating in PBIS and are affected by it. Primarily, reach indicates the number of schools (and students enrolled in those schools) that are part of the PBIS initiative, often updated yearly from previous evaluations. Answers are important to determining the scale and potential benefit of the initiative, as well as areas for expansion and sustained implementation. Results could include raw counts or percentages of schools. Evaluations often also identify other agencies or organizations that are participating (e.g., community agencies, other departments or governmental organizations).

Common Reach Questions

- a. How many districts/schools/students are involved?
- b. What levels and types of schools are participating (e.g., elementary schools, high schools, alternative schools, secure facilities)?
- c. What additional organizations (e.g., community agencies, related departments) are participating?
- d. To what extent has PBIS implementation scaled across the [district/region/state]?
- e. To what extent has PBIS implementation sustained over time?

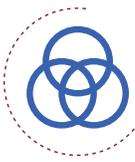
PROCESS: “WHAT IS HAPPENING WITH THE PBIS INITIATIVE?”

Process questions ask what PBIS activities or tasks have been accomplished. They serve to document

important installation and ongoing implementation activities. Typical activities that are measured include counts of professional development (e.g., training, coaching, evaluation) events and assessments of fidelity of event delivery. In addition, an important aspect of process evaluation is feedback from event participants, measured through workshop evaluations and open-ended feedback for improvement. Finally, just like feedback from event participants, process evaluation includes stakeholder input regarding the PBIS initiative, as measured by surveys and open-ended feedback from users and recipients (e.g., school personnel, students, families).

Common Process Questions

- a. What activities on the leadership team’s action plan have been completed?
- b. What are some key accomplishments?
- c. What PBIS professional development (i.e., training, coaching, evaluation) has been delivered?
- d. Who has provided the professional development?
- e. To what extent was the professional development delivered with fidelity?
- f. Have participants perceived professional development events as useful and effective?
- g. How should professional development be modified to improve fidelity or efficiency in PBIS implementation?
- h. How can the PBIS initiative be made more effective, efficient, and equitable, according to PBIS users and recipients?



CAPACITY: “WHAT IS THE ABILITY OF THE ORGANIZATION TO IMPLEMENT AND SUSTAIN PBIS?”

Answers to capacity evaluation questions assess the extent to which the host environment can build and expand a healthy PBIS initiative. Issues regarding capacity include questions and answers to support replication, sustainability, and scale-up of future PBIS implementation. One component of assessing capacity is documenting the resources assigned to the initiative, but another is assessment of other elements of capacity, which can be evaluated through measures in the same way that fidelity of PBIS is measured in schools.

Common Capacity Questions

- a. What resources are available to support PBIS adoption, installation, and sustained implementation?
- b. What capacity does the organization have to support the PBIS initiative?
- c. To what extent has the initiative improved capacity for the organization to replicate, sustain, or scale up PBIS?

FIDELITY: “ARE THE CORE FEATURES OF PBIS BEING IMPLEMENTED?”

Fidelity evaluations document the extent to which PBIS was implemented as intended at each school and each tier. These questions are best answered with validated fidelity assessments that are facilitated by an external (e.g., district) coach at least once per year. In addition, fidelity data can be used to identify model schools for local implementation demonstrations, at

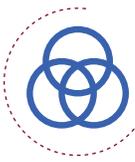
each tier and school type or level (e.g., elementary, middle, high).

Common Fidelity Questions

- a. What percent of implementing schools measured fidelity of implementation?
 - Tier 1
 - Tier 2
 - Tier 3
- b. To what extent is PBIS implemented as intended in schools?
 - Tier 1
 - Tier 2
 - Tier 3
- c. To what extent are specific practices within the PBIS framework implemented as intended in schools?
 - Social-emotional learning (SEL)/social skills programs
 - Bullying prevention programs
 - Universal screening procedures
- d. What schools can serve as model schools for local implementation demonstrations (across tiers and levels)?

OUTCOMES: “IS THE INITIATIVE ACHIEVING VALUED OUTCOMES AND WORTH SUSTAINING?”

Answers to outcomes evaluation questions document the extent to which PBIS implementation was associated with intended benefits and outcomes for students, school personnel, families, and the larger school community. The main purpose of evaluating outcomes is to provide bottom-line information to



stakeholders about initiative effectiveness. The data produced should provide results about the impact of PBIS on valued outcomes (as determined by stakeholders), some of which are listed below. This information should then lead directly into decisions regarding maintaining or increasing existing resources, expansion to other sites or areas, and how to improve initiative effectiveness.

Remember: “Do Less Better”

Select the outcomes to track that are most important to stakeholders. Combine these with any outcomes that are required for reporting (i.e., in state or federal reports) or are organizational priorities (e.g., critical initiatives) to create a core set of outcomes for evaluation. Add more outcomes only when the team has built more evaluation capacity.

Any outcomes questions require a comparison group for judging intervention effectiveness, determined during the planning process. There are numerous options for comparison, including district or state averages, averages for all schools not implementing PBIS with fidelity, averages for non-implementing schools with similar demographics, or school performance in years before PBIS implementation (pre-post comparison, with schools serving as their own comparison). Regardless of the choice, only schools implementing PBIS with fidelity should be in the PBIS group.

Common Outcomes Questions

- a. To what extent do schools implementing PBIS with fidelity show desired changes in student outcomes?

Guiding Questions for Defining Outcomes Questions

- What are the priorities of the community?
- What priorities are shared across stakeholder groups?
- What outcomes are already required for reporting (e.g., state, federal)?
- What outcomes would be meaningful for stakeholders?
- What outcomes would be considered worth investing in?
- What outcomes are important to funders who might support the work (e.g., U.S. Department of Education)?

- Discipline rates
 - Office discipline referrals
 - Suspension, expulsion, or school arrests
 - Restraint or seclusion
- Academic and related outcomes
 - Academic achievement
 - Attendance and chronic absenteeism
 - Graduation
 - Career and college readiness
- Social-emotional-behavior (SEB) outcomes
 - Social skills and social and emotional competencies
 - Mental health outcomes
 - Substance use/abuse
- Perceptions of the learning environment (e.g., student, family, school personnel)



Consider Denominators Carefully

In well-planned evaluation reports, the samples will vary depending on the question being asked. To answer each question accurately, different denominators are needed. A common evaluation logic is as follows:

Question	Denominator	Numerator
Reach	Of all schools in the organization...	...what percentage are in the PBIS initiative?
Fidelity	Of all schools in the PBIS initiative...	...what percentage are implementing with fidelity?
Outcomes	Of all schools implementing PBIS with fidelity...	...what percentage have improved outcomes?

- School climate survey scores
- School safety survey scores
- Special education decisions
 - Eligibility referrals
 - Eligibility determinations
 - Out-of-school or out-of-district placements
- Equity in each of the above outcomes, by
 - Race/ethnicity
 - Special education classification
 - Language status
 - Gender/gender identification
 - Sexual orientation

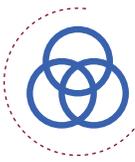
- Faculty/staff retention
- Educational policy
 - Guidance on selecting practices
 - Implementation supports

b. To what extent do schools implementing PBIS with fidelity show desired changes in other areas of schooling?

- Adult perceptions
 - School climate survey scores
 - School safety survey scores
 - Organizational health
 - Self-efficacy and burnout
- Staffing
 - Faculty/staff attendance

6. Select Evaluation Measures

Three general rules are useful for selecting measures. The first is to ensure that all measures are tied directly to the evaluation questions. The second is to consider using existing data sources (instead of collecting new data) when possible. The third general rule is to select measures that will provide credible information. This requires that the evaluation measures provide information that is reliable and valid. In some cases, this criterion is met by simple counts, such as how many school personnel attended a professional development event. In other cases, however, significant care is needed to clarify how data were collected and summarized. Reporting office discipline referral data is such an example. To interpret a measure of office discipline referrals, it is necessary to know (a) the operational definition for a behavior incident, (b) the rules used by staff to determine when an office discipline referral is issued, and (c) how office discipline referrals are reported (e.g., raw count, rate).



Below we offer examples of measures that have been useful in prior PBIS evaluations. Note that we are not recommending that all the measures below be used in a single evaluation. Instead, this summary should be considered a menu of options. Resources for measures are also available at www.pbis.org and [PBIS Assessment](#).

Measures for Reach Evaluation

Assessing reach is one of the most straightforward activities. These are counts of schools, often also disaggregated by groups of schools (e.g., grade levels served, school types). The definition of involvement may vary (e.g., having been trained, reporting implementation, active leadership team in place), but in general, it is useful to report a single metric here, as other counts (e.g., schools implementing with fidelity) will come later in the plan. In larger organizations, it is useful to use a spreadsheet or database to identify schools, points of contact, and professional development access (e.g., Tier 1 training dates, Tier 2 training dates) for the purpose of identifying cohorts.

- **Counts of schools.** This is simply the number of schools involved in the initiative for the evaluation period, often disaggregated by grade levels served (e.g., elementary, middle, high), school type (e.g., regular schools, special schools), or implementation focus (e.g., Tier 1, Tier 1 and 2, Tier 1 through 3). State evaluations may also report counts of districts or regions with implementing schools.
- **Counts of students.** Because PBIS focuses on supporting all students, assessing students reached involves obtaining and summing the total student enrollment across all participating schools. This metric is useful to provide a more humanized and

Remember: Utilize Existing Data When Feasible

The desired balance is to gather enough information to adequately inform the targeted questions but not so much information that the evaluation effort is unfeasible.

student-focused count of the reach of the PBIS initiative. More advanced evaluation plans may also include counts of students accessing Tier 2 or 3 supports.

- **Percent of schools.** It is also useful to calculate and report the percent of schools in the organization that are involved.

Measures for Process Evaluation

When building an evaluation plan, it is useful to create a formal system for documenting the training, professional development, technical assistance, technical support, materials and effort that are invested toward a targeted goal (e.g., PBIS implementation).

- **Action plan completion.** Action plan completion data indicate what accomplishments the leadership team has achieved. These data can be obtained from meeting minutes or commercial project management software (e.g., Asana, Smartsheet). They can be summarized as number of action plan steps completed during the evaluation period (or percent of scheduled action plan items completed). It is often helpful to report these descriptively as well (e.g., list of key accomplishments and a brief description of its importance).
- **Professional development offerings.** These measures document the frequency, duration, content, and participation in professional



development trainings, coaching, and other technical assistance. These data can be obtained from professional development calendars, event registration data, and coaching logs.

- **Professional development fidelity.** An often overlooked aspect of process evaluation is evaluating the extent to which the professional development was delivered as intended. Fidelity can be assessed through a checklist of critical features of effective PD delivery, completed as self-assessment or as an observation by a peer or supervisor (see Appendix A for an example). Due to their cost, observations may be completed on a sample (e.g., 10%) of activities. These data are reported as a percent of critical features delivered, and also descriptively through strengths and suggestions for improvement.
- **Professional development recipient perceptions.** Participant feedback can also serve as an indirect measure of PD fidelity. These data are most often collected via a feedback survey at the end of the event. Key questions include overall satisfaction and the extent to which the content, format, and delivery was useful, relevant, and effective, as well as the extent to which recipients report acquiring new knowledge and skills (see Appendix B for an example). Data are often summarized with a percent of respondents reporting high overall satisfaction, and also descriptively through strengths and suggestions for improvement. General and specific markers of satisfaction help to improve future professional development opportunities.
- **Stakeholder Input and Satisfaction Survey (SISS; fieldtest.pbisassessment.org).** An example of a tool used to obtain detailed descriptive feedback in each school is the SISS suite, which includes separate surveys for from school personnel, students, and families. Each survey includes a

mix of yes/no and open-ended questions to learn about how they perceive PBIS and get anonymous suggestions about how it can be improved. It is delivered online via an anonymous link that can be distributed or completed in a computer lab. Automated reports generate the percent satisfaction (from the yes/no questions) and a list of comments that can be summarized for action planning.

Measures for Capacity Evaluation

Implementation of PBIS includes not just establishing the practices, systems, and data needed in individual schools, but building the capacity of a district, region, or state to sustain and scale the full PBIS framework. As with schools measuring fidelity of implementation, there are multiple measures for districts and states to measure organizational capacity. These include the District Systems Fidelity Inventory (DSFI), the District Capacity Assessment (DCA), the State Systems Fidelity Inventory (SSFI), and the State Capacity Assessment (SCA).

- a. **District Systems Fidelity Inventory (DSFI; fieldtest.pbisassessment.org).** The DSFI is currently being piloted for assessing district capacity to implement, scale, and sustain PBIS, based on the logic of the PBIS Implementation Blueprint. It is a 56-item assessment completed annually with the district team and an external technical assistance provider, and each item is rated according to a rubric as in place, partially in place, or not in place. The DSFI has 9 subscales: Leadership Teaming, Stakeholder Engagement, Resource Alignment, Funding, and Allocation, Policy and Systems Support, Workforce Capacity, Professional Development, Coaching and Technical Assistance, Evaluation, and Local Implementation Demonstrations. The DSFI produces a total score based on the percentage of total points allocated, subscale scores based on the



proportion of points for the items in each subscale and a report with individual item scores. The primary purposes of the DSFI are to (a) indicate the current capacity of a district to implement, scale, and sustain PBIS, and (b) identify specific areas of improvement (e.g., items for action planning) that would increase district capacity. Documentation of the technical adequacy of the DSFI is currently in progress.

- b. District Capacity Assessment (DCA; [sisep.org](http://www.sisep.org)).** The DCA was published by the State Implementation and Scaling Up of Evidence-based Practices (SISEP) Center (www.sisep.org). The DCA is intervention-neutral, meaning that it can be used to assess district capacity for implementing the PBIS framework or any other intervention using the same tool and scoring. The DCA is completed by a district leadership team with an external coach and can be completed in 60 to 90 min. The DCA has 26 items scored as In Place, Partially in Place, or Not in Place across nine sub-scales: Leadership, Action Planning, Performance Feedback, Selection of Staff, Training of Personnel, Coaching, Decision-Support System, Facilitative Administration, and Systems Intervention. The instrument generates a total score, sub-scale scores, item-scores, and an action plan for use by the district leadership team to improve district capacity. Documentation of the technical adequacy of the DCA v.7.0 is available through Ward et al., 2019, available on www.sisep.org.
- c. Regional Capacity Assessment (RCA; [sisep.org](http://www.sisep.org)).** The RCA was developed by the SISEP Center to assist regional education agencies (e.g., area educational agencies, BOCES) in their efforts to support districts in their use of practices supported by evidence or effective innovations, including PBIS. Regional implementation teams need to know how well their system of support can help local districts and schools implement PBIS and determine if the system is sustainable. The RCA lists the core

features of a sustainable system of support. The RCA Technical Manual provides background information on the technical adequacy (i.e., validity and reliability) of the instrument.

- d. State Systems Fidelity Inventory (SSFI; fieldtest.pbisassessment.org).** The SSFI is based on content in the PBIS Implementation Blueprint. It is completed by a PBIS state leadership team, which may or may not be situated within the State Educational Agency (SEA). The SSFI is a 45-item measure with nine sub-scales: Leadership Teaming, Stakeholder Engagement, Funding and Alignment, Policy, Workforce Capacity, Training, Coaching, Evaluation, and Local Implementation Demonstrations. The items are scored as in place, partially in place, or not in place and generate a total score, subscale scores, and an items report.
- e. State Capacity Assessment (SCA; www.sisep.org).** Like the DCA and RCA, the SCA was published by the SISEP Center (www.sisep.org) as an intervention-neutral capacity assessment tool. The SCA is completed by a state leadership team. It is available on www.sisep.org.

Measures for Fidelity Evaluation

A central question for PBIS evaluations is if the initiative has resulted in individual schools implementing the core features of PBIS. There are numerous fidelity assessment measures, but the more recently released and recommended tool for assessing the core features of Tier 1, 2, and 3 PBIS is the *PBIS Tiered Fidelity Inventory (TFI)*. Other fidelity measures (described in Appendix C) include the *Benchmarks of Quality (BoQ)*, *Early Childhood Benchmarks of Quality (EC-BoQ)*, *Facility-wide Tiered Fidelity Inventory (FW-TFI)*, *Individual Student Systems Evaluation Tool (ISSET)*,



Interconnected Systems Framework—Implementation Inventory (ISF-II), *School-wide Evaluation Tool (SET)*, *Self-Assessment Survey (SAS)*, and *Team Implementation Checklist (TIC)*. All of these measures are available at no cost through both www.pbis.org and [PBIS Assessment](#).

[PBIS Assessment](#) allows users to download the measures, enter scores, and generate graphic reports for total scores, subscale scores, and items for monitoring over time.

a. PBIS Tiered Fidelity Inventory (TFI; PBIS Assessment). The [TFI](#) (Algozzine et al., 2014) is completed by a school PBIS implementation team with an external PBIS coach or coordinator, and is divided into three scales (Tier 1, 2, and 3), which can be completed together (i.e., all three tiers) or independently (i.e., one or two tiers). The Tier 1 scale is designed to assess how well the core features of Tier 1 PBIS are in place in a school. There are 15 items in the Tier 1 section, with each item tied to a core feature of PBIS (note that completion of the Tier 1 section of the TFI includes a walkthrough, comprised of brief interviews with faculty/staff and students). The Tier 2 section of the TFI has 13 items and provides an index of the core features of Tier 2 PBIS systems. The Tier 3 section of the TFI has 17 items and provides an index of the individualized social-emotional-behavior supports within a school (note that three randomly selected individualized support plans are reviewed for documentation of the assessment and plan development features). Schools are assessed as “meeting” basic criteria for Tier 1 PBIS implementation when they score at least 70% on the Tier 1 scale score (there are no established criterion scores for Tiers 2 or 3). Item reports for each tier are used for action planning and ongoing improvement efforts. It is important for the team to focus less on achieving a particular score and

more on identifying items for celebration and improvement. The TFI has been validated through numerous studies (Massar, McIntosh, & Mercer, 2019; McIntosh et al., 2017; Mercer, McIntosh, & Hoselton, 2017).

The recommended schedule of TFI administration depends on the school’s implementation status. As one example, a school that has chosen to implement PBIS, formed a leadership team, and committed to process would use the TFI three times per year.

- **Fall:** In October or November, a leadership team (with a district PBIS coach) can complete a full TFI, in which all three tiers are examined. The school team produces a “Total Score per Tier” and item reports that indicate specific areas where targeted improvement would be most beneficial. Based on the results, the leadership team selects one focus tier for targeted improvement for the next three months, and specific actions that will be performed. The full TFI assessment typically requires 60 to 90 minutes.
- **Winter:** In January or February, the team conducts a second TFI evaluation, but only for one tier. If the team is focusing on Tier 1 implementation, then only the items from Tier 1 are reviewed. If the Fall assessment indicated that Tier 1 was implemented at criterion and Tier 2 was the selected focus, then only the Tier 2 items would be reviewed. The winter assessment typically requires 15 to 30 minutes. The purposes of the winter assessment are to confirm the focus tier for improvement and select specific items for change in the next three months.
- **Spring:** In April or May, the team conducts their third TFI review, but again they assess only the focus tier.



A school that has been using PBIS for several years and achieved at least 70% Total Score implementation on all three tiers may choose to conduct only the Fall Assessment.

b. Model school/district/region identification.

Because having a set of local implementation demonstration sites is an effective strategy for sustainability (McIntosh, Kelm, & Canizal Delabra, 2016), evaluation plans often include the identification of model schools or districts through a set of criteria (e.g., documented fidelity of implementation, high-quality products, and positive outcomes). It is useful to identify a range of model sites, including across levels (e.g., elementary, middle, high), tiers (i.e., 1, 2, 3), and community contexts (e.g., urban, rural). Many organizations have adopted formal recognition systems as a strategy to increase submission of data for evaluation and allow for identification of potential model sites.

Measures for Outcomes Evaluation

Evaluations of PBIS typically include multiple measures of the impact of PBIS on student success in a variety of domains. In addition, evaluations may also include measures of the effect of PBIS adoption on other organizational features of schools. Some evaluations may be restricted to specific measures most related to the impetus for implementing PBIS (e.g., reductions in exclusionary discipline). However, we recommend broadening the scope of evaluation to other valued outcomes.

To obtain the highest quality data, it is preferable to select (a) direct observation measures over perception measures and (b) measures of positive behavior (i.e., what you want to see) over measures of unwanted behavior (i.e., what we want to reduce). In practice,

A Note About Recognition Systems

Recognition systems can have unintended negative side effects. Remember their key purposes: a) to identify model sites, b) to reinforce the data collection efforts of individuals from schools and districts, and c) to increase the submission of quality data to the larger organization for evaluation efforts. Be careful that being recognized could become so valued that individuals will feel encouraged (explicitly or implicitly) to inflate their fidelity scores. However, this risk can be mitigated by encouraging submission of accurate scores (not necessarily high scores). In addition, recognition systems can be used to shape school behavior. For example, Florida's PBIS Project requires that schools show progress in reducing disciplinary inequities to achieve recognition, and the Georgia DOE requires that schools have zero incidents of corporal punishment to receive their highest status.

these two considerations are often difficult to combine with our earlier recommendations of using existing measures and focusing on efficiency. The practical choice often is a compromise that fits the resources, interests, and abilities of the evaluation team and stakeholders. The following are commonly-used measures for evaluating outcomes (Appendix C includes some less-common options).

- a. Office discipline referrals (ODRs).** Most schools track the type and frequency of student unwanted behaviors that result in a referral to school administrators (the office). This practice creates a reliable metric of levels of problem behavior in the school if (a) school personnel use operationally defined definitions of problem behavior, (b) have a common standard for what is and is not referred



to the office (versus managed locally), and (c) have a reliable system for recording and reporting data (Irvin, Tobin, Sprague, Sugai, & Vincent, 2004; McIntosh, Campbell, Carter, & Zumbo, 2009). We recommend using ODRs per school day as a primary metric for an individual school, with ODRs per 100 students per school day as the metric of greatest utility for comparison across schools.

In addition to basic calculations of annual levels, ODRs can be extremely beneficial for decision making when school and district personnel have access to reports that document ODR rates disaggregated by (a) type of problem behavior, (b) location of incident, (c) time of day, (d) student, and (e) behavior motivation (e.g., to obtain or avoid attention, objects, access to activities). The School-wide Information System (SWIS; www.pbisapps.org/swis) is a web application for data entry, summary, and reporting developed at the University of Oregon that has proven useful in managing ODRs and is offered at operating cost to schools. The SWIS Suite includes an application for managing data at Tier 1, another application (CICO-SWIS) for Tier 2, and a third application (I-SWIS) for Tier 3.

- b. Suspension/expulsion/arrest and use of restraint or seclusion.** In addition to ODRs, there are more serious disciplinary procedures that can be used to assess effects of PBIS on exclusionary discipline, including in-school and out-of-school suspensions, expulsions, school (i.e., on-campus) arrests, and incidents of restraint or seclusion. Most schools are legally required to maintain records of these incidents. To make these data most helpful for decision making, it is useful to report these data in at least three forms (using suspension as an example): (a) the number of students receiving at least one suspension, (b) the number of suspensions, and (c) the number of school days lost

due to suspension.

- c. Discipline disproportionality.** Equity in educational outcomes can be assessed for any student measure, but the most common outcome to monitor is exclusionary discipline (e.g., ODRs, suspensions). Although there are a number of ways to measure discipline disproportionality (McIntosh, Barnes, Morris, & Eliason, 2014), we recommend calculating and reporting the Risk Index (e.g., number of Black students receiving ODRs / total number of Black students), Risk Ratio (e.g., the Risk Index for Black students / the Risk Index for all other groups of students, and Rates per group (e.g., ODRs issued to Black students / total number of Black students). Risk Ratios are the most commonly reported and easily explained, but Rates per subgroup are most stable for progress monitoring and provide an index of both levels and inequities in exclusions (Girvan, McIntosh, & Smolkowski, 2019). See the [PBIS Discipline Disproportionality Data Guidebook](#) (McIntosh et al., 2014) for more details.
- d. Academic achievement.** Research shows that schools implementing Tier 1 PBIS with fidelity show improved academic achievement, although these effects are typically seen after three or more years of implementation due to the indirect effects (Kim, McIntosh, Mercer, & Nese, 2018; Lee & Gage, 2020). Given these findings, it may be useful for multi-year evaluations to include measures of student academic outcomes. Common measures include state high-stakes achievement measures or curriculum-based measures, often calculated as the percent of students meeting or exceeding expectations.
- e. Attendance and chronic absenteeism.** Because PBIS is intended to make schools more safe and positive environments for students, it is common for PBIS evaluations to focus on attendance. However, the operational definition for attendance may vary across

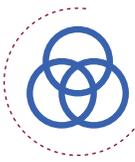


schools and districts (e.g., present for all classes, present for at least one class), so it is important to document how it is calculated. Common metrics include average daily attendance and chronic absenteeism, both reported as percent. Many schools already have high average daily attendance (e.g., at or above 95%), so it may be more beneficial to focus on reducing chronic absenteeism.

- f. Social-emotional-behavioral competence.** Student social-emotional-behavioral competence is typically assessed with rating scales or [screening tools](#) completed by school personnel (e.g., teacher-completed universal screenings), and sometimes by students as a self-report. Most rating scales are commercially published surveys of social skills or social-emotional competencies.
- g. School Climate Survey (SCS; [PBIS Assessment](#)).** The [School Climate Survey](#) suite provides an important perception measure for evaluating less observable outcomes of PBIS. The suite includes surveys for three different stakeholder groups (students, family members, and school personnel). Each survey is designed to be completed anonymously (either online or hard copy, in English or Spanish) and includes a brief set of demographic questions and items assessing the school environment. The Elementary version has 9 items, the Middle/High school version has 11 items

(a long form includes 36 items), the Family version has 21 items, and the School Personnel version has 29 items. Reports (on [PBIS Assessment](#)) produce mean ratings of school climate (higher scores are more positive), and scores can be disaggregated by demographic groups, providing an easy way to assess equity in perceived school climate. The SCS is typically completed once per year but has been used twice per year for progress monitoring or pre-post evaluation of PBIS implementation. If the team wishes to assess a sample of students instead of all of them,

- h. School Safety Survey (SSS; [PBIS Assessment](#)).** The School Safety Survey was developed by Sprague, Colvin, and Irvin (2002) and is designed to assess the overall perception of school safety by school personnel. The SSS items include 17 Risk Factors and 33 Protective Factors. SSS results are used to identify both perceived school safety and possible steps to reduce risk factors and increase protective factors of the school environment.
- i. Faculty/staff turnover.** An under-examined measure for PBIS evaluation is faculty and staff mobility. In evaluations that span multiple years it has been helpful to establish the rate at which teachers, building administrators, and other staff leave a school (or district).



Conducting Evaluations

Once an evaluation has been planned with sufficient detail, the process of collecting, organizing, storing, and summarizing the evaluation data occurs. The evaluation plan indicates what questions will be asked, what data will be collected, and when and how it will be collected.

A key way to ensure effective and efficient evaluation is to follow the prescribed steps in the evaluation plan and ensure that is completed as designed. Conducting an evaluation (or a significantly revised evaluation) for the first time often takes more time than anticipated, due to negotiating new relationships and processes. It is important to document the steps taken, availability of and access to



data, and any lessons learned in following the plan to ensure that the next cycle of evaluation becomes more efficient and routinized. Keeping the evaluation questions in mind can help maintain focus for evaluators.

Often in the course of conducting evaluations, new information will emerge that necessitates a renegotiation of the original plan.

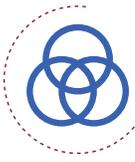
For example, data that seemed readily available may be either difficult to access or of insufficient quality to answer questions. In other situations, new data sources may emerge that will help answer existing or additional valued questions. In all cases, it is important for the evaluation team to communicate proactively with stakeholders to negotiate a revised plan.

Guiding Questions for Ensuring Evaluations are Conducted as Planned

- Were planned evaluation activities completed within the identified timeline?
- Were planned evaluation activities completed within the budget?
- Were resources allocated to conduct the evaluation?
- Were all required data for completing the evaluation obtained?
- Were all evaluation questions addressed and answered?
- Did the evaluation questions need to be revised?

Assess the Quality of Data for Evaluation

Evaluation plans may need to change if data are incomplete or insufficient for answering the questions. The problem of incomplete data can be prevented through providing schedules of data collection to sites and monitoring rates of data submission periodically instead of waiting until the evaluation period. Alternatively, the data from existing sources may be of such questionable accuracy that the questions cannot be answered. In these situations, it may be necessary to collect new data with a subset of schools to be able to answer the evaluation questions.



Finally, it is important for evaluators to conduct the process with care. Stakeholders may be uneasy with the concept of evaluation, and this unease could lead to resistance or harmed relationships. Evaluators must always be able to identify system strengths, describe the purpose of evaluation, and stress the value of conducting evaluations for improving valued outcomes (e.g., increased efficiency or effectiveness of services delivered to students).



Reporting Evaluation Results

After the evaluation is conducted, the real work begins. Each evaluation report will be shaped by the purpose, questions, and needs of the stakeholders. Care is needed to report the data with clarity, frame the data within the larger issues of concern, and ensure that readers appreciate the practical and technical limitations associated with educational evaluations.

There are many approaches to reporting evaluation results successfully, and each evaluation report will be tailored to the purposes of the evaluation and the needs of the stakeholders. Examples of state evaluation reports using different measures and adapting to stakeholder evaluation needs are provided for PBIS Missouri (Appendix E1), California PBIS Coalition (Appendix E2), Wisconsin PBIS (Appendix E3), and Florida PBIS (Appendix E4).



Regardless of the specifics, PBIS evaluation reports follow a logical outline, such as described in the following sections:

- (1) **Introduction**,
- (2) **Process**,
- (3) **Results**, and
- (4) **Summary and Recommendations**.

1. Introduction

An effective introduction prepares the reader for the report by briefly describing the context for evaluation.

It will describe the purpose of the evaluation, the goals and objectives of the PBIS initiative, a brief definition of the PBIS framework for unfamiliar readers, the key stakeholders, and the evaluation questions. Generally, it is better to be brief and straightforward so the reader sees the key information instead of losing focus reading text that is useful but not critical to understanding the report.

One example is the PBIS Missouri Annual Report.

CONTEXT

Evaluation of the context details the goals, objectives, and activities of the program. Context serves as a foundation for identifying required resources, assessing expected and actual implementation, and analyzing expected and actual outcomes and evidence of performance (Algozzine et al., 2010, p. 3). The answers to the following questions show evidence of our action plan and the people who provided and received support through MO SW-PBS for 2018–2019. This annual report is guided by the framework laid out in the "Evaluation Blueprint for School-Wide Positive Behavior Support." Questions and categories are taken directly from this document.

Question 1

What are/were the goals and objectives for MO SW-PBS implementation?

The 2015–2018 MO SW-PBS three-year Action Plan includes six primary goals that are reviewed annually. The goals and supporting objectives are revised and updated as data indicate appropriate. Each goal is addressed in more detail within this report. They are:

1. MO SW-PBS is the social behavioral, three-tiered intervention utilized in the Statewide System of Support (SSOS).
2. MO SW-PBS Consultants sustain capacity to provide professional development and technical assistance for social behavioral interventions across three tiers of support.

3. MO SW-PBS provides standardized professional development and technical assistance for school and/or district teams and administrators.
4. MO SW-PBS utilizes a systematic process for data collection and evaluation of implementation fidelity and student outcomes.
5. MO SW-PBS implements systems for replication, sustainability, and improvement.
6. MO SW-PBS utilizes a communication plan targeted to stakeholders.

MO DESE Goals: 1, 2, 3, 4



The California PBIS Coalition Evaluation Report provides a formal listing of their evaluation questions to structure the report:

- Who will provide support for PBIS implementation?
- Who will receive support during PBIS implementation?
- What professional development was part of CA PBIS implementation support?
- Who participated in the professional development?
- Are California schools implementing PBIS with fidelity?
- To what extent does implementation of PBIS result in safer schools?
- Is PBIS resulting in positive outcomes for CA students?
- To what extent is PBIS implemented throughout the state and once implemented with fidelity, to what extent do schools sustain PBIS implementation?
- To what extent does CA PBIS implementation support changes in educational/discipline policy?

2. Process

Like the introduction, this section can be brief and to the point. The content of this section is intended for readers who want to know precise details about what data were collected and gain a sense of the reliability and validity of the measures used and the general evaluation approach. Such information helps build trustworthiness and replicability of the evaluation process to give confidence in its results. Subheadings can include (a) Measures (i.e., what tools were used to collect data) and (b) Procedures (i.e., how the data were collected and converted into information).

Add Vignettes to Tell the Story of the Initiative

Effective evaluation reports not only describe results with precision but also tell the story behind the data. One approach for balancing the need for documenting large-scale evaluation precision with understanding practical outcomes for students and educators has been to intersperse quantitative results with vignettes of one school, or one sub-topic, to personalize the larger evaluation content.

3. Results

Results are typically reported by evaluation question or evaluation question area.

REACH: "WHO IS PARTICIPATING IN PBIS?"

Reach questions are intended to identify who is participating in PBIS and are affected by it. Primarily, reach indicates the schools (and students enrolled in those schools) that are part of the PBIS initiative, often updated yearly from previous evaluations. Answers are important to determining the scale and potential benefit of the initiative, as well as areas for expansion and sustained implementation. Results could include raw counts or percentages of schools. Evaluations often also identify other agencies or organizations that are participating (e.g., community agencies, other departments or governmental organizations).

Presenting reach information down to the student level (number of students receiving Tier 1 PBIS; i.e., all students in each school) is an especially helpful



way to communicate reach to a range of audiences. Increasingly, these measures are perceived as important policy metrics.

Avoid Data Overload

A common report error is providing too many tables and figures. If readers need to try too hard to make sense of the report, there is less energy available to act on the information. If needed, consider providing extensive tables or charts in an appendix or technical report.

When reporting reach, consider information that is important for communicating the scale, level, and location where implementation has taken place.

This will help stakeholders to see what is connected to their interest. It also provides an opportunity to communicate to leadership what is current status and what might be expanded in the future. Simple counts are useful, but most reports will maximize readability with visuals that capture the scale of the initiative. Some examples follow:

Charts to display increase in reach over time. For example, the California PBIS Coalition Evaluation report described change in the scaling of PBIS across the state over an 18 years period (see below).





Maps to show geographic reach. Sometimes reach is reported through maps with color-coded districts identified or through use of “heat maps” with spots that differ in size depending on the size of

participation. The PBIS Missouri Annual Report uses two maps to show the spread of schools using PBIS by district across the state over time.

Missouri SW-PBS School Counts by Districts

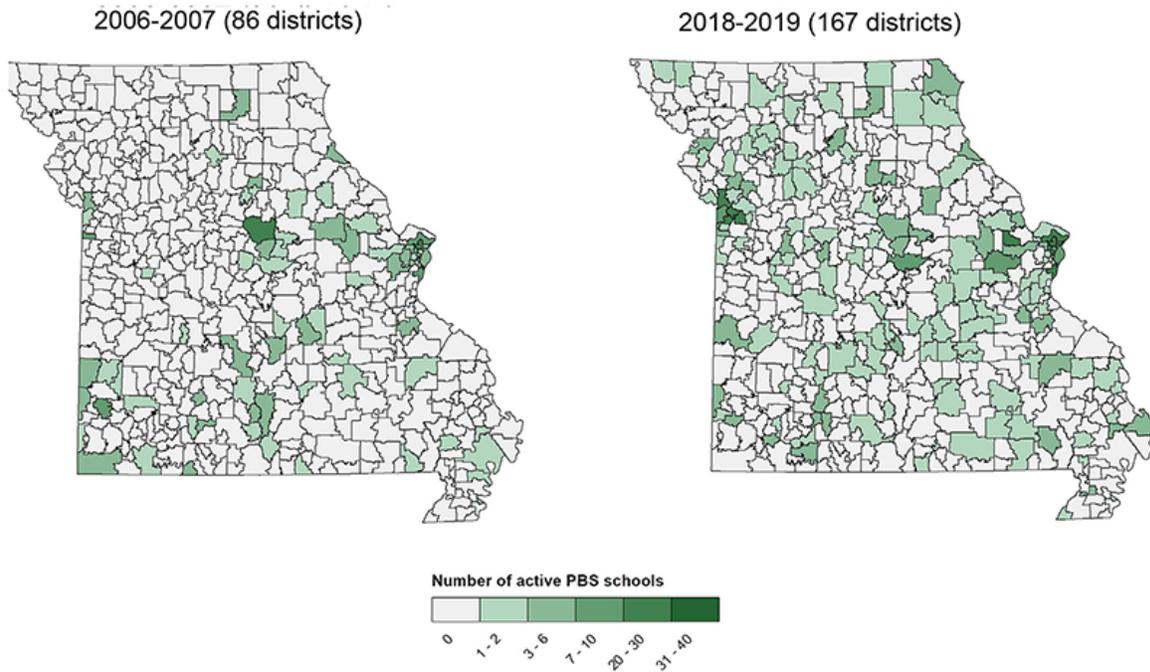
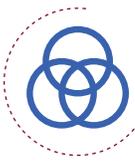


Figure 11b

Over time, MO SW-PBS has expanded to provide training and technical assistance to schools located across the state of Missouri. Additionally, participating schools across all three tiers can be found in every RPDC across the state, providing exemplars for regional schools to visit and opportunities for within region training and networking.



PROCESS: “WHAT IS HAPPENING WITH THE PBIS INITIATIVE?”

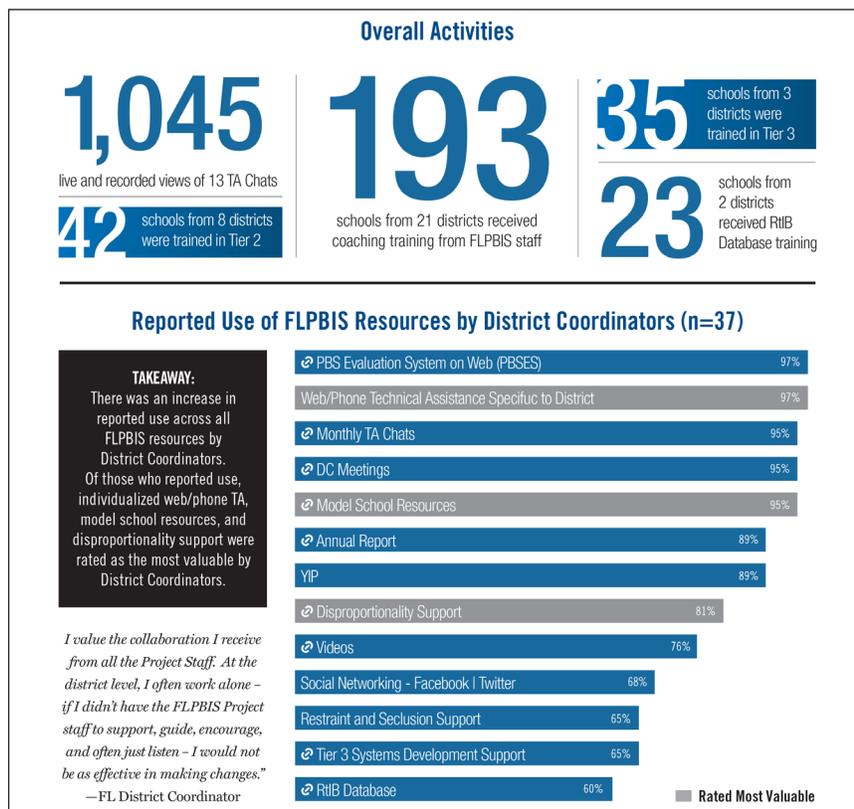
The Process section reports the actions taken to support the PBIS initiative during the evaluation period. Before readers can evaluate whether PBIS implementation was effective, the reader needs to know what was implemented, and how the implementation process occurred. It is here that the evaluation team can indicate how well the implementation process followed the PBIS Implementation Blueprint. Subsections may include (a) Team Accomplishments, (b) Training and Coaching Offered (and uptake), and (c) Participant Evaluations.

Team Accomplishments. Evaluation reports sometimes identify the actions and achievements

of the organization’s PBIS leadership team (or school leadership teams). This section can highlight particularly meaningful milestones that have been achieved and are not apparent through reach, fidelity, or outcomes questions.

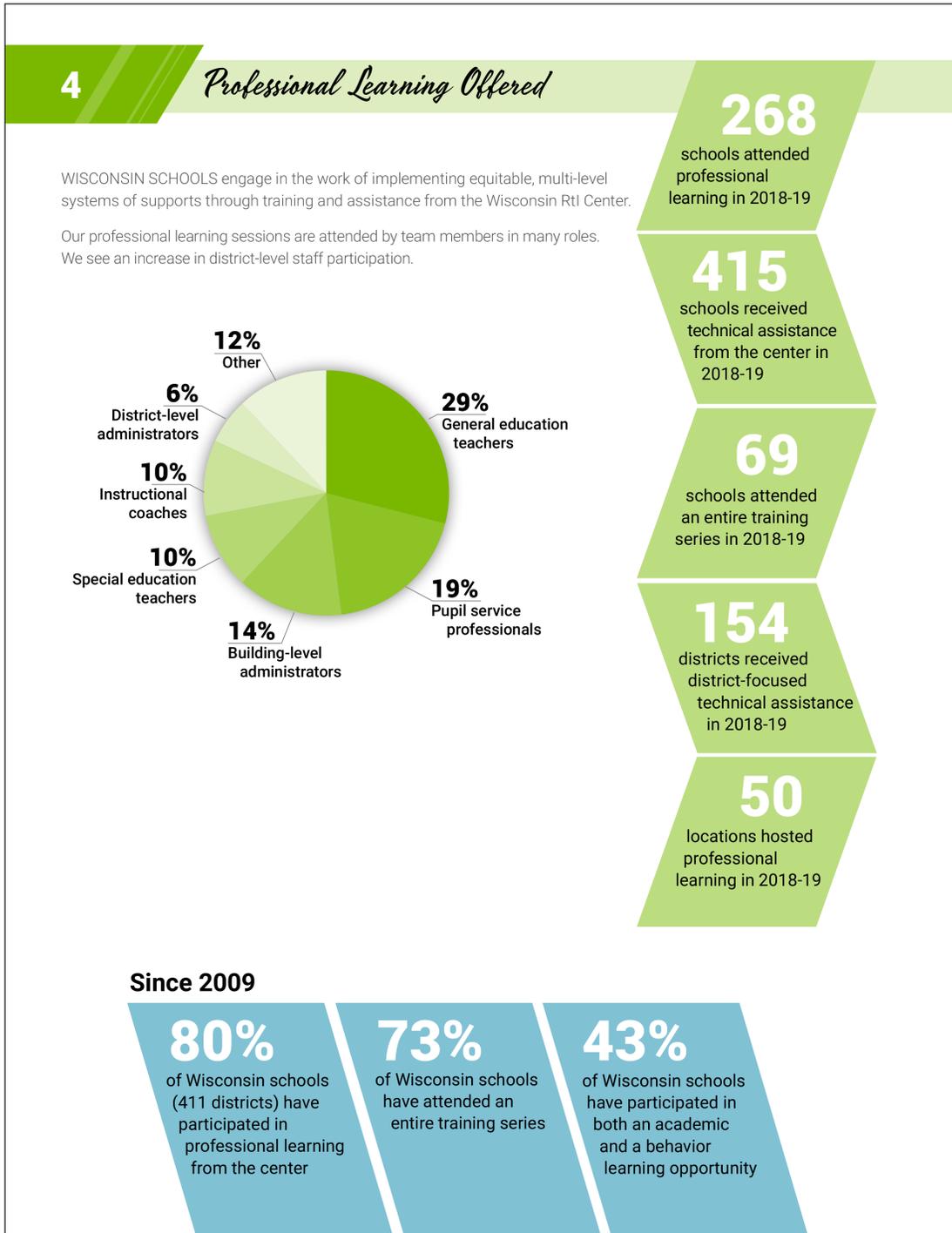
Training and Coaching Offered (and uptake). Reports often document the number and type of professional development opportunities and technical assistance offered to participating schools and districts. Many reports document uptake of offerings, often by number of attendees, schools, or districts that participated.

The FLPBIS Annual Report provides infographic-style depictions of trainings and coaching uptake. It also includes a testimonial regarding the value of the technical assistance.





The Wisconsin Rtl Center’s report shows uptake by role and a cumulative count of schools that have ever received training.





Participant evaluations. In addition to provision and uptake, strong evaluation reports indicate the quality of the training and coaching offered, most likely in terms of participant evaluation forms.

The New Bedford Public Schools Annual Report provides the results of participant evaluations. They provide clear evidence that training recipients found it valuable. The open-ended comments provide a personal touch to complement the quantitative data.

181 INSTRUCTIONAL STAFF PARTICIPANTS | SATISFIED WITH QUALITY OF TRAINING.

97.06%

Strongly Agree/Agree

“This training gave me new ideas on how to help my class in the areas that I was having a hard time in.”

NEW BEDFORD PBIS TEAM MEMBER

“As an administrator, I am planning on using the PBIS resources that were provided to the group to assist teachers in a non-evaluative way on how to improve their classroom management.”

NEW BEDFORD PBIS TEAM MEMBER

144 PARAPROFESSIONAL PARTICIPANTS | SATISFIED WITH QUALITY OF TRAINING.

99.6%

Strongly Agree/Agree

“I really enjoyed this training series. I found the presenters to be very engaging. The content of the training was very clear. The presenters took care to make sure the participants were engaged in the training.”

NEW BEDFORD PARAPROFESSIONAL

“As a Paraprofessional getting all this knowledge on how and when to approach a behavior was and will be of great use in our future.”

NEW BEDFORD PARAPROFESSIONAL

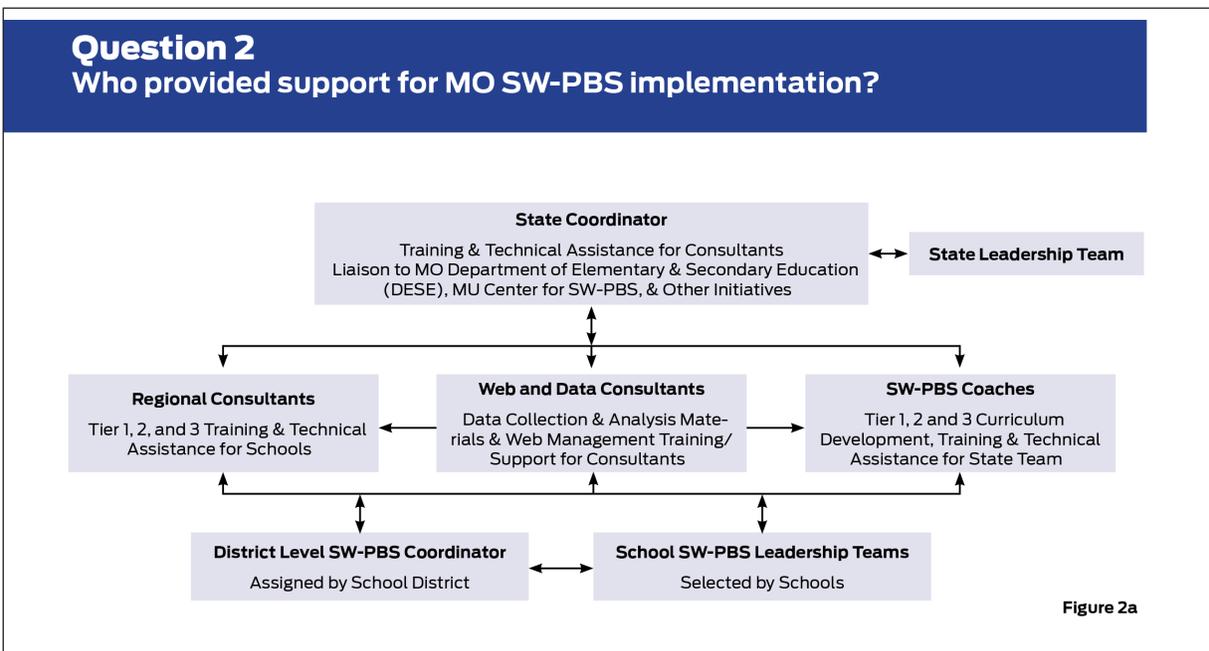


CAPACITY: “WHAT IS THE CAPACITY OF THE ORGANIZATION TO IMPLEMENT AND SUSTAIN PBIS?”

Capacity sections of the report typically document two aspects of the initiative: (a) Structures for Supporting PBIS Implementation and (b) Formal Measurement of Capacity.

Structures for Supporting PBIS Implementation.

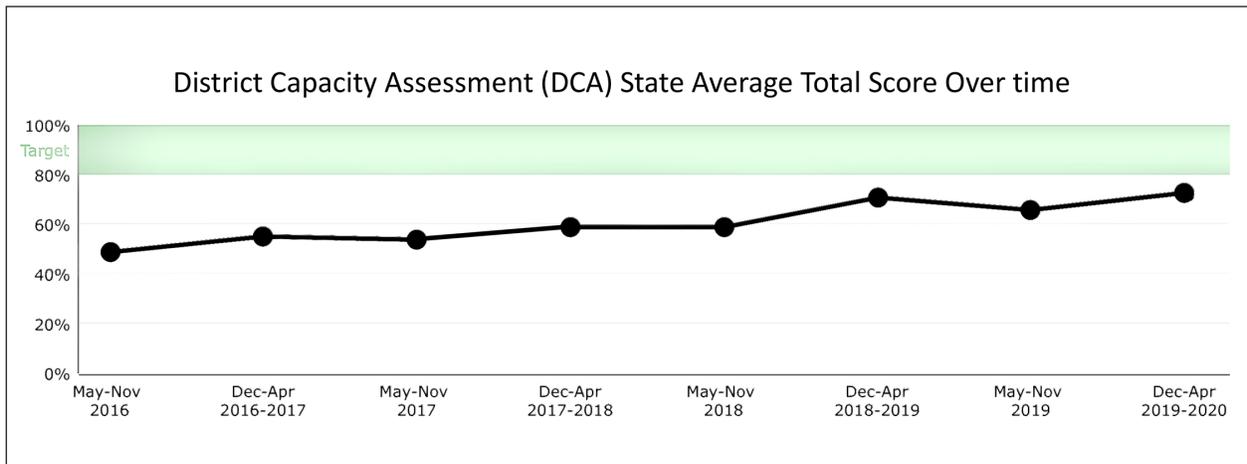
Evaluation reports describe what organizational supports for the PBIS initiative exist. These are displayed often in organizational charts. The PBIS Missouri Annual Report provides one example of displaying the structure of support in the state.





Formal Measurement of Capacity. Increasingly, reports are emphasizing scores from capacity measures like the District (e.g., DSFI, DCA) or the State (e.g., SSFI, SCA). It is helpful to display these data across time to evaluate progress in capacity development.

The following figure is a chart from a state dashboard to displays the average District Capacity Assessment total scores per year for districts participating in the state initiative.





FIDELITY: "ARE THE CORE FEATURES OF PBIS BEING IMPLEMENTED?"

Reporting the fidelity with which schools are implementing PBIS is a cornerstone of PBIS evaluation reports.

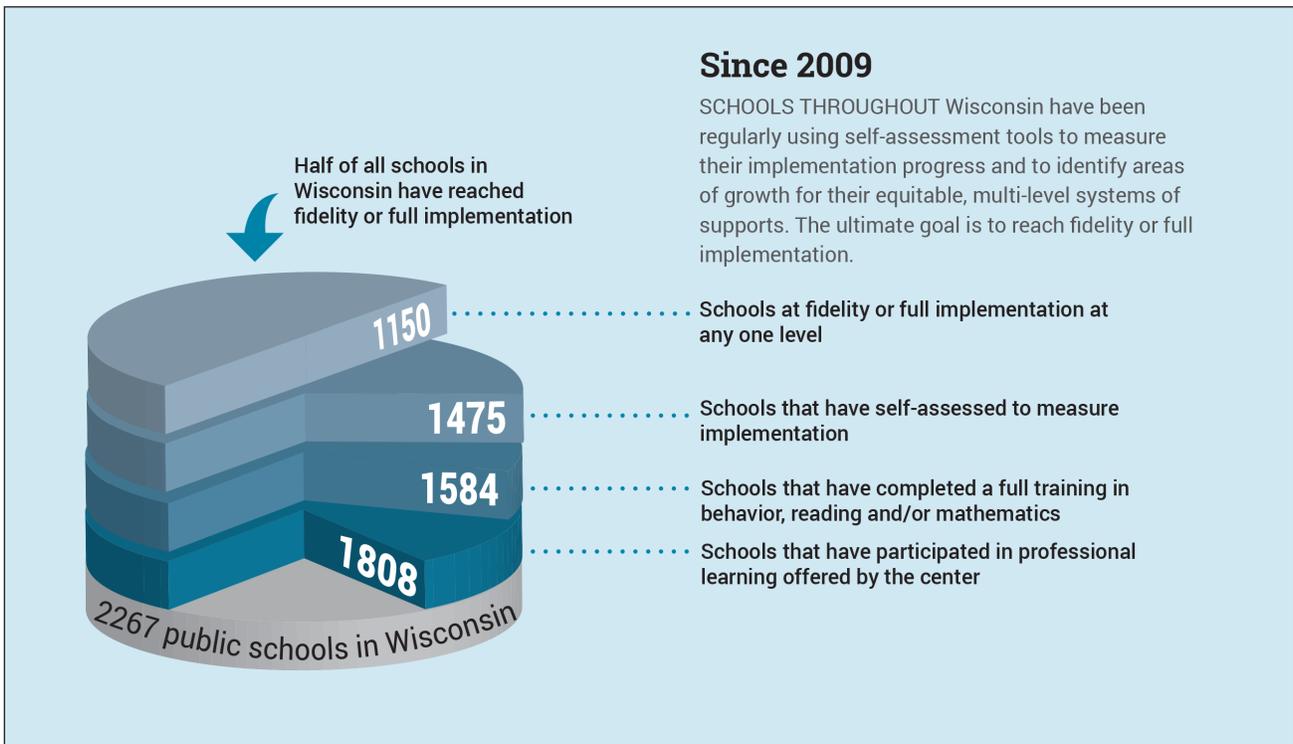
Increasingly, the TFI is being used as the sole measure of fidelity of implementation, but other measures are not uncommon. Fidelity sections often include the following subsections: (a) Schools Measuring Fidelity, (b) Fidelity Scores, and (c) Model Schools.

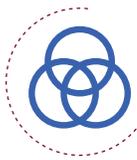
Schools Measuring Fidelity. A common preliminary question is how many (or what percent of) schools

are measuring fidelity of implementation. These are answers through counts or percentages of schools, either (a) in the organization's scope or (b) of schools involved in the initiative.

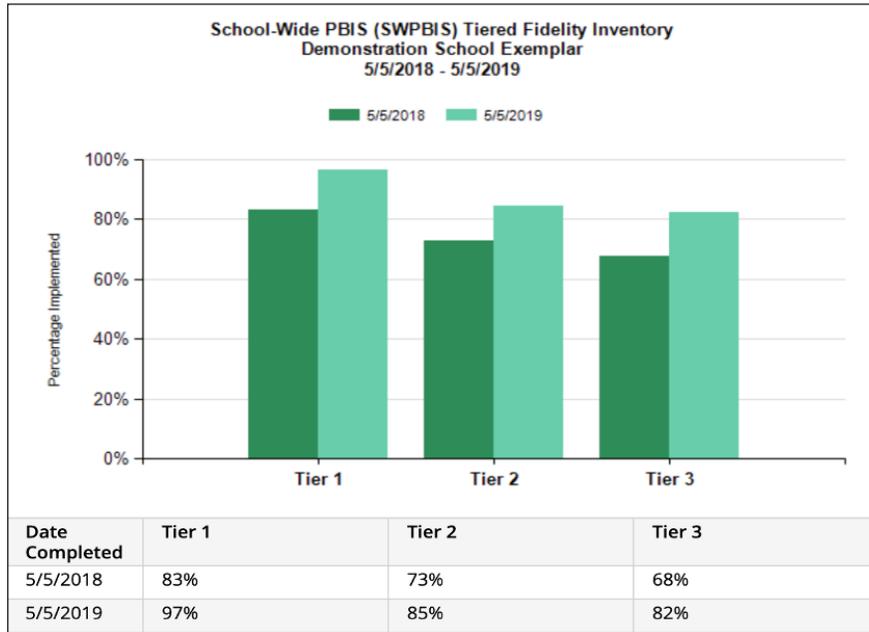
Fidelity Scores. Beyond simply measuring fidelity, it is important for reports to include the actual fidelity scores. These are reported either as (a) average scores or (b) percent of schools with adequate fidelity.

The Wisconsin RtI Center's annual report documents schools trained, schools measuring fidelity, and schools meeting a fidelity criterion.

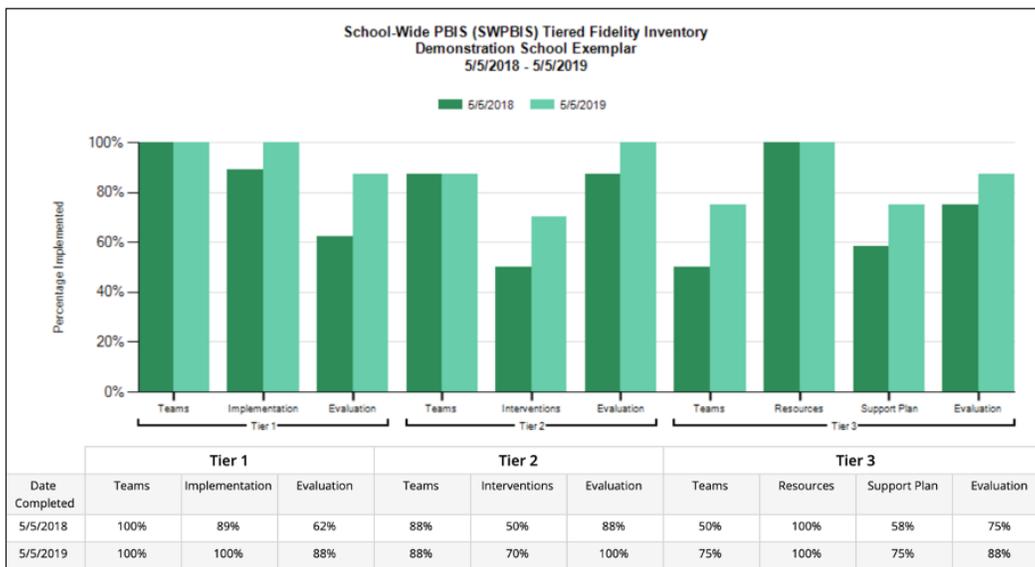




Example of reporting PBIS fidelity for one school. Displaying the scale scores by tier over time allows teams to recognize progress and determine areas for improvement. Adding an indicator of the fidelity criterion (70% for the Tier 1 scale of the TFI) will also facilitate the interpretation of scores.

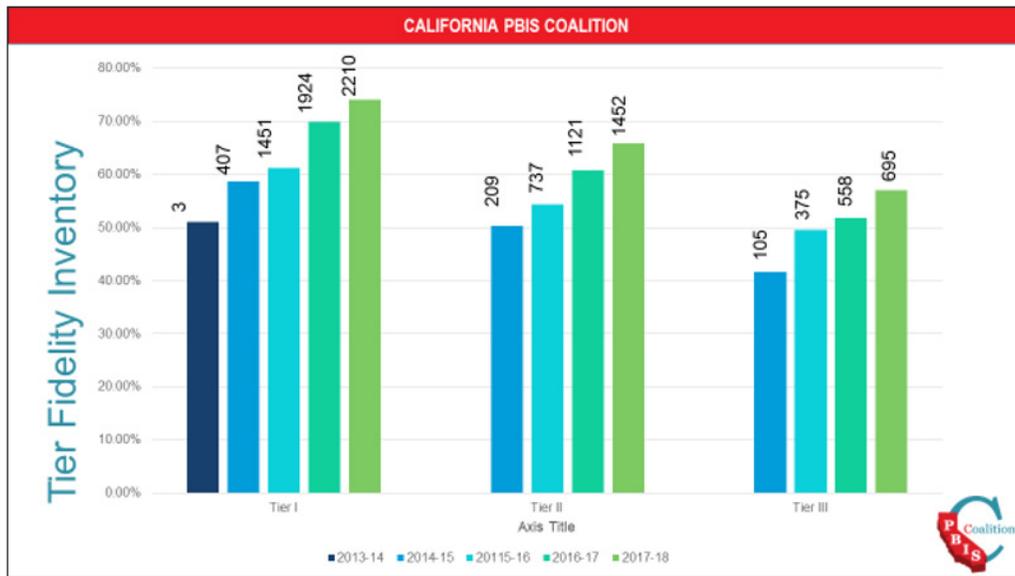


In addition to reporting overall fidelity scores by tier, it is helpful to report scores by subscale. This chart allows district and school teams and other stakeholders to identify aspects of strength and aspects for improvement.



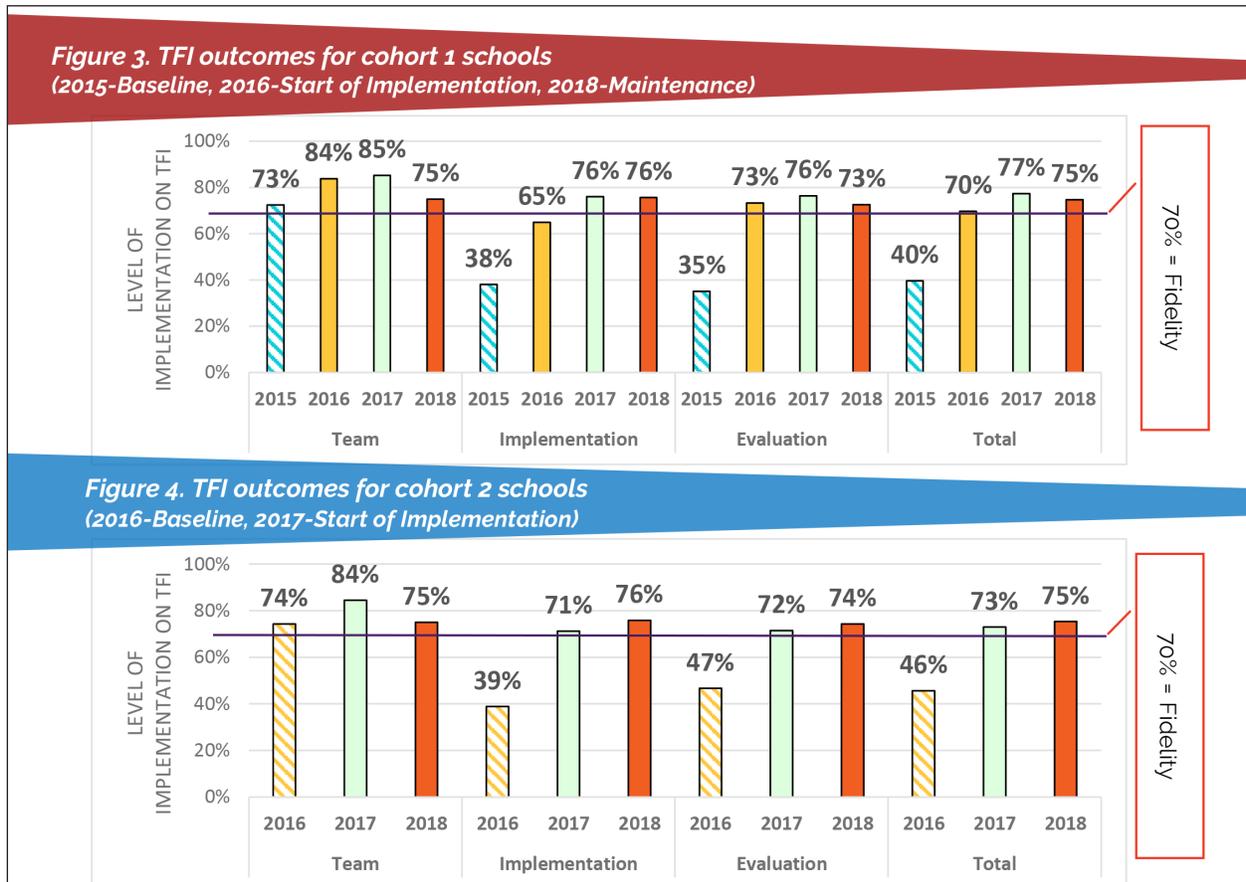


Example of reporting PBIS fidelity across a district or state. Districts or states may also report average fidelity for a collection of schools. The example below is from the California PBIS Coalition evaluation report. They summarize the mean TFI total score per tier for schools reporting fidelity data across five years. With the 70% fidelity criterion for Tier 1 implementation, these data show that in 2017-18, most California schools in the coalition met criterion.





When schools are trained in cohorts, consider also reporting fidelity by cohort. This report from Massachusetts allows the state to review the progress of each cohort and adjust supports as needed.





Model Schools. Fidelity sections often also identify model schools (or sometimes districts). These are schools that have applied for recognition for their implementation efforts. This part of the report is a common place to include a vignette of a single school or district’s implementation story.

The Wisconsin Rtl Center publishes information about its recognized schools in its annual report.

Recognized schools program

SCHOOLS ACROSS the state implementing equitable, multi-level systems of supports have been acknowledged for their efforts through our recognized schools program. Our recognition system commends schools beginning their journeys and congratulates those that are successfully sustaining AND extending them.

This past year, we were pleased to honor 670 unique schools. To view the entire list of recognized schools, visit <http://bit.ly/RecognizedSchools>.

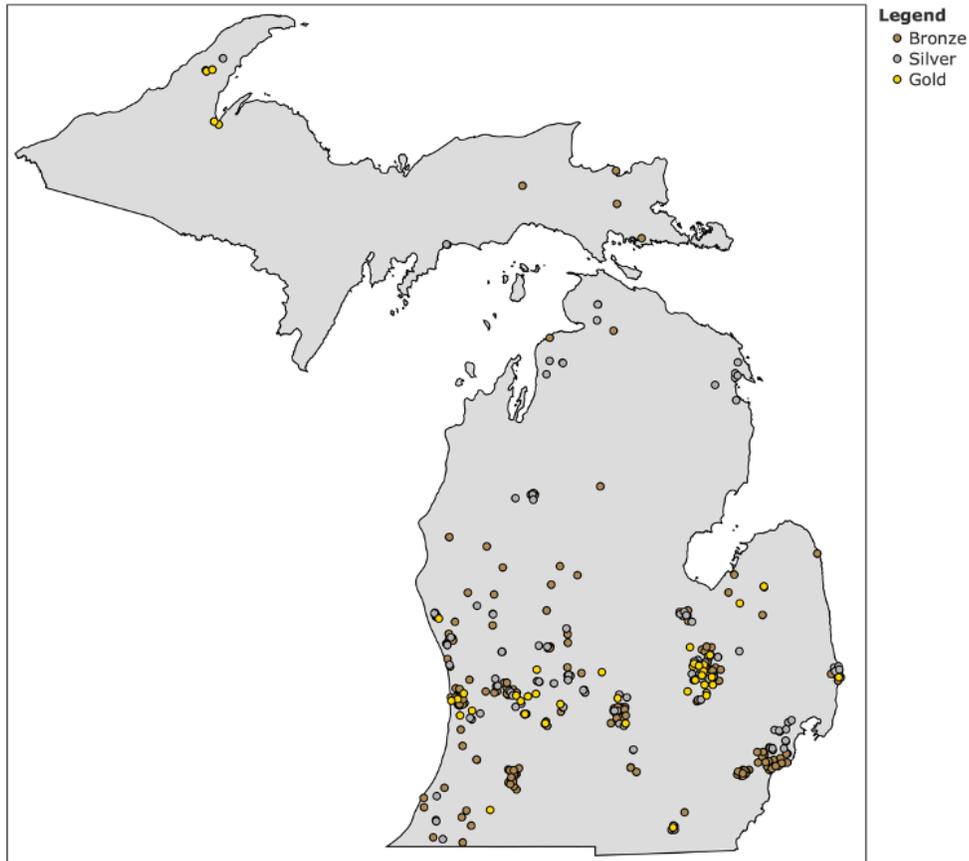


Number of schools at PLATINUM (all content areas)	0
Number of schools at GOLD for behavior and mathematics	1
Number of schools at GOLD for behavior and reading	2
Number of schools at GOLD for mathematics and reading	2
Number of schools at GOLD (all content areas)	6

	Behavior	Mathematics	Reading
Number of schools at SILVER	261	69	82
Number of schools at BRONZE	352	58	69



The following example from Michigan demonstrates a variation by illustrating recognition of schools (e.g., Bronze, Silver, Gold) on a map.



OUTCOMES: “IS THE INITIATIVE ACHIEVING VALUED OUTCOMES AND WORTH SUSTAINING?”

PBIS evaluation reports typically report on the impact that implementing PBIS with fidelity has on student outcomes. Care is needed to avoid causal inferences within evaluation reports, but given the existing well-controlled research documenting the effects of PBIS, evaluators may feel confident in highlighting descriptive results that demonstrate the effects can be realized in the organization’s schools. Typically, results

are compared using one or more of the following methods: (a) comparing schools to themselves (e.g., the same school pre vs. post implementation), (b) comparing schools by fidelity (e.g., high-fidelity schools vs. low-fidelity schools), and (c) comparing schools by involvement in the initiative (e.g., high-fidelity schools vs. all other schools in the state or district). It is important to remember that only schools implementing PBIS with fidelity should be in the PBIS group.



Example of comparing schools to themselves. A useful way to document contextualized change is to assess outcomes by implementation cohort. Such analyses can clarify the outcomes that can be expected as part of early implementation (e.g., first 3 years), sustained implementation (e.g., 4 to 9 years), and long-term

sustained implementation (e.g., 10 or more years). Consider the way the Florida PBIS Evaluation report (below) presented information on change in the office discipline referral rates for 766 schools before and after their first year of implementing PBIS.

Change in Discipline Rates per 100 Students after Their First Year of Implementation

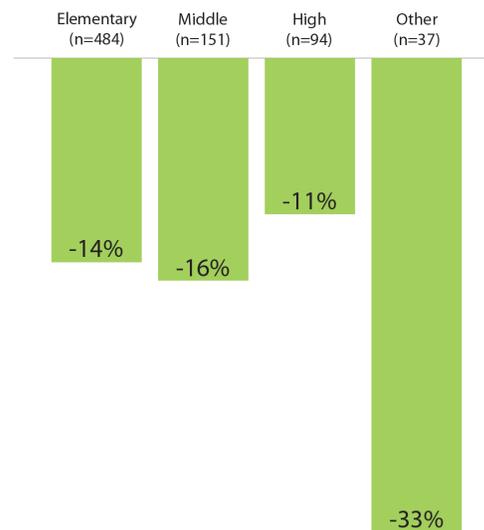
Newly trained PBIS schools can have an impact on student outcomes as early as their first year of implementation.

Note: The data represent schools that reported discipline data both before and after their first year of implementation. They do not represent a particular set of school years. "Other" schools include those with irregular grade ranges such as K-8. Alternative/center schools are not included due to the variable nature of their data.

TAKEAWAY:

Schools see decreases in ODRs, ISSs, and OSSs across most school types from baseline to one-year of implementation.

OFFICE DISCIPLINE REFERRALS (ODRs)





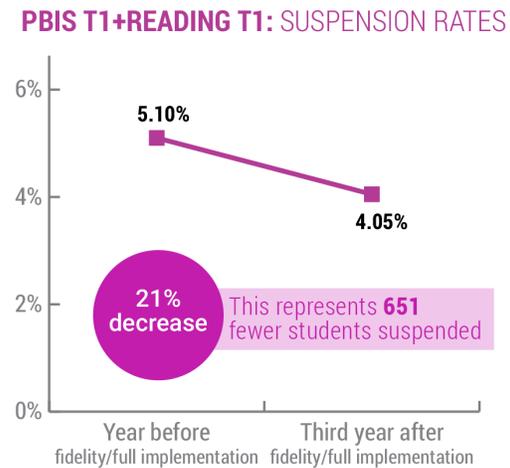
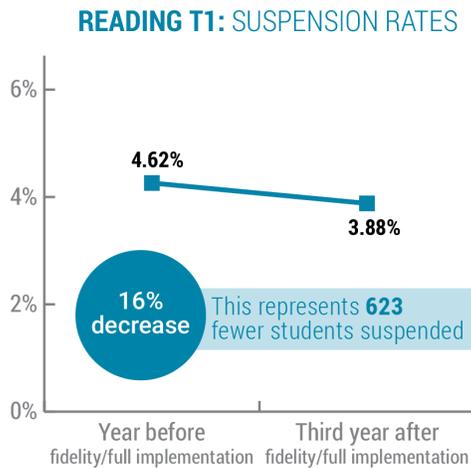
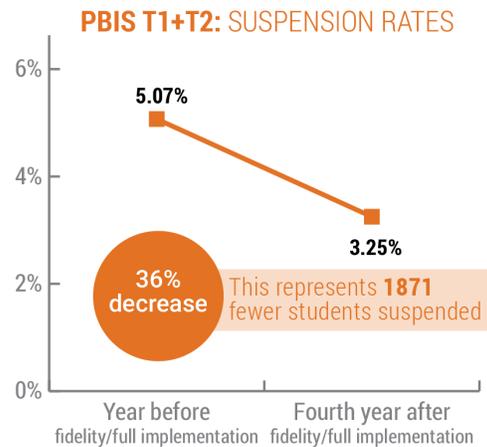
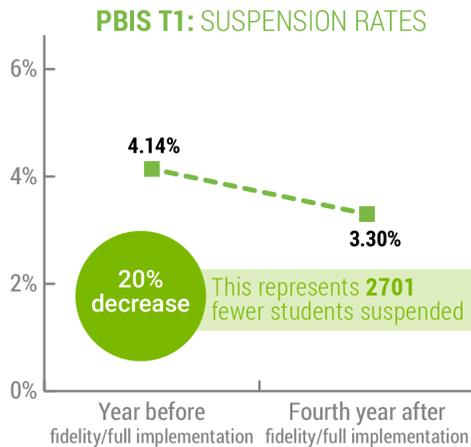
The Wisconsin RtI Center’s annual report displays effects of PBIS after 3 years of implementation.

Do schools sustaining at fidelity/full implementation over time see positive changes in student outcomes and close gaps for underserved students?

WE LOOKED at suspension rates for Wisconsin schools before they reached fidelity/full implementation and contrasted those with their suspension rates after sustained implementation. Specifically, we compared the year before schools reached fidelity/full implementation with their third or fourth year at fidelity/full implementation.

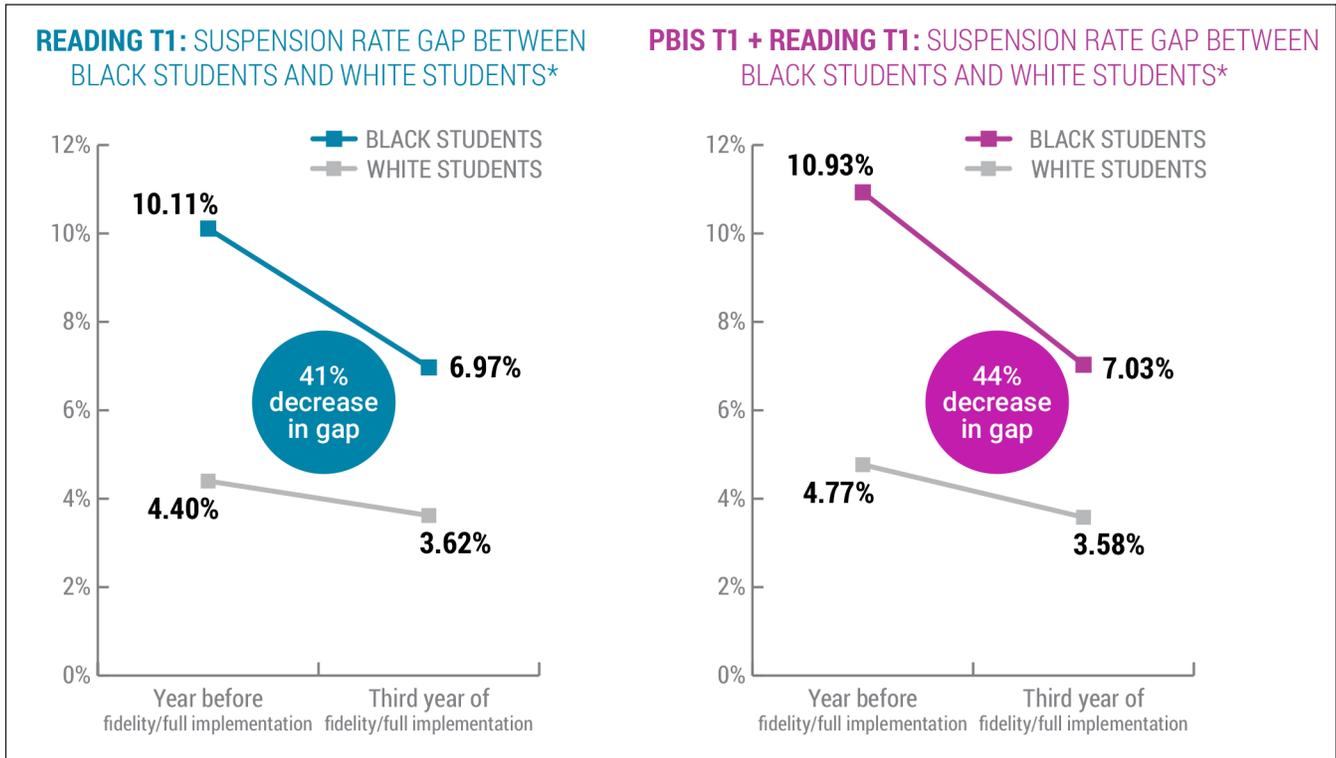
In all of the last three years, we have seen significant overall suspension rates decrease for high implementing schools.

OVERALL SUSPENSION RATE: % OF STUDENTS WITH AT LEAST 1 OUT OF SCHOOL SUSPENSION



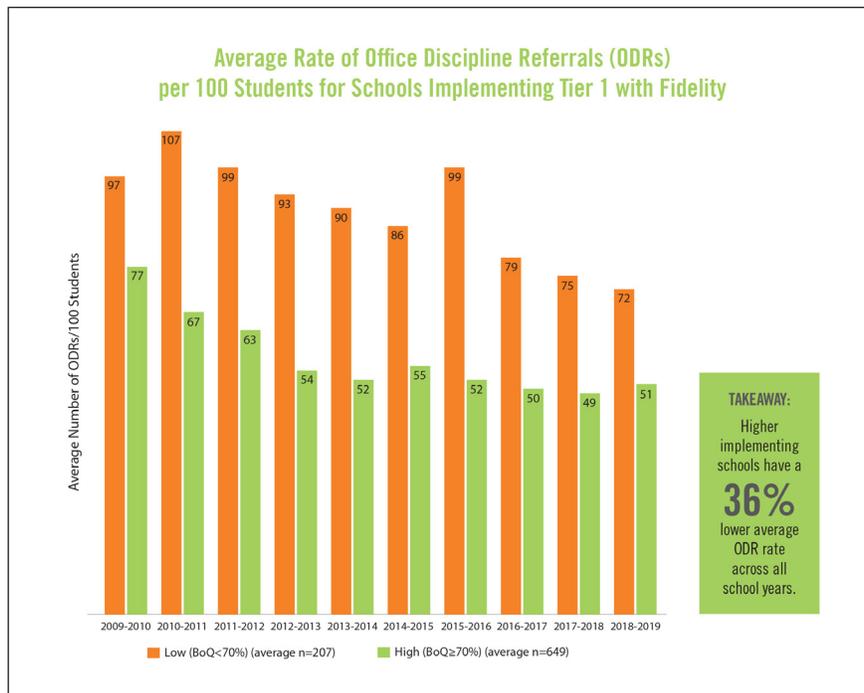


The report also examines racial equity in outcomes using the same method.

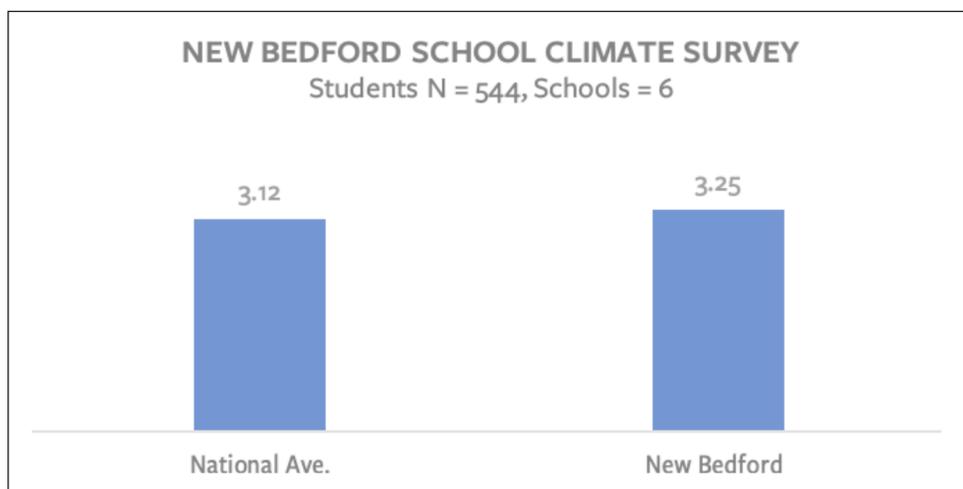




Example of comparing schools by fidelity. In addition to assessing outcomes for schools over time, another common comparison is comparing the outcomes of high and low-fidelity schools. In this example from the FLPBIS Annual Report, ODRs rates are compared to high and low-fidelity schools over time.



Example of comparing schools by involvement in the initiative. In this example from the New Bedford Public Schools Annual Report, school climate survey scores in PBIS schools are compared to the national average.

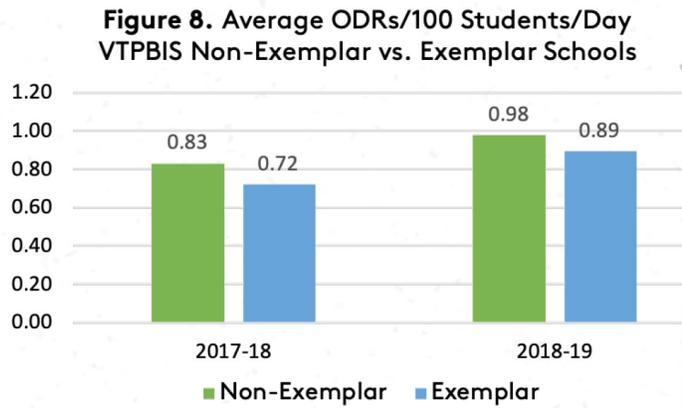




Contextualizing outcomes. One key consideration, especially for long-existing PBIS initiatives is placing small fluctuations in outcomes by year in context. As an example, schools may experience a 90% reductions in suspensions in the first two years of implementation and then experience a 100% increase in suspensions

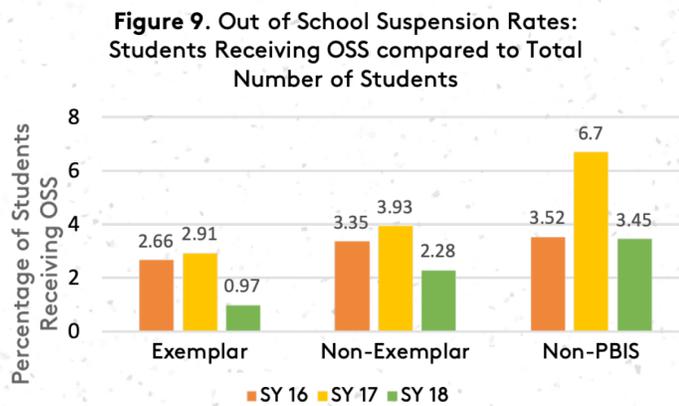
in the third year, Although this data point may seem alarming, it is still an 80% reduction from pre-implementation. It is worth investigating, but it is not by any means an indication that PBIS is not effective. See the following example from the Vermont Annual Report:

While there is an overall increase in major ODRs across PBIS schools using SWIS from last year to this year, Exemplar schools continue to report fewer ODRs than non-Exemplar schools (Figure 8). This will be a priority focus of the VTPBIS State Team in the upcoming year.



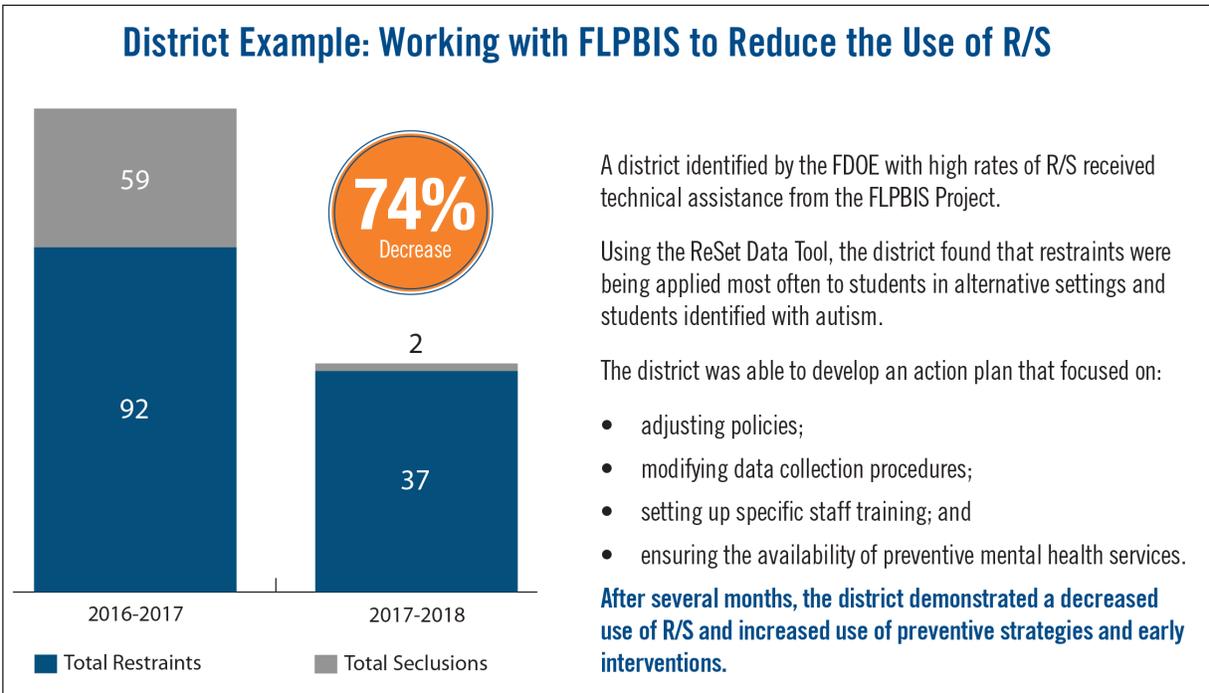
Example of comparing schools by involvement in the initiative. The Vermont PBIS Annual Report also adds another comparison, high-fidelity schools, low-fidelity schools, and other schools in the state. Such data are commonly available for suspensions and academic achievement.

Another indicator of impact is examining Out-of-School Suspension (OSS) rates. While data for the 2018-2019 school year are not yet available, we were able to compare OSS rates across the last three years. When we compared the rate of OSS across VTPBIS Exemplar schools, other VTPBIS schools, and non-VTPBIS schools, the analysis found that Exemplar schools consistently had the lowest OSS rates and non-PBIS schools had the highest rates (Figure 9).





Example of an outcomes vignette. Persuasive outcomes sections include not only documentation of overall effects on important outcomes, but also vignettes to highlight specific outcomes of interest that may not be a target for all schools. For example, the FLPBIS Annual Report includes the following vignette related to reductions in the use of restraint and seclusion through PBIS.



The inclusion of stories from stakeholders can also be powerful, as shown from the Vermont PBIS Annual Report.

“At our January 2019 Data Dive meeting, we noted that school-wide, the motivation behind 45% of ODRs for defiance (major and minor) and disruption (minor) was task avoidance. This data mirrored staff observations and concerns. To address this issue, the Universal team developed sequential lessons... Topics included self-management, identifying strengths and how they were achieved, perseverance, and grit. The fourth and fifth grade classes used restorative circle times to review SWIS data about task avoidance and defiance and to share their thoughts about the data, the impact of behavior on self and others, and how and why to change behavior. From the first to the second half of the year, the total number of SSFs (ODRs) for task avoidance-defiance (major and minor), and disruption decreased by 34% (31% decrease in majors, 36% decrease in minors.)”

- Dothan Brook School, **Eight Year Exemplar**



Yet another example of excellent outcomes evaluation is systems-level change related to the PBIS initiative from PBIS Missouri. The report describes how districts are using PBIS to meet state performance plan standards.

Question 12

To what extent did SW-PBS implementation change educational/behavioral policy?

Missouri School Improvement Plan (MSIP) and State Performance Plan (SPP) indicators help to shape the content of school district policy through their Comprehensive School Improvement Plans (CSIP). These multi-year plans identify goals and indicators to guide areas of improvement and determine desired outcomes to demonstrate achievement.

Participating Missouri school districts are increasingly including SW-PBS in these plans. Currently, 43 districts, approximately 26% of participating districts, have committed to a district-wide implementation, either by having a district level team that meets at least twice yearly to address SW-PBS implementation, or by participating in the DCI or the School Climate Transformation Grant (SCG). Some schools use SW-PBS to address CSIP indicators including:

- orderly and safe schools
- school climate
- Data-Based Decision-Making
- professional development
- appropriate services for all children
- high school transition
- support of parental involvement

MO SW-PBS staff members have been actively involved in the state-level alignment group that

is working to develop guidance for all schools and districts in District Continuous Improvement (DCI). The purpose of this group is to enhance the collaboration as outlined through the SSOS work. One of the outcomes of this work group has been to develop and pilot online training content. MO SW-PBS has been at the forefront of online course development in the state of Missouri.

MO SW-PBS actively supports the “Show Me Success” goals of The DESE. The implementation of MO SW-PBS includes work in early childhood education and secondary transition. Training, networking opportunities, and resource development were areas of focus for the MO SW-PBS team during the 2018-2019 school year. Training across all three tiers supports best instructional practice by Missouri educators. MO SW-PBS has systematically worked to streamline and align training with all other initiatives of The DESE in an effort to maximize resource utilization. MO SW-PBS personnel have been active participants in state level dialogue and development of a framework for Multi-Tiered Systems of Support (MTSS) in an effort to outline best practice in supports across behavioral and academic domains.

MO SW-PBS Goals: 1, 2, 3, 4, 5, 6;

MO DESE Goal: 1, 2, 3, 4;

SPP Indicator: 1, 2, 4, 5, 6, 7, 8, 9, 10, 13, 14



4. Summary and Recommendations

The report should conclude with summative information about the evaluation. At a minimum, the summary can restate the purpose and evaluation questions, with definitive answers to each question.

Overall, readers will want to know the initiative’s status, how successful it is, and what specific changes should be made to improve it. Most evaluation reports include a set of recommendations that arise from the evaluation questions and results.

- To what extent is the PBIS initiative reaching enough students?
- Are the current training and coaching structures sufficient to support implementation with fidelity?
- If implemented poorly, what is needed to implement well?
- If effective, what do we need to sustain it?

- If ineffective, what needs to change, or should it be abandoned?
- What action items should be considered for short-term (e.g., 1 year) and long-term action plans (e.g., 5 years)?

Recommendations and action plans based on the evaluation findings may be included in either internal or external-facing documents. For example, a state or district may review the evaluation results and develop an internal action plan to improve implementation in specific areas. Alternatively, the leadership team may determine that it is important for external stakeholders to understand the goals and action steps developed and may choose to include goals or action steps in a published evaluation report. The following example from Vermont’s Annual Report provides recommendations for next steps as goals within an action plan.

TABLE 2. SY 20 ACTION PLAN FOR SUSTAINABILITY

SY 20 GOALS	SY 20 ACTION STEPS
Help build and support SU/SD-level capacity to manage PBIS by aligning state-level initiatives and expectations	<ul style="list-style-type: none"> • Create state-level executive team to focus on connections between social/emotional/behavioral initiatives • Align messaging around vtMTSS and PBIS • Monthly communication from VTPBIS State Team to SU/SD Coordinators
Build connections between PBIS implementation fidelity and student outcomes	<ul style="list-style-type: none"> • Introduce “TFI Learning Walks” to improve implementation fidelity • Review PBIS fidelity and student outcome data quarterly. Provide feedback to SU/SDs and create professional development based on data trends • Introduce additional social/emotional/behavioral student outcome measures to ascertain student success, beyond problem behavior data
Improve student and family voice in PBIS	<ul style="list-style-type: none"> • Offer opportunities for networking around student and family engagement • Continue to promote restorative practices to enhance engagement • Consider ways for the VTPBIS State Team to engage students and families as stakeholders



Using Results for Continuous Improvement

The final stage in the evaluation cycle involves how the report is disseminated and used in action planning. The ultimate purpose and value of the evaluation process is to determine the results associated with PBIS implementation and then use this information to guide next steps.

Evaluation helps teams determine what to continue doing, what to stop, and what to modify to produce better, faster, and more meaningful results. The steps in this stage are (1) Share Results with Stakeholders, (2) Use Evaluation Results to Improve the Initiative, and (3) Identify Improvements in the Evaluation Process for the Next Evaluation Cycle.

1. Share Results with Stakeholders

It is important to communicate the evaluation report outcomes with stakeholders. Remember that stakeholders provided support (in the form in



increasing priority, visibility, allocation of resources, and participation) for PBIS implementation. Sharing results broadly provide several functions:

- Demonstrate respect and acknowledgment for collaborating and supporting this work
- Report back on what value was added through the initiative
- Encourage continued support for improving and expanding implementation
- Learn what activities and outcomes are more valued by different stakeholder groups

Simply posting a full evaluation report on a website is unlikely to serve these functions well. It is wise to identify how best to share findings in ways (both products and communication avenues) that are tailored to the interests of specific stakeholder groups. Consider what evaluation questions and messages would make these stakeholders interested in this work. Typically, all are interested in improved student

Examples of Tailoring the Message to Stakeholders

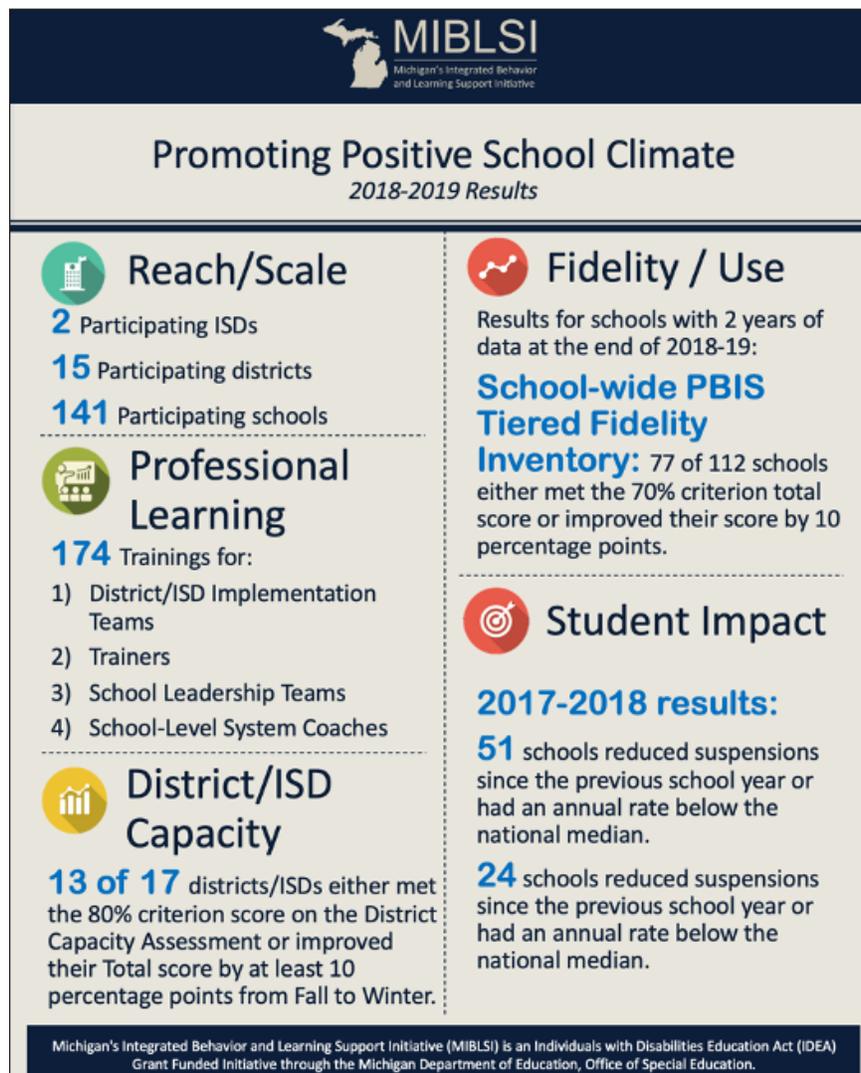
In well-planned evaluation reports, the samples will vary depending on the question being asked. To answer each question accurately, different denominators are needed. A common evaluation logic is as follows:

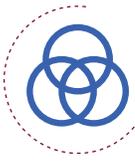
Product Variations	Venue Variations
<ul style="list-style-type: none"> • Full evaluation report with appendices • Executive summaries • At-a-glance reports (e.g., presenting outcomes only) • Infographics • Brief (< 5 min.) videos 	<ul style="list-style-type: none"> • Initiative website • Social media postings • Presentations to state commissioners or school boards • Town halls • Family focus groups • Local news stories



outcomes. Some are also interested in specific results (e.g., educator outcomes, return on investment). Some will respond more positively to numbers and charts, and others will be compelled by illustrative case studies that tell a story. Length is also an important consideration. For example, a state level executive leader requested that updates be provided in a manner that is concise enough to be viewed on his cell phone screen. Tailoring the message to each stakeholder group can increase the likelihood of increased engagement in the initiative.

The following figure is an infographic from Michigan of summarizing PBIS evaluation findings tailored to a funder.





2. Use Evaluation Results to Improve the Initiative

The primary purpose of evaluation is to determine the success of the current plan and resources, then revise the plan as needed to better meet the organization's goals. The evaluation may find that the plan is being effective as is, and stakeholders are pleased with the growth in reach, capacity, fidelity, and effectiveness. In addition, stakeholders may be interested in finding resources or strategies to accelerate that growth. Alternatively, results may show the need to change action plans to meet the initiative's goals.

For more information on using evaluation results for decision making and planning, consider the following resources available at PBIS.org:

- [PBIS Implementation Blueprint](#)
- DSFI and SSFI action planning guides (district and state teams)
- [TFI Action Planning Tool](#) (school teams)

Guiding Questions for Using Evaluation Results for Decision Making

- Are the outcomes still valued by stakeholders?
- How does PBIS relate to other priorities and initiatives?
- How can we make PBIS implementation better, more efficient, and more sustainable?
- Can we allocate supports to the areas that need it by reducing support for areas with positive outcomes?
- If additional resources are needed, have we identified with precision what areas need to be addressed?
- How do we use this evaluation information to embed support for PBIS within existing systems?



Examples of Using Evaluation Results

Two examples from the state of Michigan describe how evaluation results have been used to modify aspects of the PBIS initiative.

Example 1. After reviewing four years of evaluation data, we noticed that schools were no longer implementing PBIS at previous levels of fidelity. Further examination of evaluation data indicated that districts were not supporting school implementation efforts through strategies such as resource allocation, policy and practice alignment, and visible commitment. Based on this information, we moved from a school-by-school initiative to a district level initiative with support for implementing in schools. This resulted in longer lasting and better quality implementation efforts.

Example 2. Our initiative focuses on developing local implementation support capacity that leverages and connects schools, districts, and regional educational agencies. During our early attempts at aligning work across this education cascade, we utilized a linear approach. We first developed capacity at the regional level to support local districts in PBIS implementation. Then we moved on to develop district implementation capacity to support school level PBIS. Finally, we developed school-level teams and classroom PBIS implementation. We found through our evaluation that this approach was taking too long to get to the school and classroom level. The teams at district and regional levels grew weary of the effort without seeing impact on student outcomes. Based on evaluation information we changed our model from a linear approach to more of a simultaneous approach, in which pilot sites are developed while also developing other levels of the system. This new model, often referred to as a transformation zone (Fixsen, Blase, & VanDyke, 2012), allowed us to demonstrated impact faster while also developing local capacity.

3. Identify Improvements in the Evaluation Process for the Next Evaluation Cycle

Finally, evaluations should provide valuable information for improving the next evaluation to take place. It is important to take note of any modifications that

would improve the evaluation plan's effectiveness or efficiency. Examples include changes in measures, timing of measure administration or submission, methods for increasing participation, or ways to partner with stakeholders more meaningfully.

Guiding Questions for Improving the Next Evaluation

- What do we wish we could have evaluated during the first evaluation cycle?
- What aspects of the initiative have changed since the start of the evaluation, and what should change for the next evaluation?
- What changes to the evaluation questions would make the evaluation more meaningful (e.g., changes to target audience, content, timing)?
- What evaluation questions can be eliminated?
- What are emerging or issues that are important to key stakeholders, and how might the next evaluation address these issues?



Summary

Evaluation is a critical part of any PBIS implementation effort. The planning, conducting and reporting of evaluation information has improved both its core features and the process by which those features are adopted by schools, districts, and states. The primary purpose of the *Evaluation Blueprint* is to provide teams faced with a PBIS evaluation with a framework for evaluation planning, useful evaluation measures, and lessons learned from prior evaluation efforts.



References

- Algozzine, R. F., Barrett, S., Eber, L., George, H., Horner, R. H., Lewis, T. J., Sugai, G. (2014). *SWPBIS Tiered Fidelity Inventory*. Eugene, OR: OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. Available at <http://www.pbis.org>.
- Bradshaw, C. P., Koth, K., Bevans, K. B., Jalongo, N., & Leaf, P. J. (2008). The impact of school-wide positive behavioral interventions and supports on the organizational health of elementary schools. *School Psychology Quarterly*, 23, 462-473.
- Bryson, J. M. (2018). *Strategic planning for public and nonprofit organizations: A guide to strengthening and sustaining organizational achievement (5th ed.)*. Hoboken, N.J.: J. Wiley & Sons.
- Cohen, R., Kincaid, D., & Childs, K. E. (2007). Measuring school-wide positive behavior support implementation: Development and validation of the Benchmarks of Quality. *Journal of Positive Behavior Interventions*, 9, 203-213.
- Fixsen, D., Blase, K., & Van Dyke, M. (2012). *From ghost systems to host systems via transformation zones*. Washington, DC: U.S. Department of Education.
- Girvan, E. J., McIntosh, K., & Smolkowski, K. (2019). Tail, tusk, and trunk: What different metrics reveal about racial disproportionality in school discipline. *Educational Psychologist*, 54, 40-59. doi:10.1080/00461520.2018.1537125
- Hoy, W. K., & Feldman, J. (1987). Organizational health: The concept and its measure. *Journal of Research and Development in Education*, 20, 30-38.
- Irvin, L. K., Tobin, T. J., Sprague, J. R., Sugai, G., & Vincent, C. G. (2004). Validity of office discipline referral measures as indices of school-wide behavioral status and effects of school-wide behavioral interventions. *Journal of Positive Behavior Interventions*, 6, 131-147.
- Kim, J., McIntosh, K., Mercer, S. H., & Nese, R. N. T. (2018). Longitudinal associations between SWPBIS fidelity of implementation and behavior and academic outcomes. *Behavioral Disorders*, 43, 357-369. doi:<https://doi.org/10.1177%2F0198742917747589>
- Lee, A., & Gage, N. A. (2020). Updating and expanding systematic reviews and meta-analyses on the effects of school-wide positive behavior interventions and supports. *Psychology in the Schools*. doi:10.1002/pits.22336
- Maslach, C., Jackson, S. E., & Schwab, R. L. (1986). Maslach burnout inventory: Educators survey manual.
- Massar, M. M., McIntosh, K., & Mercer, S. H. (2019). Factor validation of a fidelity of implementation measure for social behavior systems. *Remedial and Special Education*, 40, 16-24.
- McIntosh, K., Barnes, A., Morris, K., & Eliason, B. M. (2014). *Using discipline data within SWPBIS to identify and address disproportionality: A guide for school teams*. Eugene, OR: OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. University of Oregon.
- McIntosh, K., Campbell, A. L., Carter, D. R., & Zumbo, B. D. (2009). Concurrent validity of office discipline referrals and cut points used in schoolwide positive behavior support. *Behavioral Disorders*, 34, 100-113.
- McIntosh, K., Kelm, J. L., & Canizal Delabra, A. (2016). In search of how principals change: A qualitative study of events that help and hinder administrator support for school-wide PBIS. *Journal of Positive Behavior Interventions*, 18, 100-110. doi:10.1177/1098300715599960
- McIntosh, K., Massar, M., Algozzine, R. F., George, H. P., Horner, R. H., Lewis, T. J., & Swain-Bradway, J. (2017). Technical adequacy of the SWPBIS Tiered Fidelity Inventory. *Journal of Positive Behavior Interventions*, 19, 3-13.



Mercer, S. H., McIntosh, K., & Hoselton, R. (2017). Comparability of fidelity measures for assessing tier 1 school-wide positive behavioral interventions and supports. *Journal of Positive Behavior Interventions, 19*, 195-204.

Splett, J. W., Perales, K., Al-Khatib, A. M., Raborn, A., & Weist, M. D. (2020). Preliminary development and validation of the Interconnected Systems Framework-Implementation Inventory (ISF-II). *School Psychology, 35*(4), 255-266.

Sprague, J. R., Colvin, G., & Irvin, L. K. (2002). *The School Safety Survey (2nd ed.)*. Eugene, OR: Institute on Violence and Destructive Behavior.

Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783-805.



Appendix A: Sample Professional Development Fidelity Assessment Form

The following form is used by the Center on PBIS to assess the extent to which trainings and other TA events (e.g., meetings) are delivered with accuracy and in keeping with research on adult learning.

Preparation					
TA Provider Activities	No	Partial	Yes	NA/unable to observe	Comments
1. Facilitator gathered background knowledge on group and needs					
2. Facilitator sent agenda and materials / any pre-work reminders ahead of time					

Delivery					
TA Provider Activities	No	Partial	Yes	NA/unable to observe	Evidence or Example
3. Greet and build rapport with attendees before presentation					
4. Share session expectations and objectives					
5. Deliver content accurately and with high quality					
6. Actively engage participants (e.g., elicit sharing, match pace to audience, use personal examples, invite questions throughout)					
7. Check for understanding and provide opportunities to practice if applicable for each strategy taught					



Delivery					
TA Provider Activities	No	Partial	Yes	NA/unable to observe	Evidence or Example
8. Respond to comments or questions with thanks and positive tone					
9. Finish with big ideas and next steps (and homework as applicable)					
10. Meet intended learning outcomes or objectives					
11. Provide and collect attendee evaluations					

OPTIONAL Self-Reflection

- As the **Facilitator**, use this self-reflection to capture your thoughts for future decision-making.
- As an **Observer**, consider using this section to give narrative feedback for the facilitator of the event.

What went well?

What could be done better next time?

What are the follow-up tasks?



Appendix B: Sample Professional Development Evaluation Form

The following form is used by the Center on PBIS to obtain participant feedback from TA events (e.g., trainings, meetings). It includes items required by OSEP and additional items that are useful for performance feedback and iterative improvement.

PBIS Quality, Relevance, Usefulness (QRU) Survey Items

CENTER EVENTS

Please indicate your agreement with the following statements about the materials you reviewed, using the rating scale below:

0 = Strongly Disagree

1 = Disagree

2 = Agree

3 = Strongly Agree

Name of event you are evaluating: _____

Statement	Rating
1. The content presented in the [event I attended/training I received] was consistent with current research or special education policy.	
2. The content presented in the [event I attended/training I received] addressed current educational problems or issues that are important to me.	
3. The content presented in the [event I attended/training I received] can be readily and successfully used by me in my work.	
4. The presenters were effective in presenting the material.	

What did you find most helpful?

What changes would you recommend?



Appendix C: Additional Measures for Evaluation Plans

Fidelity of Implementation Measures

In addition to the TFI, the following PBIS fidelity of implementation measures are currently used by schools and included in evaluation plans.

a. *Benchmarks of Quality-Revised (BoQ)*. The [BoQ](#) was developed by Center on PBIS personnel at the University of South Florida, with an emphasis on identifying detailed information that can guide professional development, coaching, and technical assistance efforts. The BoQ is completed by the PBIS school leadership team with their external coach and is composed of 53 items (e.g., PBIS team has administrative support, PBIS team meets at least monthly, PBIS team has established a clear mission/purpose). Each item is rated by each member of the team on a 0 to 3 scale with quality descriptors provided for each level of each item to facilitate scoring. The coach then summarizes team member scores, resolves inconsistencies, and generates a final score for each item. The percentage of possible points is used to assess PBIS Tier 1 fidelity. A score of 70% or higher is considered sufficient fidelity to expect improvement in student behavior. The reliability and validity of the BoQ have been documented (Cohen, Kincaid, & Childs, 2007).

b. *Early Childhood Benchmarks of Quality v2.0 (EC-BoQ)*. Efforts to extend positive behavior support to preschool settings have demonstrated important, positive effects in the past decade (Fox, Hemmeter, Jack, & Perez Binder, 2017). One key element associated with this process is measurement of the core features being used in preschool settings using the [EC-BoQ](#). The EC-BoQ addresses many of the core features familiar to PBIS and is composed of 41 items (each scored

2 = In place, 1 = Partially in place, or 0 = Not in place). The EC-BoQ is organized around seven content areas: Leadership Team, Staff Buy-in, Family Engagement, Program-wide Expectations, Professional Development, Responding to Challenging Behavior, and Monitoring outcomes. Reports documenting a Total Score, Sub-scale scores, and an Item summary are provided. The EC-BoQ is available at www.pbis.org and [PBIS Assessment](#).

c. *Facility-wide Tiered Fidelity Inventory v0.2 (FW-TFI)*. Currently in pilot testing, the [FW-TFI](#) is being adopted in a growing number of juvenile justice and other residential programs throughout the country. The FW-TFI was adapted from the School-wide Tiered Fidelity Inventory and reflects the unique language, structure, and priorities of secure care and residential facilities. It measures the extent to which residential care facility personnel are applying core features of Tier 1, Tier 2, and Tier 3 PBIS throughout the facility. Initial content validation data from juvenile justice and residential facility personnel indicate that the FW-TFI items and scoring criteria are appropriate and applicable for these settings, and that the instrument is user-friendly and efficient.

The FW-TFI is divided into three sections: Tier 1: Universal FW-PBIS Features (16 items), Tier 2: Targeted FW-PBIS Features (13 items), and Tier 3: Intensive FW-PBIS Features (17 items). Each section can be used separately or in combination to assess the extent to which core features are in place for each tier. In addition, the FW-TFI is an appropriate tool for monitoring fidelity, regardless of whether programs are implementing PBIS facility-wide, or in a single program only (e.g., Education).



d. Interconnected Systems Framework—Implementation Inventory (ISF-II). The ISF-II provides a valid and reliable measure of the extent to which education and school mental health partners are integrating PBIS and School Mental Health at the school level. The Interconnected Systems Framework (ISF) is a structure and process to integrate PBIS and School Mental Health within school, district, and state educational systems. The goal is to blend resources, training, systems, data, and practices to improve outcomes for all children and youth. There is an emphasis on prevention, early identification, and intervention of the social, emotional, and behavior needs of students. Family and community partner involvement is critical to this framework. The purpose of the ISF-II is to assess the extent to which these key features of ISF are being implemented at the school level to support action planning and continuous improvement. The ISF-II is completed by members of integrated team(s) in a school site and typically takes 30 minutes to complete. It is strongly recommended that the Inventory be completed with a trained facilitator or coach. Training is available at www.midwestpbis.org. The results of completing the Inventory will guide action planning and assist with implementation.

The ISF-II includes 54 items distributed across three tiers in the domains of 1) implementation of school-wide PBIS; 2) teaming; 3) collaborative planning and training; 4) family and youth engagement; 5) intervention selection, implementation, and progress; and 6) school-wide data based decision making. A total score across tiers and domains is generated along with a percentage of total points possible. A score of 80% or higher is considered sufficient fidelity to expect improvement in student social-emotional-behavior outcomes. Total points and percentage of points

possible are also calculated for each tier and each domain within tiers. The reliability, usability, and validity of the ISF-II for measuring implementation have been documented (Splett, Perales, Al-Khatib, Raborn, & Weist, 2020).

- e. Individual Student Systems Evaluation Tool (ISSET).** The [ISSET](#) is a formal research tool designed to determine the extent to which the core features of Tier 2 and 3 PBIS are in place in a school. Data are collected by an external, trained evaluator who (a) visits and tours a school; (b) conducts interviews with an administrator, behavior support team leader, and staff members; and (c) reviews permanent products, including behavior support plans. The ISSET has nine features that are assessed, within three subscales: Foundations, Targeted Interventions, and Intensive Individualized Interventions.
- f. School-wide Evaluation Tool (SET).** The [SET](#) is a formal research tool designed to determine the extent to which the core features of Tier 1 PBIS are in place in a school. Data are collected by an external, trained evaluator who (a) visits and tours a school, (b) interviews at least 10 faculty/staff, (c) interviews 10-15 students, (d) interviews the building PBIS leadership team, and (e) reviews permanent products from training, discipline management and coordination.

The SET has seven features that are assessed: School-wide expectations defined, School-wide expectations taught, Reward System, Behavioral Violations system, Data/Decision-making, Management, and District Support. Each feature of the SET includes between 2 and 8 items (total of 28 items) each of which are assessed as “in place = 2,” “partially in place = 1” or “not in place = 0.” A percent implemented score is determined for each of the seven features, and a mean across



the features is used to index the “Total Score” for the SET. A Total Score at or above 80% is used to document “criterion fidelity.” The level at which improvement in student outcomes can be expected.

It typically takes the majority of a day for a full SET evaluation to be conducted at a school. Results from the SET have been demonstrated to be valid and reliable (Horner et al., 2004), and SET data have been used in numerous research reports to document adoption of Tier 1 PBIS.

- g. Self-Assessment Survey (SAS).** The [SAS](#) was one of the first PBIS measures developed in the 1990s. Before use of the multi-tiered triangle became common, PBIS content was organized around core features at four levels: (a) school-wide (b) non-classroom, (c) classroom, and (d) individual student. Both before and after PBIS implementation efforts were made, it was helpful to gauge (a) the extent to which all faculty and staff in a school perceived themselves to be implementing these core features, and (b) the level of importance the faculty and staff gave to improving implementation of these features.

The SAS is a low cost/ low effort option for summarizing the perceptions of faculty and staff in a school. A common recommendation is to consider launching PBIS implementation only if at least 80% of faculty and staff identify features in the School-wide level of the SAS as in need of improvement and a medium to high priority.

The “School-wide” section of the SAS includes 18 items. Each item is rated based on current status (in place, partially in place, or not in place) and perceived priority for improvement (high, medium, low). When all (or most) faculty and staff in a school have scored each item (either using paper/

pencil, or online) a graphic summary is created depicting the proportion of respondents rating items In place, Partially in place, or Not in place, and the proportion of respondents rating the same items as High, Medium or Low priority. As noted above, schools that are most successful adopting PBIS reach staff agreement that the School-wide section of the SAS has at least 80% of respondents rating features as partially or not in place, and medium or high need for improvement.

The “non-classroom,” “classroom,” and “individual student” sections of the SAS have 9, 11, and 8 features respectively. Each is scored similarly to the school-wide section and graphic results are developed with similar summary of “status” and “priority.” The “non-classroom,” “classroom” and “individual student” sections of the SAS have become less commonly used as newer fidelity measures have emerged, but the “school-wide” section continues to be viewed as helpful given that the SAS is the only measure where input is obtained from the full faculty/staff of a school. Evaluation efforts focusing on providing a voice for the full faculty/staff may find the SAS useful.

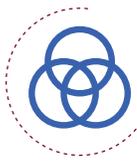
- h. Team Implementation Checklist v3.1 (TIC).** The [TIC](#) is a brief self-assessment of 22 items associated with installation and initial implementation of PBIS. It is designed as a progress monitoring tool for school teams implementing Tier 1 PBIS. The TIC should be completed by the full team (preferably with the district coach). Information from the TIC should be used to guide development of an action plan. Once the school team has reached 80% implementation, the team can switch to more comprehensive measures, such as the BoQ or TFI.



Less Common Outcome Measures

The following measures are not commonly used in evaluations but have been used to evaluate effects of PBIS, most commonly as part of research studies.

- a. Organizational Health Inventory (OHI).** The OHI (Hoy & Feldman, 1987) is a survey completed by school personnel assessing nine factors: Direction, Innovation and Learning, Leadership, Coordination & Control, Capabilities, Motivation, Work Environment, Accountability, and External Orientation. The OHI has been used in formal research studies as opposed to evaluations, but implementation of PBIS has been shown to improve the organizational health of schools (Bradshaw, Koth, Bevans, Jalongo, & Leaf, 2008).
- b. Teacher efficacy.** Evaluators examining PBIS implementation have on occasion asked if the implementation of PBIS is related to teacher's self-perception of how effective they were with their students. The Teachers' Sense of Efficacy Scale (TSES; Tschannen-Moran & Woolfolk Hoy, 2001) and the Maslach Burnout Inventory–Educators Survey (MBI-ES; Maslach, Jackson, & Schwab, 1986) have been used to indicate effects of PBIS implementation on teacher outcomes.



Appendix D: Sample Evaluation Plan

Evaluation Question	Measures	Metric/Data Produced	Collection Cycle	Possible Decisions Based on Evaluation
Reach				
What schools are receiving support (i.e., in the initiative)?	<ul style="list-style-type: none"> School Profile Form Schools Implementing (PBIS Assessment/Eval) 	<ul style="list-style-type: none"> Name, address, contact, enrollment, grade levels, demographics, start year Number of schools reporting fidelity 	<ul style="list-style-type: none"> When a school begins implementation Updated annually 	Do we need to implement PBIS in more schools or specific schools?
How has PBIS sustained over time?	<ul style="list-style-type: none"> Schools Implementing (PBIS Assessment/Eval) 	<ul style="list-style-type: none"> Number of schools reporting fidelity per year 	<ul style="list-style-type: none"> Updated annually 	Do we need to change the initiative to improve implementation capacity?
Process				
What steps are the Leadership Team taking?	<ul style="list-style-type: none"> PBIS State Systems Fidelity Inventory (SSFI) Action Plan Items Report 	<ul style="list-style-type: none"> Number of action plan goals met 	<ul style="list-style-type: none"> Completed annually 	
Who in the state/district is providing training?	<ul style="list-style-type: none"> Trainer List 	<ul style="list-style-type: none"> List of trainers 	<ul style="list-style-type: none"> Updated annually 	Do we have sufficient internal training capacity?
Who in the state/district is providing coaching?	<ul style="list-style-type: none"> Coaches' Survey List of coaches' meetings and trainings provided 	<ul style="list-style-type: none"> Used to identify training needs Number of events and coaches present at each 	<ul style="list-style-type: none"> Collected annually Collected after each training event 	Do we have sufficient internal coaching capacity?
How are district coaches supported?	<ul style="list-style-type: none"> Coaches' Survey List of coaches' meetings and trainings provided 	<ul style="list-style-type: none"> Used to identify training needs Number of events and coaches present at each 	<ul style="list-style-type: none"> Collected annually Collected after each training event 	Do we need to better support coaches in their roles?
What professional development has been delivered?	<ul style="list-style-type: none"> List of events, training content, and school teams participating 	<ul style="list-style-type: none"> Number of events and teams present at each 	<ul style="list-style-type: none"> Collected after each PD event 	



Evaluation Question	Measures	Metric/Data Produced	Collection Cycle	Possible Decisions Based on Evaluation
Was professional development perceived as useful by participants?	<ul style="list-style-type: none"> Workshop Evaluation Form 	<ul style="list-style-type: none"> % reporting that the event was useful 	<ul style="list-style-type: none"> Collected after each PD event 	Do we need to provide more practical and relevant strategies/tools to training participants?
How often is input obtained from stakeholders at each school?	<ul style="list-style-type: none"> Stakeholder Input and Satisfaction Survey (SISS): Family, School Personnel, Student 	<ul style="list-style-type: none"> % of schools administering SISS to all 3 stakeholder groups 	<ul style="list-style-type: none"> Annually (by end of November) 	Do key stakeholders feel like they are providing valuable input and that their voices are heard?
How can the PBIS initiative be improved according to users and recipients?	<ul style="list-style-type: none"> Stakeholder Input and Satisfaction Surveys (SISS): Family, School Personnel, Student 	<ul style="list-style-type: none"> % satisfaction Descriptive feedback 	<ul style="list-style-type: none"> Collected in each school annually 	What actionable steps can the leadership team to make the initiative more relevant and valuable for stakeholders?
Capacity				
What resources are available to support PBIS?	<ul style="list-style-type: none"> Organizational FTE Budget information 	<ul style="list-style-type: none"> List of resources (individuals, funding sources) 	<ul style="list-style-type: none"> Updated annually 	Are we allocating (or reallocating) adequate resources for PBIS implementation?
What capacity of the state to support PBIS?	<ul style="list-style-type: none"> PBIS State Systems Fidelity Inventory (SSFI) 	<ul style="list-style-type: none"> Percent of implementation 	<ul style="list-style-type: none"> Completed annually 	What can we do to ensure that there is adequate state level support for PBIS implementation?
What capacity of each district to support PBIS?	<ul style="list-style-type: none"> PBIS District Systems Fidelity Inventory (DSFI) 	<ul style="list-style-type: none"> Percent of implementation 	<ul style="list-style-type: none"> Completed annually 	What can we do to ensure that there is adequate district level support for PBIS implementation?



Evaluation Question	Measures	Metric/Data Produced	Collection Cycle	Possible Decisions Based on Evaluation
Fidelity				
What is the fidelity of implementation of PBIS at each school?	<ul style="list-style-type: none"> PBIS Tiered Fidelity Inventory (TFI): Tiers 1, 2, 3 	<ul style="list-style-type: none"> % of schools implementing each Tier with fidelity (criterion: 70%) 	<ul style="list-style-type: none"> All tiers completed by end of September, then progress monitoring with one tier quarterly 	How can we support educators in implementing PBIS correctly and consistently enough to produce meaningful student outcomes?
What is the fidelity of implementation of related practices (e.g., SEL program) at each school?	<ul style="list-style-type: none"> Lesson calendar and check-off Observation of lessons 	<ul style="list-style-type: none"> % of lessons taught on time (goal: 100%) % of lesson's critical features observed (goal: 90%) 	<ul style="list-style-type: none"> Submitted annually by each school team Submitted annually by each school team 	
What schools can serve as local implementation demonstrations (across tiers and levels)?	<ul style="list-style-type: none"> PBIS Tiered Fidelity Inventory (TFI): Tiers 1, 2, 3 School Profile Form 	<ul style="list-style-type: none"> List of schools and contact information Brief case studies for evaluation report 	<ul style="list-style-type: none"> Updated annually 	How can we showcase the effects of PBIS in a range of schools?
Outcomes				
If PBIS is implemented to criterion, is there improvement in social & academic outcomes for students?	<ul style="list-style-type: none"> School Outcome Data Summary Form <ul style="list-style-type: none"> Attendance Office Disc. Referrals Suspensions Academic Achievement 	<ul style="list-style-type: none"> % reductions in: <ul style="list-style-type: none"> Office Disc. Referrals Suspensions ALSO: Student/Admin. time recouped % increases in: <ul style="list-style-type: none"> Attendance % of students meeting academic expectations 	<ul style="list-style-type: none"> Collected Annually 	How do we communicate these result in a way that produces further support for the initiative?
Do schools implementing PBIS to criterion have more positive school climate?	<ul style="list-style-type: none"> School Climate Survey (SC Suite) <ul style="list-style-type: none"> Student Family School Personnel 	<ul style="list-style-type: none"> Mean school climate rating 	<ul style="list-style-type: none"> Annually (by end of March) 	



Additional Resources

[Florida PBIS Project Evaluation Report 2018-19](#)

[PBIS Missouri Evaluation Report 2018-19](#)

[Wisconsin RtI Network Evaluation Report 2019-20](#)

[New Bedford Public Schools \(MA\) Evaluation Report 2020](#)