Design Options for an Impact Evaluation of Schoolwide Positive Behavior Supports

Final Design Report

Suyapa Silvia, RTI International
Jonathan Blitstein, RTI International
Joni Splett, University of South Carolina
Linda Dusenbury, Tanglewood Research, Inc.

William Hansen, Project Director, Tanglewood Research, Inc.

Lauren Angelo, Project Officer, Institute of Education Sciences

April 2013

CONTENTS

APPENDIXES

EXHIBITS

1.0 STUDY BACKGROUND AND PURPOSE

This document was prepared for the Institute of Education Sciences (IES) as part of the Design of an Impact Evaluation of Schoolwide Positive Behavior Supports (SWPBS) contract. The report presents preliminary design options for an impact evaluation of SWPBS, which is a multitiered, systematic framework intended to improve classroom and school climate; student social, emotional, and behavioral competence; and academic achievement for all students (Horner, Sugai, and Anderson 2010). The motivation for studying SWPBS is discussed in Section 1, which provides background information, the prevalence of problem behaviors, and the federal support for SWPBS. The study's research questions are also found in Section 1. Section 2 provides an overview of the SWPBS framework and components as well as examples of implementation approaches and targeted interventions that could be implemented within the SWPBS framework. Section 3 describes the available training and implementation support for SWPBS. Section 4 presents several design options for an impact evaluation. Finally, section 5 addresses other design considerations. Additional study products are appended to this report as follows: Appendix A: Literature Review Memo; B: Provider Memo; and C: Technical Working Group (TWG) Meeting Summary.

1.1. Problem Behaviors and Their Impact on Academics and the School Environment

Negative behaviors, including nonphysical aggression and more serious violent behaviors, are prevalent in a wide variety of social and physical contexts, including schools. The most recent data from the National Center for Education Statistics (Robers, Zhang, and Truman 2012) show that in 2009-10, 74 percent of public schools reported one or more violent incidents of crime, whereas 68 percent reported one or more other types of incidents. Some 23 percent of public schools reported daily or weekly bullying among students during the 2009-10 school year, whereas 9 percent reported "widespread disorder in classrooms." Among high school students, in 2009, 31 percent reported that they had been in a physical fight on school property in the past year. Student problem behaviors also impact school staff and the classroom environment. During the 2007-08 school year, 7 percent of elementary school teachers and 8 percent of secondary school teachers reported being threatened by a student. Moreover, in 2007-08, 39 percent of secondary school teachers and 33 percent of elementary school teachers reported that student misbehavior interfered with their teaching. In response to student misbehavior, 39 percent of public schools took at least one serious disciplinary action against a student, including suspensions, transfers to specialized schools, and removals from school for the remainder of the year.

The short- and long-term impacts of students' negative behavior on the learning environment and on student outcomes is well documented. Disruptive behaviors have been

shown to contribute to loss of instruction time (Arnold 1997) and academic problems, including poor achievement, grade retention and school dropout (Tremblay et al. 1992; Wagner 1995). Furthermore, co-occurring behavior and academic problems are associated with later negative outcomes such as poor academic progress and postsecondary outcomes (Wagner 1995), receipt of special education services, conduct problems, criminal arrest, and antisocial behaviors (Darney et al. 2013; Moffitt 1993; Reinke et al. 2008; Schaeffer et al. 2004).

Schools have an opportunity to teach and reinforce appropriate behaviors and skills as well as address negative behavior, and the classroom environment plays an important role. Kellam and colleagues (1994) demonstrated the role of classroom orderliness and teacher classroom management on later aggressive behavior. The study found that aggressive first-grade boys in orderly classrooms had significantly lower odds of being highly aggressive in middle school compared with aggressive first-grade boys in chaotic classrooms. Another study found that negative relationships with teachers and peers in the kindergarten classroom increased the likelihood of developing a pattern of escalating externalizing behavior (Silver et al. 2010).

A number of risk factors for problem behaviors have been identified that may be amenable to intervention in the school setting; they include poor social skills, perceived low social competence, and negative attitudes toward school. Further studies suggest that positive school bonding is a protective factor against antisocial behaviors; that is, students who form a positive school bond are less likely to engage in problem behaviors (Dryfoos, 1990; Hawkins and Weis, 1985). Simons-Morton and colleagues (1999) found that school bonding, perceived school climate, and school adjustment were all negatively associated with problem behaviors. Studies also suggest that positive school bonding may be promoted through teaching social skills, developing social competence, and improving school climate (Hawkins and Catalano, 1990; Schaps and Battistich, 1991).

The SWPBS framework is one strategy for addressing and preventing problem behaviors and promoting protective factors in school settings. The next section provides background information on the federal policy and support for this approach.

1.2. Federal Policy and Support for SWPBS

In the 1997 amendments to the Individuals with Disabilities Education Act (IDEA), Individualized Education Program (IEP) teams were required to consider the use of positive behavior interventions and supports for any student whose behavior impedes his or her learning or the learning of others. This requirement applies to students in prekindergarten through 12th grade. In 2004, IDEA was amended to permit, and sometimes require, schools to use IDEA Part B special education funds to implement educational and behavioral services and supports to provide Coordinated Early Intervening Services (CEIS). CEIS are provided to students in

kindergarten through 12th grade (with an emphasis on students in kindergarten through third grade) "who are not currently identified as needing special education or related services, but who need additional academic and behavioral support to succeed in a general education environment" (IDEA 34 C.F.R. 300.226). The CEIS provision in IDEA allows for up to 15 percent of funds to be used this way. Funds may be used for both professional development and for direct services related to academic and behavioral interventions. During the 2008-09 school year, approximately 60 percent of districts using IDEA Part B funds for CEIS were implementing behavioral interventions (Bradley et al. 2011).

The Technical Assistance Center on Positive Behavioral Interventions and Supports (PBIS) (TA Center) has been supported by the Office of Special Education Programs in the U.S. Department of Education since the 1990s. Its mission is to assist states and school districts in large-scale implementation of SWPBS to achieve reduction in problem behavior and enhanced learning environments.

1.3. Purpose and Research Questions

The purpose of this design study was to assess the feasibility of large-scale implementation and impact evaluation of SWPBS using a randomized controlled trial design and to prepare a design-options report. SWPBS is considered a promising approach for managing behavior at the individual student, classroom, and whole-school levels and for improving school climate and academic performance. Although some information is available regarding the effectiveness of this approach, more rigorous studies are needed.

The design study's key research questions are shown in exhibit 1. The remainder of this report addresses each of the research questions: section 2 addresses questions 1 and 2; sections 3 and 4 address questions 3 and 4, respectively; and section 5 presents additional issues to consider in the design and implementation of SWPBS.

Exhibit 1. Key research questions for the design study

1. What are the components of SWPBS universal supports that research demonstrates have the most promise in improving the behavior and academic performance of students?
2. Are there available and exemplary universal support interventions as well as targeted interventions that can be implemented within the SWPBS framework?
3. What is the feasibility of training schools in SWPBS universal supports and targeted interventions?
4. What evaluation design options are feasible to study the impacts of SWPBS universal supports and targeted interventions?

2.0 SWPBS: OVERVIEW AND CORE FEATURES

This section begins with an overview of the SWPBS framework and its key components. The section then provides examples of approaches to SWPBS implementation and targeted programs that could be implemented within the SWPBS framework. This section addresses research questions 1 and 2:

- What are the components of SWPBS universal supports that research demonstrates have the most promise in improving the behavior and academic performance of students?
- Are there available and exemplary universal support interventions as well as targeted interventions that can be implemented within the SWPBS framework?

2.1. Overview and Key Components of SWPBS

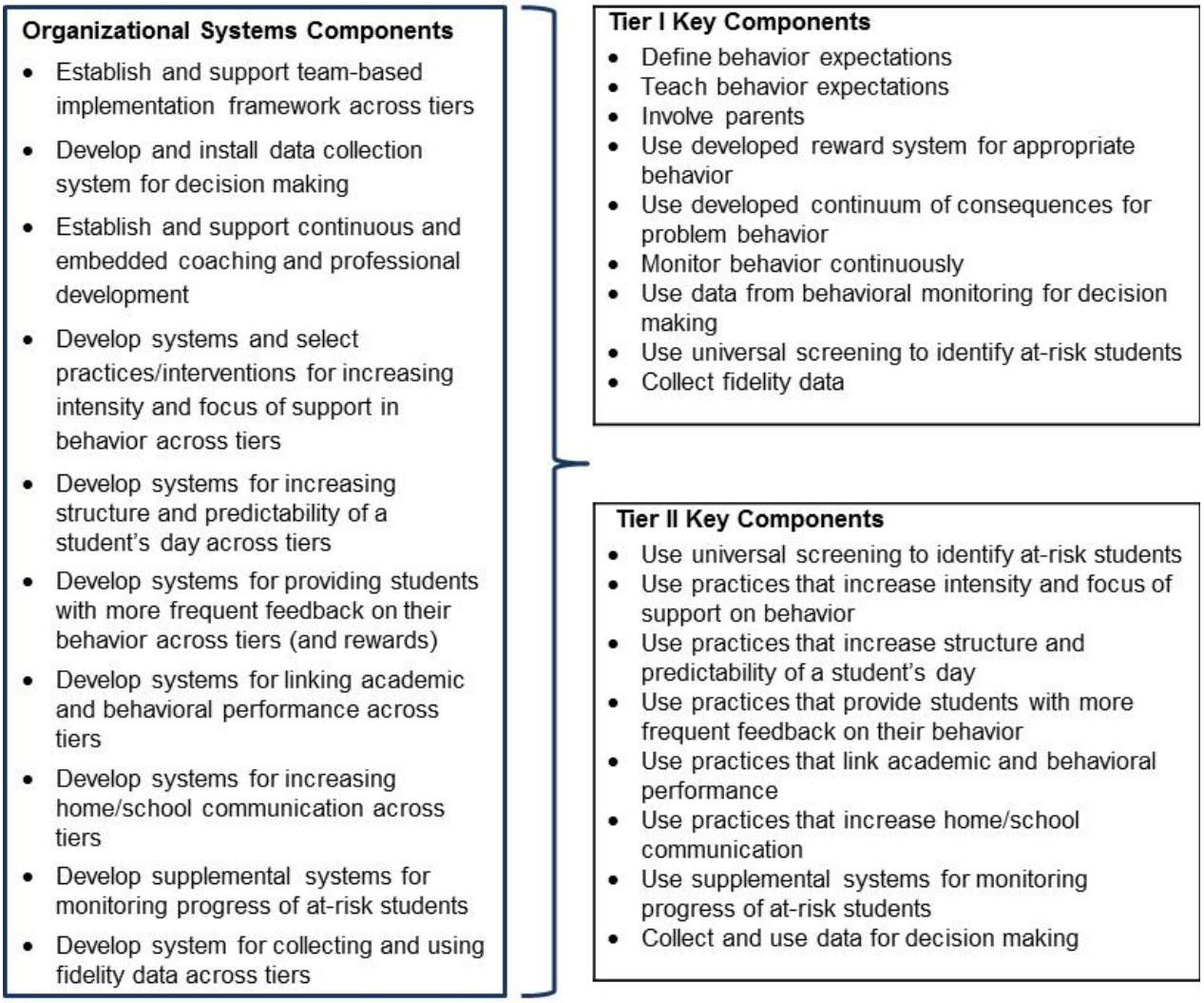
SWPBS is a three-tiered, systematic framework intended to improve classroom and school climate; student social, emotional, and behavioral competence; and academic achievement for all students (Horner, Sugai, and Anderson 2010). SWPBS is based upon the assumption that student academic and social behavior is linked to the school environment and climate (Zins and Ponti 1990). It promotes the use of effective practices from behavioral science and social-emotional instruction to redesign the school environment and improve school climate. SWPBS is also commonly referred to as "Positive Behavior Interventions and Supports" following the language in IDEA (Bradshaw, Mitchell, and Leaf 2010a) and Positive Behavior Supports (Horner et al. 2009). While these terms refer to the same core elements that compose SWPBS, SWPBS is used here to be inclusive of the framework's *schoolwide* nature.

The first tier, commonly referred to as primary or universal, is implemented schoolwide for all students and includes proactive strategies to teach and reinforce positive behaviors. Students not responding to Tier I supports may receive additional, targeted intervention services. Tier II services, also commonly referred to as secondary or targeted services, are more specialized interventions and supports for those students not responding to Tier I supports. Common Tier II interventions include Check-In/Check-Out, Check and Connect, First Step to Success, and group counseling (Horner et al. 2010). Additional individualized support may be provided in Tier III, commonly referred to as tertiary or intensive or indicated, to students who did not or are unlikely to respond to Tier I and II supports. Common Tier III interventions include functional behavioral assessments, individual mental health services, and special education referrals (Horner et al. 2010). Because students' needs are assessed regularly and support levels are tied to need, additional interventions and supports can ideally be delivered as early as possible.

All tiers are integrated into a SWPBS framework with four basic operating elements: data, practices, systems, and outcomes (Sugai 2007). Using these elements, school leadership teams tailor the delivery of evidence-based behavioral and social-emotional supports at all three tiers so that all students, including those with or at risk of behavior concerns, can experience academic and behavior success. The "data" element refers to collecting and analyzing data or information to facilitate data-based decision making for purposes such as (a) universal screenings and progress monitoring of targeted and selected interventions, (b) evaluation of the evidence base for a particular practice, and (c) assessing the implementation fidelity of a given practice. The "practices" element of SWPBS encourages schools to adopt practices at all three tiers that are based on theoretically sound principles and empirically based interventions. These practices are aligned and integrated into a continuum of supports accessible to all students. In general, all students benefit from Tier I level interventions; however, some students may require additional supports through Tier II/III practices and interventions. The third element, "systems," focuses on enhancing sustainable, quality implementation by emphasizing team-based action planning, training and coaching to fluency, and continuous data-based evaluation. SWPBS is commonly implemented through a school-based leadership team comprising teachers, administrators, and other pupil support staff. In addition, coaching and training capacity are established to provide training and technical assistance throughout the systems change process. The final element, "outcomes," refers to student academic and behavior outcome indicators (Sugai 2007) that are relevant and measureable.

In addition to the four basic elements, SWPBS is characterized by key components or principles of practice. The study team identified key components of the SWPBS framework based on a review of the literature; these components were further modified with input from the Technical Working Group. These key organizational and Tier I and II components are summarized in exhibit 2. A number of key organizational system components are developed to support implementation of the key Tier I and II components, including, for example, a support system for SWPBS (e.g., leadership team, data collection system), systems for supporting behavior monitoring and progress, and a system for collecting implementation fidelity data. Development of these organizational systems by the school teams drives the design and implementation of the core tier components. The teams implement and use these systems schoolwide in Tier I and with more intensity and focus in Tier II. Across SWPBS tiers, these components guide the practice and activities of school personnel.

Exhibit 2. SWPBS key components



In keeping with the tiered nature of SWPBS, Tier I components are implemented and applied schoolwide to all students in a school. Tier II components represent a more intensive and targeted application of the Tier I components and are provided to smaller groups of students. For example, an initial task of many school leadership teams is to establish a core set of prosocial skills they want to teach and encourage schoolwide (e.g., "safe, respectful, and responsible learners"). They operationalize what these skills look like in various school settings such as the hallway, gym, cafeteria, bathrooms, and classroom, and develop a system for teaching and reinforcing them, including signage throughout the building, classroom lessons, and token reinforcers (e.g., Warrior bucks). Once these initial steps are taken and the school year begins, the team begins to regularly review available data regarding schoolwide behavior problems and academic achievement to identify environmental modifications that may be necessary (e.g., increased adult supervision) and students in need of additional support. Students who are

identified as needing Tier II or III services are also monitored by school teams through data collection on their progress; this monitoring facilitates data-based decision making at the team level. Thus, the SWPBS key components are implemented across the multitiered continuum of academic and behavioral interventions.

In the next section, we provide an overview of representative framework implementation approaches and describe examples of Tier I and Tier II programs that could be implemented within the SWPBS framework. While Tier III services also form part of the SWPBS systematic approach, these will not be explicitly featured in any of the design options for an impact evaluation of SWPBS and thus are not discussed further in the remainder of this report.

2.2. Framework Implementation Approaches

Various approaches to training schools in and implementing the SWPBS framework are available to schools and are currently in use at the national, state, district, and school levels. As part of this contract, a literature review was conducted to identify implementation approaches (see Attachment A: Literature Review Memo). The study team considered several factors in identifying approaches: (1) the approach promotes, implements, and supports the SWPBS framework; (2) the approach focuses on social, emotional, and behavioral supports; and (3) the approach has at least preliminary supporting evidence of effectiveness and/or is in widespread use. The following four examples of framework implementation approaches are featured in this report:

- Technical Assistance Center for PBIS (TA Center) (supported by the Office of Special Education Programs; OSEP)
- Project Achieve (developed by Dr. Howard Knoff)
- Safe & Civil Schools (developed by Dr. Randy Sprick)
- Best Behavior (developed by Dr. Jeffrey Sprague).

It should be noted that these four approaches represent only some of the positive behavior support approaches currently employed by schools.

2.2.1. Key Differences Across Approaches

Although all four implementation approaches address the same key components or principles, there is some variation in how these key components are implemented. One such difference relates to the sequencing of training for school staff. Variations include (1) whether a leadership team is trained and then is responsible for training the larger school staff (TA Center, Safe & Civil Schools, Best Behavior) or if training is provided to the larger school staff by the provider (Project Achieve); (2) whether the training and implementation in Tier II is rolled out in the first year following training in the framework and Tier I (Project Achieve), rolled out in the

second year (TA Center, Best Behavior), or rolled out in the third year (Safe & Civil Schools); and (3) the intensity of the training (e.g., Project Achieve provides specialized training for each building-level committee focused on a particular aspect of implementation). Other salient differences include whether the training in Tier II is focused on all at-risk students (TA Center) and whether multiple supports are implemented based on students' needs (Project Achieve). Another difference across approaches is in the Tier I and II programs employed. Safe & Civil Schools and Project Achieve have each developed their own Tier I and II programs, and it appears that schools using these two implementation approaches only implement Tier I and II programs affiliated with the particular approach. Other implementation approaches, including the TA Center and Best Behavior do not specify particular Tier I or II programs that schools must adopt. Instead, they provide training on how to implement or coordinate integration of a variety of programs within the framework.

One of the three design options (design option 3) discussed in section 4 provides an opportunity to potentially contrast two such approaches to training schools in SWPBS or offer a variant on a key aspect of training. This design would provide information about the differential impacts on student and school outcomes of distinct approaches to training schools in SWPBS; no studies have examined such a contrast.

2.2.2. Evidence Base for Approaches

The study team examined the evidence base for each of the four approaches and included at most two studies that met the inclusionary criteria for each approach (e.g., includes a comparison group, assigned a minimum of five schools per condition, and includes pretest-posttest measurement with statistical analyses to test for significant differences).

The TA Center on PBIS approach reports positive effects in both behavior (Waasdorp, Bradshaw, and Leaf 2012) and academic outcomes for reading (at time 2 only) (Horner et al. 2009) based on two experimental studies. Project Achieve reports positive effects in both academic and behavior outcomes, but results are based on only one, quasi-experimental study (Harding et al. 2008). Results of a quasi-experimental study of Safe & Civil Schools indicated a positive effect on state-standardized test scores over four years of treatment (Madigan and Cross 2009). Finally, an experimental study of Best Behavior did not find significant intervention effects on behavior outcomes (Silvia et al. 2011). A quasi-experimental study reported percent change in behavior outcomes favoring the treatment schools but did not report statistical tests. The impact of Best Behavior on academic outcomes was not tested in either study. (See Attachment A: Literature Review Memo for details about each of these studies.)

2.3. Tier I and Tier II Programs

As described earlier, Tier I (primary, universal) support programs are implemented schoolwide for all students and include proactive strategies to teach and reinforce positive behaviors. Tier II (secondary, targeted) programs are provided to students not responding to Tier I supports. These students receive additional, specialized intervention services. In practice, extending the SWPBS framework to include Tier II and providing a Tier II intervention can be viewed as an additional component and often occurs after training in and implementation of the framework and Tier I program. The third design option described in section 4 provides an opportunity to examine the value of the Tier II component as an "add-on" to the framework and Tier I component.

A review of the literature identified examples of behavior intervention programs targeting students' social, emotional, and behavioral competence from kindergarten to grade 12 that could potentially be implemented as Tier I or Tier II (or both) interventions within the SWPBS framework. Note that not all of these programs have been associated with or implemented within the SWPBS approach. These Tier I and Tier II programs can be grouped according to method of delivery, as follows:

- Tier I Classroom Lesson Programs—These programs include a discrete number of lessons delivered in the classroom to promote and teach a specific behavior or skill. Examples include: Lions Quest Skills for Adolescence, Positive Action, Too Good for Drugs, and Too Good for Violence.
- Tier I Classroom Lesson Programs, with additional Tier II Small Group Component—These programs are similar to those in the first category but also include Tier II small group or individual activities intended for students in need of additional behavioral supports. Examples include: Incredible Years, Second Step Violence Prevention Program, Skillstreaming, and Stop and Think Social Skills Program.
- Other Tier I Programs—Additional Tier I programs include the Olweus Bullying Prevention Program, which has classroom lessons but also has additional schoolwide components, and CHAMPS, which targets teacher classroom management and does not include any direct student components.
- Tier II Small Group Lesson Programs—These programs include those in category 2 above and deliver a discrete number of lessons in a small group setting to address a specific problem behavior. An example of a Tier II program that is not also a Tier I program is the Coping Power Program, which includes child group meetings and parent group or individual sessions.

- Tier II Individual, Check-In Programs—These programs are delivered individually and involve students checking in with a mentor daily to weekly and receiving regular (multiple times daily to weekly) behavioral feedback. Check & Connect and Check-In/Check-Out are two programs commonly used within the SWPBS framework.
- Other Tier II Programs—In addition to having a school-based intervention component, First Steps to Success includes lessons taught to parents and delivered by parents to their child as well as a behavioral feedback system. An additional example is CHAMPS, described above under category 3.

Exhibit 3 lists these programs and indicates their potential use at Tier I, Tier II, or both. The exhibit also provides information on the age/grade levels targeted by the intervention and identifies which SWPBS approach(es) each program has been used with previously (based on a review of the literature and interviews with providers). As noted earlier, some of the approaches have developed their own particular Tier I and Tier II programs (e.g., Project Achieve, Safe & Civil Schools), while the TA Center on PBIS providers are purposefully flexible in training schools to implement whichever programs the school has identified to address their needs.

The study team examined the evidence base for Tier I and Tier II programs, details of which are included in the Literature Review Memo (Attachment A). For all but three of the Tier I programs, the study team was able to identify at least one study that met the inclusionary criteria (comparison group, pretest-posttest measurement, and statistical test of significance). Most of these studies were conducted in elementary grades, and most used an experimental design. All such Tier I programs with at least one rigorous study have demonstrated positive behavioral effects but only one (Positive Action) has tested and reported positive academic effects.

All four Tier II–only programs and one program with combined Tier I and Tier II components had at least some documented evidence of effectiveness on behavioral outcomes. None reported intervention effects on academic achievement, although one program reported positive effects on academic behaviors (e.g., completing assignments). Most evaluations were conducted with elementary grade students and used an experimental design.

Exhibit 3. Examples of Tier I and Tier II programs

Examples of relevant programs	Tier I	Tier II	Age/grade	Previous use with SWPBS framework implementation approaches: TA Center on PBIS	Project Achieve	Safe & Civil Schools	Best Behavior
Lions Quest Skills for Adolescence	✓		Grades 6 to 9, ages 11 to 14				
Positive Action	✓		Grades K to 12	✓			
Too Good for Drugs	✓		Grades K to 8	✓			
Too Good for Violence	✓		Grades K to 8	✓			
Incredible Years	✓	✓	1 month to 12 years	✓			
Second Step	✓	✓	PK to middle school, ages 4 to 14	✓			✓
Skillstreaming	✓	✓	Preschool to grade 12	✓			
Stop & Think Social Skills Program	✓	✓	Preschool to grade 8	✓	✓		
Olweus Bullying Prevention Program	✓	✓	Grades K to 9	✓			✓
CHAMPS	✓	✓	Grades K to 12			✓	
Coping Power Program		✓	Grades 4 to 6	✓			
Check & Connect		✓	Grades K to 12	✓			
Check-In/Check-Out		✓	Grades K to 8	✓			
First Steps to Success		✓	Preschool to grade 2	✓			

3.0 TRAINING AND IMPLEMENTATION SUPPORT

In this section, we provide information about various approaches to training and implementation of SWPBS. As we will discuss, there are different approaches and strategies for training school staff to implement the multitiered SWPBS framework. Note that the review of training providers is not exhaustive and is meant only to provide examples of the types of training that are available. For additional details, refer to the Provider Memo (Attachment B). This section addresses research question 3: What is the feasibility of training schools in SWPBS universal supports and targeted interventions?

3.1. Training Providers for the SWPBS Approach and Tier I and Tier II Programs

As described earlier, several approaches are available to schools for implementing the SWPBS framework; this report focused on four examples: TA Center on PBIS, Project Achieve, Safe & Civil Schools, and Best Behavior. Although each has developed specific training for schools, all but the TA Center provide that training through their own organizations. By contrast, training for the TA Center's approach is decentralized and includes at least 10 state-level training providers, each of which offers their own particular approach to delivering training on SWPBS.

Training for specific Tier I or Tier II programs or practices is typically separate from that of the SWPBS framework training due to differences in the school staff they target and the type of training. For example, while training in the SWPBS framework typically involves the school's leadership team, training for specific Tier I or Tier II interventions would target only those teachers or staff implementing the program. In many cases, the tier-specific training is provided by trainers associated with the specific program, and in other cases, this training may be provided by the SWPBS approach provider, particularly when the Tier I or Tier II program is developed in conjunction with the SWPBS approach. An example of the latter is CHAMPS, a Tier I program that is affiliated with Safe & Civil Schools.

3.2. Training in the SWPBS Framework and Tier I/II Support

Training in the SWPBS framework is typically provided to a group of 4-10 school staff who form a leadership team. The team meets monthly or more and is responsible for implementing and training other school staff to implement the SWPBS approach. The leadership team may include:

- an administrator;
- representative grade level teachers;
- the assistant principal or other individual involved in handling discipline referrals;
- personnel in support roles not requiring a teaching or related credential (e.g., transportation, nutrition services);

- a behavioral health specialist (e.g., school counselor, mental health professional) and/or others already involved in providing behavioral supports to at-risk students;
- a parent; and
- a coach (someone who understands the research and practice related to SWPBS and is available to support and monitor the leadership team and ensure faithful implementation of the SWPBS approach).

Not all SWPBS approaches include the training of a leadership team who then trains the entire school staff. Project Achieve, for example, includes the creation of multiple building-level committees (including a leadership team), each of which is focused on a particular aspect of implementation and receives specialized training (e.g., curriculum/instruction, data systems, or parent and community involvement). It is assumed that Project Achieve involves more school staff and more extensive training overall than other approaches, given the need to train multiple subcommittees instead of one school leadership team, as is the case with other approaches.

For approaches focused on training a school leadership team, training provided to the leadership team covers implementation of the SWPBS framework and also includes methods and strategies to support Tier I and Tier II implementation (e.g., defining and teaching positive behaviors at Tier I, and universal screening methods to identify at-risk youth at Tier II). As noted earlier, training for *specific* Tier I or Tier II programs is typically separate and involves those teachers or staff implementing the particular program. In the case of Project Achieve, however, the entire school staff is trained in the Tier I/II program used by that approach (i.e., Stop & Think Social Skills Program).

3.3. Progression of Training in SWPBS

The sequencing and length of time for training in SWPBS varies among the four approaches. The approaches typically involve a 3-year implementation plan (Best Behavior's is slightly shorter, at 2-3 years) with training and ongoing technical assistance and coaching throughout the process. The four approaches require that the school demonstrates readiness to implement, usually prior to implementation; however, in the case of Project Achieve, schools must establish readiness outside of the 3-year plan. Readiness typically involves demonstrating district, principal, and/or staff support for SWPBS; establishing leadership teams or committees; and identifying a method for tracking data.

Training in the implementation of the framework and Tier I typically occurs in the first year for all four approaches, whereas the sequencing of training in Tier II varies by approach. For example, the TA Center typically conducts Tier II training after Tier I fidelity has been met (typically in year 2). Tier II training is also typically completed in year 2 for Best Behavior but not until year 3 for Safe & Civil Schools. For Project Achieve, training in Tier II typically begins in the spring semester of the first year and continues in the second and third years along with Tier III training.

4.0 DESIGN OPTIONS

This section addresses research question 4: What evaluation design options are feasible to study the impacts of SWPBS universal supports and targeted interventions? The section begins with an overview of the rationale for selecting design options. It then presents three design options for an impact evaluation of SWPBS. Finally, the section provides sample size calculations for various scenarios.

4.1. Background/Rationale for Selecting Design Options

Several approaches for implementing the SWPBS framework—including the TA Center for PBIS, Project Achieve, Safe & Civil Schools, and Best Behavior—have been evaluated. Results of these evaluations have been mixed but are generally positive, with most evaluations involving 10–37 schools and conducted in one or a limited number of school districts. The findings of the literature review supporting this design report (see Attachment A) suggest there is sufficient evidence from local evaluation studies to warrant a larger evaluation of the SWPBS framework.

There are various approaches to training schools in the implementation of the SWPBS framework, and there are several key aspects of SWPBS training that differ across training providers. However, studies have neither compared different approaches to training schools in SWPBS against each other nor compared key aspects of the training to see what type of training is most effective.

In some common training models, training in extending the SWPBS framework through Tier II occurs after the framework (including Tier I) is implemented. Schools are sometimes given the flexibility to select the Tier II intervention(s) that match their needs. Although evaluations of specific interventions for at-risk students have shown positive results across academic outcomes, externalizing behavior, and internalizing behavior, no studies have evaluated the impact of training in Tier II within the context of SWPBS implementation.

4.2. Study Design

The study design team considered three design options, each tailored to a specific evaluation question (see exhibit 4). All of the study design options are in the class of experimental designs known as group-randomized trials; these designs involve random assignment of schools to study condition with at least some program outcomes measured at the individual level. These designs use controlled application of an intervention or treatment and randomization of schools to study condition to provide evidence of program effect. They offer strong internal validity, but they may not provide external validity (i.e., generalization) if the set

Exhibit 4. Design options and hypotheses for an evaluation of SWPBS

Design option	Treatment	Counterfactual	Alternative counterfactual
Option 1: Treatment-Control Design	SWPBS Approach + Tier 1 + Tier 2 (μ_A)	Business as Usual (μ_B)	
Evaluation Hypotheses		$H_1 : \mu_A \neq \mu_B$	
Option 2: Constructive Treatment Design	SWPBS Approach + Tier 1 +Tier 2 (μ_A)	SWPBS Approach + Tier 1 (μ_B)	Business as Usual (μ_C)
Evaluation Hypotheses		$H_1 : \mu_A \neq \mu_B$	$H_2 : \mu_A \neq \mu_c$ $H_3 : \mu_B \neq \mu_c$
Option 3: Comparative Treatment Design	SWPBS Approach(A) + Tier 1 + Tier 2 (μ_A)	SWPBS Approach (B) + Tier 1 + Tier 2 (μ_B)	Business as Usual (μ_C)
Evaluation Hypotheses		$H_1 : \mu_A \neq \mu_B$	$H_2 : \mu_A \neq \mu_c$ $H_3 : \mu_B \neq \mu_c$

of schools is selected to be homogeneous in order to maximize statistical power. Group-Randomized Trials, like all randomized experiments, provide a basis for excluding plausible alternative explanations of program effects.

Group-randomized trials conducted in school settings also have characteristics of field studies that must not be ignored. Unlike highly controlled laboratory studies, developers of group-randomized trials must be prepared to account for variations in school settings and in program implementation and fidelity, which may influence program effects.

Implementing studies that randomize schools to experimental conditions presents some challenges. This method requires a large population of schools to serve as the sampling frame. Some school officials may not want to participate in a randomized evaluation if their schools may be placed in a control group in which they do not receive the benefit of the intervention. Others may want to ensure that everyone in the neediest classrooms or schools receives the intervention; this is a form of selection bias that would limit the utility of the study's findings.

4.2.1. Design Option 1: Treatment-Control Design

Design option 1 addresses the research question, "What is the impact of training schools in the implementation of the SWPBS approach, including Tier I supports and Tier II interventions?" This design option employs a group-randomized trial that compares SWBPS

training in a selected group of schools to business as usual in a second group of schools. Under this design, schools are recruited to participate and are randomly assigned to receive SWPBS training and support or to a business as usual (i.e., control) condition. There are a range of potential treatments because there are different implementation approaches and various SWBPS Tier I support programs and Tier II interventions. Under design option 1, the business as usual condition serves as the counterfactual comparison. The primary evaluation hypothesis (H_1) asserts that measured behavioral and academic outcomes will differ between schools that receive SWPBS training and support and schools conducting business as usual.

Although most studies to date on the efficacy of SWPBS programs have used "business as usual" comparisons, no large-scale (e.g., multistate, multidistrict) evaluations of SWPBS have been conducted. Most studies have taken place in one locale and within a limited number of school districts. This design affords an opportunity to build directly from the established literature base and expand our understanding of the SWPBS approach by evaluating its ability to train and implement programs on a large scale. Assuming a recruitment process that includes schools and school districts from different parts of the country, this design also allows the assessment of how regional variation and contextual influence serve as barriers or facilitators to successful program outcomes. Primary challenges to successful evaluation of this design include recruiting schools for a randomized study, retaining schools assigned to the control group, heterogeneity in the operationalization of business as usual among schools in the control condition, and monitoring the implementation and effects of SWPBS-like programs that may arise over the course of the study period.

4.2.2. Design Option 2: Constructive Treatment Design

Design option 2 addresses the research question, "What is the benefit of training schools in targeted interventions (i.e., Tier II) once the SWPBS framework and universal supports are in place?" This design option employs a group-randomized trial that randomly assigns schools to (group A) receive training and support in an SWPBS approach that includes universal supports (Tier I) and targeted interventions (Tier II) or to (group B) receive training and support in an SWPBS approach limited to universal supports (Tier I). The constructive treatment design assumes both groups receive training and support in the same SWPBS approach and the schools limited to universal support (group B) serve as the counterfactual. The primary evaluation hypothesis (H_1) asserts that measured behavioral and academic outcomes will differ between schools that receive SWPBS training and support that include both universal supports (Tier I) and targeted interventions (Tier II) and schools that receive SWPBS training and support that only include universal supports (Tier I). Additionally, a third condition representing business as usual (i.e., control condition) can be added and schools can be randomized across the three study

conditions. The business as usual condition (group C) can serve as the counterfactual for both treatment conditions, providing two alternative evaluation hypotheses: (H_2) asserting that measured behavioral and academic outcomes will differ between schools that receive SWPBS training and support that includes universal supports and targeted interventions (group A) and schools conducting business as usual; and (H_3) asserting that measured behavioral and academic outcomes will differ between schools that receive SWPBS training and support that includes universal supports (group B) and schools conducting business as usual.

Targeted interventions involve supplemental progress monitoring for at-risk students and developing systems for increasing the intensity and focus on behavioral supports. Theoretically, this approach is supported by the SWPBS logic model but has yet to be evaluated empirically. The literature review conducted to support this report (see Attachment A), reported no studies that examined the benefits of adding targeted intervention to an existing program offering SWPBS framework and universal supports. Primary challenges to successful evaluation of this design include a priori estimation of anticipated program effects with little guidance available on the additive benefits of targeted interventions.

4.2.3. Design Option 3: Comparative Treatments Design

Design option 3 addresses the research question, "What is the impact of training schools in the implementation of one SWPBS approach vs. a different SWPBS approach?" This design option employs a group-randomized trial that compares two fully implemented SWPBS approaches. Under this design, schools are recruited to participate and randomly assigned to receive training and support in one SWPBS approach (group A) or to receive training and support in an alternative SWPBS approach (group B). This design could involve comparing two different approaches to training schools in SWPBS or varying the training on a key dimension of training. The comparative treatment design does not provide a traditional counterfactual. The primary evaluation hypothesis (H_1) asserts that measured behavioral and academic outcomes will differ between schools that receive SWPBS training and support (group A) and schools that receive the alternative SWPBS training and support (group B). Identification of treatment impacts will be expressed as the benefit of one program over the other. Based on a substantive understanding of selected SWBPS approaches, a priori specification of one treatment as the indicated treatment and the second treatment as the referent may be warranted. A third condition representing business as usual (group C) can be added, and schools can be randomized across the three study conditions. The business as usual condition can serve as the counterfactual for both treatment conditions, providing two alternative evaluation hypotheses: (H_2) asserting that measured behavioral and academic outcomes will differ between schools that receive SWPBS training and support (group A) and schools conducting business as usual; and (H_3) asserting that

measured behavioral and academic outcomes will differ between schools that receive SWPBS training and support (group B) and schools conducting business as usual.

This design presents potentially important opportunities to assess how differences in training and implementation influence program outcomes at the student-level and school-level. There have been no studies that report the comparative effectiveness of two SWPBS treatments. Primary challenges to successful evaluation of this design include a priori estimation of anticipated program effects with little guidance available on the comparative effectiveness of different SWPBS approaches. Additionally, minimum detectable effects are likely to be smaller, requiring substantially larger number of schools.

4.3. Sample Size Options

Statistical power calculations guide a number of key decisions in the design and execution of any research study. Power is the probability of observing a statistically significant difference, where such a difference exists. In other words, power indicates the probability that an intervention effect will be judged to be statistically significant, given the assumptions of the specified model. Statistical power is a function of sample size, the assumed magnitude of the intervention impact, the form of the statistical model, and the anticipated level of random variation in the measure of the outcome. In the following, we provide information from the published literature on the assumed magnitude of the intervention impact, the form of the statistical model, and the anticipated level of random variation in the measure of the outcome to determine the minimum sufficient sample size for study planning purposes.

4.3.1. Minimum Detectable Effects

The assumed magnitude of the intervention impact or minimum detectable effect (MDE) is typically obtained from effect size (ES) estimates reported in the published literature. For example, studies of SWPBS approaches have reported ESs of 0.12–0.17 for teacher reports of student disruptive behavior and ESs of 0.08 to 0.11 for measures of emotional regulations in elementary school settings (Bradshaw et al. 2010a; Bradshaw et al. 2012; Waasdorp et al. 2012). These studies also suggest ESs in the range of 0.12–0. 27 for school-level outcomes such as office referrals and suspensions. A review of Positive Action, a Tier I program, reported ESs ranging from 0.23 for tobacco use to 0.93 for binge drinking, with an average ES of around 0.50 for problem behaviors (Flay and Allred, 2003; Flay et al. 2006). Reports on the effectiveness of Positive Action also indicate ESs for academic outcomes in reading and math of 0.73 and 0.34, respectively (Flay et al. 2006). Studies of SWPBS approaches reported improvements in academic outcomes equivalent to an ES of 0.54 for math; however, this trend did not reach an accepted level of statistical significance (Bradshaw et al. 2010a). A review of First Steps to Success, a Tier II program, reported ESs of 0.76 and 0.62 for external (problem) behaviors and

social outcomes, respectively; this review also noted an ES of 0.32 for academic outcomes (Walker et al. 1998; Walker et al. 2009). The ESs found in the review of First Steps are slightly larger than the finding reported by Mytton and colleagues (2006). These researchers conducted a systematic review of the literature on secondary prevention programs and reported ESs of 0.36 to 0.42 for aggressive and violent behavior in primary school settings (Mytton et al. 2006). While most ES data is reported for student-related outcomes, ES information on teacher's classroom management is also worth consideration. One study reported findings from The Incredible Years, another Tier II program, on a composite measure of classroom management. Data from this study indicate ESs ranging from .35 to .63 for treatment groups compared to a control condition (Webster-Stratton et al. 2004). The reported ESs are from studies that compared SWBPS interventions to business as usual comparison conditions. Accordingly, they are reasonable estimates for H_1 under design option 1 and for H_2 and H_3 under design options 2 and 3. Additionally, these estimates are from well controlled trials in a small number of geographies. Moving to a large-scale evaluation would increase program complexity and sampling heterogeneity which could lead to smaller program impacts. There are no available estimates from studies directly comparing SWPBS approaches to guide study planning. It is reasonable to assume that program impacts obtained by directly comparing competing interventions (i.e., H_1 under options 2 and 3) will be smaller.

4.3.2. Intraclass Correlation Coefficients

In a group-randomized trial, schools are the unit of assignment and are nested in study condition; observations are taken in the individuals nested within schools. Repeated observations on the same individuals can lead to additional levels of nesting. This hierarchical arrangement leads to correlation in the structure of the data. The index of this correlation, the intraclass correlation coefficient (ICC), indicates the proportion of variation in the response of a given individual that is associated with group membership. Unless the extra variation is accounted for in study planning, the study is likely to be underpowered (Murray and Hannan 1990; Zucker 1990). For the purposes of estimating sample size, we employ small ICCs ranging from 0.05 to 0.30. In general, student self-reported measures tend to be associated with smaller ICCs while measures of academic achievement, school climate, and teacher-reported behavioral observations tend to be associated with larger ICCs. Exhibit 5 provides ICCs from the published literature.

Exhibit 5. Intraclass Correlation Coefficients (ICCs) for teacher- or student-reported outcomes for students

Source	Outcome domain	ICC	Level of clustering	Grade level/school type
Teacher	Social contact[1]	0.33	Classroom	3rd grade
	Emotional regulation and social aggression[1,2]	0.13–0.30	Classroom	3rd grade
	School climate[3]	0.35	School	5th grade
	Academic emphasis[3]	0.54	School	5th grade
Student	School climate[3]	0.28	Classroom	5th grade
	Academic emphasis[3]	0.12	Classroom	5th grade
	Violence (non-weapon)[4]	0.01	School	Middle Schools
	Relational victimization[4]	0.01	School	Middle Schools
	Math achievement[5]	0.21–0.24	School	3rd grade
	Reading achievement[5]	0.22–0.27	School	3rd grade
	Bullying (victimization)[6]	0.02	School	Primary schools
	Bullying (perpetration)[6]	0.04–0.08	School	Primary schools
	Violent behavior[7]	0.02	School	Middle schools

[1]Fraser et al. 2005.
[2]Emotional regulation includes dimensions from the TOCA-R including cognitive concentration, authority acceptance, and social competence.
[3]Mitchell et al. 2010.
[4]Silvia et al. 2010.
[5]Hedges and Hedberg 2007.
[6]Shaw and Cross 2012.
[7]Janega et al. 2004.

4.3.3. Sample Size Estimates

In Equation 1, derived from Murray (1998), $\hat{\sigma}_y^2$ represents the estimated variance of the outcome and the ICC is combined with $\hat{\sigma}_y^2$ to determine within-group and between-group components of variation.

$$\left(\hat{t}_{\beta-1}+\hat{t}_{\alpha/2}\right)^2 \frac{4T\left[\left(\hat{\sigma}_y^2\left(1-\hat{ICC}_{m:g:c}\right)\left(1-\hat{r}_{yy(m)}\right)\hat{\Theta}_m + m\left(\hat{\sigma}_y^2\left(\hat{ICC}_{m:g:c}\right)\right)\left(1-\hat{r}_{yy(g)}\right)\hat{\Theta}_g\right)\right]}{m\hat{\Delta}^2} \quad (1)$$

The expression $\left(\hat{t}_{\beta-1}+\hat{t}_{\alpha/2}\right)$ represents the critical values from the *t*-distribution. Here, *T* is a multiplier that indicates the number of time points included in the analysis, *m* represents the number of observations (e.g., children) per school, and $\hat{\Delta}$ is the estimated program impact. Equation 1 also includes parameters reflecting the statistical model that will be used to control measured variation and increase statistical precision. Specifically, Θ_m and Θ_g, representing a reduction in the member and group variances due to the inclusion of covariates. Additionally,

$r_{yy(m)}$ and $r_{yy(g)}$ represent the reduction in variation associated with repeated measures on the same persons and groups, respectively. Exhibit 6 provides sample size estimates that determine the number of schools needed to achieve a type-II error rate of 0.20 (i.e., 80% statistical power) under two modeling scenarios:

- Scenario A: Assumes the use of a nested cross-sectional model, with repeated measures on schools and a refreshed sample of 300 students at each data collection period. This scenario would be appropriate for estimating the population average effects of SWPBS.
- Scenario B: Assumes the use of a nested cohort model, with repeated measures on schools and on 100 students tracked across grades within schools over the study period. This model would be appropriate for assessing how individuals change over time as a result of exposure to SWBPS.

The sample size estimates in Exhibit 6 assume intervention impacts will be modeled as a two condition comparison and assessed with a one degree of freedom test statistic. These estimates also assume a balanced and homoscedastic two condition design so that for any MDE and ICC, the tabled number of schools (g) implies $g/2$ schools per condition. Additional considerations are necessary for designs that include more than two conditions. Consider, for example, the case where two independent approaches to training in SWBPS are applied, each to a set of schools and another set of schools serves as a business as usual control condition. Based on the assumptions of the design and analytic framework, one could propose assigning $g/2$ schools to each condition, leading to a study that includes a total of $3g/2$ schools. Under a different interpretation of this scenario, one could propose combining the schools from the two treatment conditions and comparing the pooled mean change to the mean change among the schools in the business as usual condition. This study could be executed and evaluated by randomly assigning $g/2$ schools to the business as usual condition and $g/4$ schools to each of the treatment conditions, resulting in a balanced design for the primary hypothesis. More elaborate design options are feasible. If design options include multiple, related hypotheses, additional consideration must be given to methods for controlling the false discovery rate. Alternatively, designs may be predicated upon randomly assigning schools to condition based upon a priori expectations for intervention impacts of varying magnitudes (i.e., assigning treatments assumed to produce smaller ESs to more schools). If imbalance is introduced as a characteristic of the study design, the impact of heteroscedasticity on statistical power should be considered.

Exhibit 6. SWPBS sample size requirements

Minimum detectable effect (MDE)	Intraclass correlation (ICC)	Number of schools needed	
		Scenario A	Scenario B
0.10	0.05	92	88
	0.10	164	160
	0.15	236	232
	0.20	308	302
	0.25	378	374
	0.30	450	446
0.15	0.05	42	40
	0.10	74	72
	0.15	106	104
	0.20	138	136
	0.25	168	168
	0.30	200	198
0.20	0.05	24	22
	0.10	42	40
	0.15	60	58
	0.20	78	76
	0.25	96	94
	0.30	114	112
0.25	0.05	16	14
	0.10	28	26
	0.15	38	38
	0.20	50	50
	0.25	62	60
	0.30	72	72
0.30	0.05	12	10
	0.10	20	18
	0.15	28	26
	0.20	36	34
	0.25	42	42
	0.30	50	50

NOTE: The numbers in the table represent the total number of schools needed to identify the minimum detectable effects specified in the first column assuming a balanced design. Under both modeling assumptions, T = 2 points in time, for a two-tail, null hypothesis test comparing two conditions, with a type-I error rate of 0.05. Both scenarios include reduction in variance at the school level (0.25) and student level (0.20) due to the inclusion of covariates, as well as reduction in variance associated with repeated measures at the school level (0.75). In addition, Scenario B includes reduction in variance associated with repeated measures at the student level (0.75).

5.0 OTHER DESIGN CONSIDERATIONS

This section provides an overview of additional issues to consider when designing an impact evaluation of SWPBS. First, the section provides considerations for selecting the school/grade level for the study. It then describes methods for assessing: (1) implementation fidelity; (2) impact outcomes; and (3) variations in the impacts across subgroups.

5.1. School/Grade Level of Implementation

A critical decision for an impact evaluation of SWPBS is the school or grade level at which to implement the treatment. Although the design options presented in section 4 could be implemented at any K–12 grade level, it is reasonable to consider narrowing the range of grade levels for both cost and feasibility considerations. One factor to consider in selecting the school/grade level in which to evaluate is the grade levels for which the SWPBS approach and programs were designed. There are various programs available that were designed for various school levels. The provider information assembled through this design study indicates that all of the SWPBS approaches reviewed are designed for K–12 implementation. At Tier I, the majority of programs are designed for a combination of grade levels spanning elementary through high school grades. At Tier II, about half of the interventions are designed for elementary school only, whereas others can be implemented with combined grade levels K–12. A second consideration is the prevalence of implementation—that is, at what school/grade levels SWPBS is most often implemented across schools and districts. A scan of the literature indicates that the most prevalent level for implementation of SWPBS is elementary school. With regards to evaluations of SWPBS approaches, we found that nearly all have been conducted in elementary schools. Similarly, evaluations that focus on particular Tier I or Tier II programs are predominantly at the elementary school level. Thus, much of the knowledge around implementation of SWPBS and much of the evidence base are at the elementary school level, with significantly less focus on middle and high school levels.

A key question for consideration is whether to build a large-scale evaluation at the school/grade level where there is the best understanding and likelihood (based on previous work) that significant effects will be found, or to attempt to develop this understanding where there is less information, in this case at the middle or high school level. For a variety of reasons, including best use of resources, IES has consistently supported the view that a large-scale evaluation should be conducted only when there is able knowledge.

5.2. Implementation Fidelity

Documentation of fidelity of implementation will be critical for any impact evaluation of SWPBS due to the wide flexibility that schools have in how they adopt SWPBS strategies to their own school culture. Among the four approaches, the TA Center in particular appears to

emphasize that the SWPBS approach should be ever evolving and improving as it needs to be constantly adapted to fit the unique context of different schools systems. In this section, we first address which aspects of implementation may be considered most important for measurement and then provide an overview of available measures and tools.

5.2.1. Aspects of Implementation Fidelity to Measure

Within the context of SWPBS, it will be important to measure:

- Readiness to implement SWPBS—for example, demonstrating district, principal, and/or staff support for SWPBS; establishing functional leadership teams or committees; and identifying a method for tracking data
- Adherence to guidelines for best practice of the approach or program—for example, the extent to which schools collect and review data, have teams that meet monthly, or use rewards consistently
- Training sequence and scope for each tier (based on the specific approach and Tier I/II programs)
- Extent of training (e.g., number of training and coaching sessions; documentation of training content)
- Implementation of treatment as planned (e.g., staff self-reports of adherence; observer reports of adherence, including scheduled observations and spot checks; monthly progress by leadership team; and extent of and participation in coaching and training)
- How the implementation of behavior supports differs in treatment and control schools (impacts on practice)

5.2.2. Available Measures and Tools

A variety of measures and tools are available to document implementation of SWPBS and measure fidelity. Several of the implementation approaches to SWPBS included in this design study have also developed their own fidelity instruments. A list of sample fidelity instruments for Tier I and Tier II is presented in exhibit 7.

Exhibit 7. Sample fidelity instruments for Tier I and Tier II

Tier I	Tier II
• TA Center on PBIS – Benchmarks of Quality (BOQ) – Schoolwide Evaluation Tool (SET). – Team Implementation Checklist (TIC) – School Assessment Survey (SAS) – PIC (Positive Behavior Intervention Checklist) – Phases of Implementation (POI)	• TA Center on PBIS – Benchmarks for Advanced Tiers (BAT) – Tier II Interventions Assessment Tools – Monitoring of Advanced Tiers Tool (MATT) – Phases of Implementation (POI) – Tier II/Tier III Intervention Tracking Form – Tier II Action Plan Checklist – Individual Student Systems Evaluation Tool (ISSET)
• Best Behavior – Best Behavior Self-Assessment Survey (fidelity checklist created by Tanglewood Research) – Prevention Practices Assessment (PP-A) – Prevention Practices Survey (PP-S)	• Best Behavior – Assessment of Services for At-Risk Students (ARSSA)
• Project Achieve – PRAIISE (Project ACHIEVE Implementation Integrity Self-Evaluation)	• Project Achieve – PRAIISE (Project ACHIEVE Implementation Integrity Self-Evaluation)
• Safe & Civil Schools – Foundations Checklist & Implementation Rubric – Foundations Surveys (measure perceptions of staff, students, and parents regarding safety, climate, and discipline) – Protocol for Observing Common Areas (e.g., halls, playground, and cafeteria)	• Safe & Civil Schools – Foundations Checklist & Implementation Rubric – Interventions Checklist

5.3. Impact Outcomes

Outcomes of potential interest to an impact evaluation of SWPBS may be organized around five major areas:

- Student behavior outcomes
- Student social competency skills
- Student academic outcomes
- Teacher performance outcomes
- School climate

In the context of SWPBS implementation, **student behavior** is commonly tracked using Office Discipline Referrals (ODRs) and typically includes both major (e.g., fighting, abusive language) and minor infractions (e.g., disruption, defiance). These data are collected using either an existing school data system or one that is implemented across all participating schools. Individual student records for behavior and data on suspensions and expulsions may be obtained through the district or school office. School-level suspension/expulsion rates may also be available in states that have state-level reporting systems. Self-report surveys are another source of behavior data, although they are only appropriate beginning at the higher elementary grades. With younger students, a more appropriate method of collecting behavior data is through teacher ratings/reports of behavior. Still other methods for measuring behavior include peer-nomination procedures, which rely on students' ratings of their peers, and direct observation by researchers. Teacher and student surveys, teacher ratings, peer nominations, and direct observation methods can also be used to measure behaviors not tracked by the school such as social aggression (e.g., excluding others, spreading rumors) and prosocial behaviors (e.g., helpfulness, sharing, cooperation, empathy). Adequately capturing student behavioral outcomes will likely require a combination of multiple data collection methods and data sources.

Student social competency indicators that may be of interest include communication skills, social problem solving, conflict resolution skills, peer interaction skills, and the prosocial behaviors described previously (e.g., helpfulness, sharing, cooperation, empathy). Social competence may be measured in ways similar to student behavior including student surveys, teacher ratings, peer nominations, and independent observations.

Two types of **student academic measures** are typically available through school records: school performance measures (e.g., academic grades, end-of-grade tests or state tests, achievement tests) and school academic behaviors (e.g., tardiness, attendance).

Teacher performance outcomes include teacher practices, classroom management, teacher-child interactions, and student engagement and other behaviors. Classroom observations are particularly helpful for measuring these teacher performance outcomes in the context of the classroom environment.

School climate measures include those addressing feelings of safety, perception of fairness, feelings of inclusiveness and connectedness, opportunities for parent involvement, and perception of student support. Accurate assessment of school climate requires input from multiple members of the school community, including students, teachers, administrators and other staff, and parents. These data can be collected through self-report surveys, personal interviews, or focus groups. Certain school-level data can serve as additional sources of data on school climate such as attendance rates, teacher turnover rate, principal tenure length, academic sanctions, and rates of suspension/expulsion.

5.4. Variations Across Subgroups

A number of student subgroups may be of interest for targeted analyses in an impact evaluation of SWPBS. Subgroup comparisons would provide information about whether or not SWPBS is more or less effective for different groups of students. In some cases, there is evidence that the treatment may affect various groups differently. In other cases, it is particularly important to understand the effect of the treatment on certain subgroups. Potential student subgroups of interest include:

- Racial/ethnic groups—Subgroup comparisons would examine the potential for differences in impacts among racial/ethnic groups. Subgroup analyses may also provide pertinent information related to the issue of overidentification of students for special education based on race and ethnicity, that is the result of inappropriate identification. There are data indicating that minority students are overrepresented in office discipline referrals; suspensions and expulsions; and special education services (Kaufman et al. 2010; Raffaele-Mendez and Knoff 2003; Vincent and Tobin 2011). IDEA requires states and districts to take steps to address such inequities with respect to special education and to examine data to determine if disproportionality exists based on race/ethnicity with respect to special education services, placement in particular education settings, and incidence of suspensions and expulsions. SWPBS supports culturally responsive practices to provide opportunities for all students and there is some evidence that SWPBS may help reduce disciplinary inequity across students from different racial/ethnic backgrounds (Bradshaw et al. 2010b).
- Gender—Differences in the types of problem behaviors exhibited by boys and girls have been documented in the literature (Orpinas and Horne 2006). Other research indicates that prevention programs may impact girls and boys differently (Farrell and Meyer 1997). It will be important to understand if SWPBS interventions and supports are equally effective for all students in a school, both boys and girls.
- Students identified for special education—SWBPS as supported by IDEA is intended to address the needs of special education students within the context of the larger school population. Subgroup analyses would provide pertinent information about outcomes for special education students.
- Students identified as at risk for aggressive behaviors—Subgroup measurement and analysis would provide information about the impact of SWPBS on students already exhibiting or at high risk for problem behaviors. It will be useful to understand whether universal (Tier I) supports, which are implemented schoolwide, have an impact on at-risk students as well as the general student population. In addition, it will be important to examine the impact of Tier II programs (e.g., using design option 2;

see section 4) on at-risk students. By design, these programs provide more targeted services and are intended for students not responding to Tier I supports.

In addition to student subgroups, an impact evaluation of SWPBS could also consider subgroups among school staff, in particular those staff responsible for implementing SWPBS programs and interventions at the classroom level. For example, one could examine differences between groups of teachers with more or less classroom teaching experience on the quality of implementation of the classroom-based program and outcomes for students. Another teacher subgroup analysis could measure differences in student/classroom outcomes among teachers with different levels of prior training (including no training) in classroom management/behavior interventions. These subgroup analyses would be helpful in understanding whether teacher characteristics mediate the effects of SWPBS—that is, whether effects vary by teaching experience or specialized training in classroom management, for example.

6.0 REFERENCES

Arnold, D.H. (1997). Co-occurrence of externalizing behavior problems and emergent academic difficulties in young high-risk boys: A preliminary evaluation of patterns and mechanisms. *Journal of Applied Developmental Psychology, 18*: 317-330.

Bradley, M.C., Daley, T., Levin, M., O'Reilly, R., Parsad, A., Robertson, A., and Werner, A. (2011). *IDEA National Assessment Implementation Study* (NCEE 2011-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U. S. Department of Education.

Bradshaw, C.P., Mitchell, M.M., and Leaf, P.J. (2010a). Examining the effects of Schoolwide Positive Behavioral Interventions and Supports on student outcome: Results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions, 12*: 133-148. doi: 10.1177/1098300709334798.

Bradshaw, C.P., Mitchell, M.M., O'Brennan, L.M., and Leaf, P.J. (2010b). Multilevel exploration of factors contributing to the overrepresentation of black students in office discipline referrals. *Journal of Educational Psychology, 102*(2): 508-520.

Bradshaw, C.P., Waasdorp, T.E., and Leaf, P.J. (2012). Effects of school-wide positive behavioral interventions and supports on child behavior problems. Retrieved from http://www.pediatrics.aappublications.org doi: 10.1542/peds.2012-0243.

Darney, D., Reinke, W.M., Herman, H.C., Stormont, M., and Ialongo, N.S. (2013). Children with co-occurring academic and behavior problems in first grade: Distal outcomes in twelfth grade. *Journal of School Psychology, 51*: 117-128. doi: 10.1016/j.jsp.2012.09.005.

Dryfoos, J.D. (1990). *Adolescents at Risk: Prevalence and Prevention.* New York, Oxford University Press.

Farrell, A.D., and Meyer, A.L. (1997). The effectiveness of a school-based curriculum for reducing violence among urban sixth-grade students. *American Journal of Public Health, 87*(6): 979-984.

Flay, B.R., and Allred, C.G. (2003). Long-term effects of the Positive Action program. *American Journal of Healthy Behavior, 27*(1): 6-21.

Flay, B., Acock, A., Vuchinich, S., and Beets, M. (2006). Progress report of the randomized trial of Positive Action in Hawaii: End of third year of intervention. Available from http://www.positiveaction.net/content/PDFs/research/3rd_Year_Progress_Report_with_Academic_and_Behavior_Results.pdf

Fraser, M.W., Galinsky, M.J., Smokowski, P.R., Day, S.H., Terzian, M.A., Rose, R.A., and Guo, S. (2005). School information processing skills training to promote social competence and prevent aggressive behavior in the third grade. *Journal of Clinical and Consulting Psychology, 73*(6): 1045-1055.

Harding, M., Knoff, H.M., Glenn, R., Johnson, L., Schrag, H., and Schrag, J. (2008). *The Arkansas State Improvement Grant evaluation and outcome report to the U.S. Department of Education's Office of Special Education Programs: Improving student outcomes through the school-wide implementation of Project ACHIEVE's Positive Behavioral Support Systems*. Little Rock, AR: Arkansas Department of Education, Special Education.

Hawkins, J.D., and Weis, J.G. (1985). The social development model: an integrated approach to delinquency prevention. *Journal of Primary Prevention, 6*: 73-97.

Hawkins, J.D., and Catalano, R.F. (1990). Broadening the vision of education: Schools as health promoting environments. *Journal of School Health, 60*(4): 178-181.

Hedges, L.V., and Hedberg, E.C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis, 29*(1): 60-87.

Horner, R.H., Sugai, G., and Anderson, C.M. (2010). Examining the evidence base for schoolwide positive behavior support. *Focus on Exceptional Children, 42*: 1-14.

Horner, R.H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A.W., and Esperanza, J. (2009). A randomized, wait-list controlled effectiveness trial assessing school-wide Positive Behavior Support in elementary schools. *Journal of Positive Behavior Interventions, 11*: 133-144. doi: 10.1177/ 1098300709332067.

Janega, J.B., Murray, D.M., Varnell, S.P., Blitstein, J.L., Birnbaum, A.S., and Lytle, L.A. (2004). Assessing the most powerful analysis method for schools intervention studies with alcohol, tobacco, and other drug outcomes. *Addictive Behaviors, 29*(3): 595-606.

Kaufman, J.S., Jaser, S.S., Vaughan, E.L., Reynolds, J.S., Di Donato, J., Bernard, S.N., and Hernandez-Brereton, M. (2010). Patterns in office discipline referral data by grade, race/ethnicity, and gender. *Journal of Positive Behavior Interventions, 12*, 44-54.

Kellam, S., Rebok, G., Ialongo, N., and Mayer, L. (1994). The course and malleability of aggressive behavior from early first grade into middle school: Results of a developmental epidemiologically-based prevention trial. *Journal of Child Psychology and Psychiatry, 35*(2): 259-281.

Madigan, K., and Cross, R. (2009). *Determining the Impact of Safe & Civil Schools' School-Wide Positive Behavior Supports Program on Academic Achievement: A Nine-Year Study*. Bethesda, MD: Accountability Works.

Mitchell, M.M., Bradshaw, C.P., and Leaf, P.J. (2010). Student and teacher perceptions of school climate: A multilevel exploration of patterns of delinquency. *Journal of School Health, 80*(6): 271-279.Moffitt, T.E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100*: 674-701.

Murray, D.M. (1998). *Design and Analysis of Group-Randomized Trials*. New York: Oxford University Press.

Murray, D.M., and Hannan, P.J. (1990). Planning for the appropriate analysis in school based drug-use studies. *Journal of Consulting and Clinical Psychology, 58*(4): 458-468.

Mytton, J.A., DiGuiseppi, C., Gough, D., Taylor, R.S., and Logan, S. (2006). School-based secondary prevention programmes for preventing violence. *Cochrane Database of Systematic Reviews, 3*. Art. No.: CD004606. DOI: 10.1002/14651858.CD004606.pub2.

Orpinas, P., and Horne, A.M. (2006). *Bullying Prevention: Creating a Positive School Climate and Developing Social Competence*. Washington, DC: American Psychological Association.

Raffaele Mendez, L.M., and Knoff, H.M. (2003). Who gets suspended from school and why: A demographic analysis of schools and disciplinary infractions in a large school district. *Education and Treatment of Children, 26*, 30-51.

Reinke, W.M., Herman, K.C., Petras, H., and Ialongo, N.S. (2008). Empirically-derived subtypes of child academic and behavior problems: Co-occurrence and distal outcomes. *Journal of Abnormal Child Psychology, 36*: 759-770. doi: 10.1007/s10802-007-9208-2.

Robers, S., Zhang, J., and Truman, J. (2012). *Indicators of School Crime and Safety: 2011* (NCES 2012-002/NCJ 236021). Washington, DC: National Center for Education Statistics, U.S. Department of Education, and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice.

Schaeffer, C.M., Petras, H., Ialongo, N., Poduska,, J., and Kellam, S. (2004). Modeling growth in boys' aggressive behavior across elementary school: Links to later criminal involvement, conduct disorder, and antisocial personality disorder. *Developmental Psychology, 39*: 1020-1035. doi: 10.1037/0012-1649.39.6.1020.

Schaps, E., and Battistich, V. (1991). Promoting health development through school-based prevention: new approaches. In Goplerud, E. N. (ed.), *Preventing Adolescent Drug Use: From Theory to Practice. Office for Substance Abuse Prevention Monograph 8*, USDHHS, PHS, ADAMHA, Rockville, MD, pp. 127–181.

Shaw, T., and Cross, D. (2012). The clustering of bullying and cyberbullying behaviour within Australian schools. *Australian Journal of Education, 56*(2): 142-162.

Silver, R.B., Measelle, J.R., Armstrong, J.M., & Essex, M.J. (2010). The impact of parents, child care providers, teachers, and peers on early externalizing trajectories. *Journal of School Psychology, 48*: 555-583. doi:10.1016/j.jsp.2010.08.003.

Silvia, S., Blitstein, J., and Williams, J. (2010). *Impacts of a Violence Prevention Program for Middle Schools: Findings After Three Years of Implementation*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Silvia, E., Blitstein, J.L., Williams, J., Ringwalt, C., Dusenbury, L., and Hansen, W. (2011, May). *Impacts of a Violence Prevention Program for Middle Schools: Findings After 3 Years of Implementation* (NCEE 2011-4017). Prepared for National Center for Education Evaluation

and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Research Triangle Park, NC: RTI International.

Simons-Morton, B.G., Crump, A.D., Haynie, D.L., and Saylor, K.E. (1999). Student-school bonding and adolescent problem behavior. *Health Education Research, 14*(1): 99-107. doi:10.1093/her/14.1.99.

Sugai, G. (2007). *SWPBS: Leadership Team Guidelines* [PowerPoint slides]. Retrieved March 28, 2013, from http://www.pbis.org/presentations/default.aspx.

Tremblay, R.E., Masse, B., Perron, D., Leblanc, M., Schwartzman, A.E., & Ledingham, J.E. (1992). Early disruptive behavior, poor school achievement, delinquent behavior, and delinquent personality: Longitudinal analyses. *Journal of Consulting and Clinical Psychology, 60*: 64-72.

Vincent, C.G., and Tobin, T.J. (2011). An examination of the relationship between implementation of school-wide positive behavior support (SWPBS) and exclusion of students from various ethnic backgrounds with and without disabilities. *Journal of Emotional and Behavioral Disorders*, *19*: 217-232.

Waasdorp, T.E., Bradshaw, C.P., & Leaf, P.J. (2012). The impact of School-wide Positive Behavioral Interventions and Supports (SWPBIS) on bullying and peer rejection: A randomized controlled effectiveness trial. *Archives of Pediatrics and Adolescent Medicine, 116*: 149-156.

Wagner, M. (1995). Outcomes for youths with serious emotional disturbance in secondary school and early adulthood. *The Future of Children, 5*: 90-112.

Walker, H., Kavanagh, K., Stiller, B., Golly, A., Severson, H., & Feil, E. (1998). First Step to Success. An early intervention approach for preventing school antisocial behavior. *Journal of Emotional and Behavioral Disorders, 6*(2): 66-80.

Walker, H.M., Seeley, J.R., Small, J., Severson, H.H., Graham, B.A., Feil, E.G., Serna, L., Golly, A.M., and Forness, S.R. (2009). A randomized controlled trial of the First Step to Success early intervention: Demonstration of program efficacy outcomes in a diverse, urban school district. *Journal of Emotional and Behavioral Disorders, 17*(4): 197-212.

Webster-Stratton, C., Reid, M.J., Hammond, M. (2004). Treating children with early-onset conduct problems: Intervention outcomes for parent, child, and teacher training. *Journal of Clincal Child and Adolescent Psychology, 33(1)*: 105-124.

Zins, J.E., and Ponti, C.R. (1990). Strategies to facilitate the implementation, organization, and operation of system-wide consultation programs. *Journal of Educational and Psychological Consultation, 1*: 205-218.

Zucker, D.M. (1990). An analysis of variance pitfall: The fixed effects analysis in a nested design. *Education and Psychological Measurement, 50*: 731-738.