Journal of Positive Behavior Interventions
Volume 12 Number 3
July 2010 133-148
© 2010 Hammill Institute on Disabilities
10.1177/1098300709334798
http://jpbi.sagepub.com
hosted at
http://online.sagepub.com

Examining the Effects of Schoolwide Positive Behavioral Interventions and Supports on Student Outcomes

Results From a Randomized Controlled Effectiveness Trial in Elementary Schools

Catherine P. Bradshaw
Mary M. Mitchell
Philip J. Leaf
Johns Hopkins University

Schoolwide Positive Behavioral Interventions and Supports (SWPBIS) is a universal, schoolwide prevention strategy that is currently implemented in over 9,000 schools across the nation to reduce disruptive behavior problems through the application of behavioral, social learning, and organizational behavioral principles. SWPBIS aims to alter school environments by creating improved systems and procedures that promote positive change in student behavior by targeting staff behaviors. This study uses data from a 5-year longitudinal randomized controlled effectiveness trial of SWPBIS conducted in 37 elementary schools to examine the impact of training in SWPBIS on implementation fidelity as well as student suspensions, office discipline referrals, and academic achievement. School-level longitudinal analyses indicated that the schools trained in SWPBIS implemented the model with high fidelity and experienced significant reductions in student suspensions and office discipline referrals.

Keywords: *Schoolwide Positive Behavioral Interventions and Supports; randomized controlled trial; effectiveness research; suspensions; office discipline referrals; achievement*

Given the increased emphasis on accountability for student achievement and discipline problems resulting from the No Child Left Behind Act, local school districts and administrators are increasingly turning to schoolwide prevention models to promote a positive school climate and reduce discipline problems. Many of these programs systematically manage student behavior problems by creating schoolwide plans that clearly articulate positive behavioral expectations, provide incentives to students who meet those behavioral expectations, and establish a consistent strategy for managing student behavior problems (Horner, Sugai, Todd, & Lewis-Palmer, 2005; Sugai & Horner, 2006). Whole-school programs are attractive to local school systems because they are believed to foster an optimal learning environment for all students and encourage the use of additional supports for children with greater social-emotional and behavioral needs.

Positive Behavioral Interventions and Supports (Horner et al., 2005; Sugai & Horner, 2006; Sugai, Horner, & Gresham, 2002) is one such whole-school prevention strategy that seeks to enhance the school's capacity to prevent disruptive behavior by creating and sustaining primary (schoolwide/universal), secondary (targeted/selective), and tertiary (individual/indicated) systems of support. The three-tiered prevention model follows a public health approach (Mrazek & Haggerty, 1994) whereby

Authors' Note: Support for this project comes from grants from the Centers for Disease Control and Prevention (R49/CCR318627, 1U49CE 000728, and K01CE001333-01) and the National Institute of Mental Health (1R01MH67948-1A). The authors would like to thank the Positive Behavioral Interventions and Supports Maryland State Leadership Team for their support of this project, especially Milton McKenna and Andrea Alexander of the Maryland State Department of Education and Susan Barrett and Jerry Bloom of the Sheppard Pratt Health System. Correspondence regarding this article can be sent to Catherine P. Bradshaw, Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, 624 N. Broadway, Room 831, Baltimore, MD 21205; e-mail: cbradsha@jhsph.edu.

two levels of targeted and indicated programs are implemented to complement and build on the universal schoolwide components of the model (for a review, see Horner et al., 2005; Sugai & Horner, 2006; Sugai et al., 2002; Taylor-Greene et al., 1997; Taylor-Greene & Kartub, 2000).

The universal level of the three-tiered model, referred to as Schoolwide Positive Behavioral Interventions and Supports (SWPBIS), is being widely disseminated by the U.S. Department of Education (Knoff, 2000) and several state departments of education (e.g., Illinois, North Carolina, Colorado, Maryland, Oregon). It is estimated that SWPBIS is currently implemented in more than 9,000 schools across the United States (Horner, 2009), in at least 44 states in the United States, and in several other countries around the world (e.g., Norway, Canada, Australia). Despite the growing use and acceptance of SWPBIS and other whole-school interventions, there have been few longitudinal randomized controlled trials evaluating the effectiveness of SWPBIS. This effectiveness study used a group randomized controlled trial design to determine the effects of SWPBIS training on implementation fidelity and student outcomes, including suspensions, office referrals, and academic achievement.

Overview of SWPBIS

SWPBIS is a universal prevention strategy that aims to alter the school environment by creating improved systems (e.g., discipline, reinforcement, data management) and procedures (e.g., office referral, training, leadership) that promote positive change in staff behaviors, which subsequently alter student behaviors. The model draws on behavioral, social learning, and organizational behavioral principles (Lewis & Sugai, 1999; Lindsley, 1992), which were traditionally used with individual students but have been generalized and applied to an entire student body consistently across all school settings (Durand & Carr, 1992).

This article focuses solely on the universal SWPBIS model that is implemented by following these seven steps, which map onto the model's critical features: (a) Within the school, a SWPBIS team is formed that includes 6 to 10 staff members and an administrator, all of whom provide building-level leadership regarding the implementation of SWPBIS. The team attends annual training events, establishes an action plan for implementing SWPBIS, develops materials to support program implementation, trains other staff members, and meets at least twice a month to discuss schoolwide behavior management systems and procedures. (b) An external behavioral support coach provides on-site consultation and technical assistance regarding the implementation of SWPBIS. The coach is typically a school psychologist or guidance counselor who has prior experience working with SWPBIS and conducting functional behavioral assessments. The coach is an important member of the school-level SWPBIS team and attends at least one team meeting each month. (c) Expectations for positive student behavior are defined and known by staff and students. The school team establishes three to five positively stated schoolwide expectations for student behavior (e.g., "Be respectful, responsible, and ready to learn"), which are posted in all classrooms and nonclassroom settings. (d) Defined behavioral expectations are taught to all students. Lesson plans are developed by the school staff for teaching students the schoolwide behavioral expectations at the beginning of the school year and at least once a month thereafter. (e) A schoolwide system is developed to reward students who exhibit the expected positive behaviors. School staff members establish and use a schoolwide system for reinforcement that includes a tangible reinforcer (e.g., "high-five" or "gotcha") that is used consistently by all school staff in classroom and nonclassroom settings. (g) An agreed-upon system is created to respond to behavioral violations. Staff and administrators agree on what constitutes a classroom-managed versus an office-managed discipline problem, and students across all classrooms receive consistent consequences for disciplinary infractions. (h) A formal system is developed to collect, analyze, and use disciplinary data for data-based decision making. Disciplinary data (e.g., office discipline referrals, suspensions) are systematically collected, analyzed, and summarized in a report that is used by the SWPBIS team to make decisions regarding program implementation. All staff members receive training from the schools' SWPBIS team on the procedures for documenting discipline problems through the use of an office discipline referral form. An Internet-based data entry and reporting system called the School-Wide Information System (SWIS) was used to facilitate this process (www.SWIS.org; May et al., 2003).

Results from nonrandomized studies, which primarily used pre/post comparisons in one or two schools or large groups of schools (e.g., statewide evaluations), have shown promising effects of the universal SWPBIS model. Specifically, training in SWPBIS has been shown to lead to sustained changes in schools' internal discipline practices and systems (Barrett, Bradshaw, & Lewis-Palmer, 2008; Nersesian, Todd, Lehmann, & Watson, 2000; Taylor-Greene & Kartub, 2000). With regard to

outcomes, implementation of SWPBIS in a rural middle school was associated with a 42% reduction in office discipline referrals (Taylor-Greene et al., 1997). Similarly, a study of an urban elementary school observed reductions in discipline problems and improvements in academic outcomes following training in SWPBIS (Luiselli, Putnam, Handler, & Feinberg, 2005; also see Metzler, Biglan, Rusby, & Sprague, 2001). Larger statewide evaluations of SWPBIS have documented significant reductions in suspensions among elementary and middle schools (Barrett et al., 2008) and reductions in office discipline referrals among middle and high schools (Muscott, Mann, & LeBrun, 2008) trained in the schoolwide model.

Findings from a recent 3-year randomized trial of SWPBIS conducted by the developers of SWPBIS using a waitlist design indicated that implementation of the model was associated with improvements in students' perceptions of safety at school, an increase in third-grade reading performance, and reductions in office disciplinary referrals (Horner et al., in press). Furthermore, previous studies reporting data from the current sample of 37 elementary schools participating in a randomized effectiveness trial indicated significant improvements in the school staff members' perceptions of the schools' organizational health after training in SWPBIS (Bradshaw, Koth, Bevans, Ialongo, & Leaf, 2008; Bradshaw, Koth, Thornton, & Leaf, in press) as well as a significant reduction in students' need for and use of school-based counseling services (Bradshaw, Mitchell, & Leaf, 2008). For additional information on the evidence base for SWPBIS see the National Technical Assistance Center on Positive Behavioral Interventions and Supports (2007).

The aim of this article is to extend the recent effectiveness research on SWPBIS by examining the intervention effects on behavioral and academic outcomes for students across 5 years. Given the increased accountability for ensuring positive student outcomes, it is important to examine if there were significant changes occurring in schools that were randomly assigned to receive training in SWPBIS as compared to schools randomly assigned to not receive training. In an effort to better understand how training in the SWPBIS schoolwide reform model results in sustainable changes in the schools' management procedures (Doolittle, 2006; McIntosh, Horner, & Sugai, 2009), we also examined both the fidelity and sustainability of SWPBIS as implemented over the course of 4 years. Examining the multiyear effects of SWPBIS on student outcomes as compared to multiple years of "care as usual" using a rigorous randomized controlled trial design is essential for determining whether SWPBIS is an evidence-based program.

Method

Data

The data for this study came from a longitudinal group randomized effectiveness study of SWPBIS conducted to determine the impact of the model on discipline problems, student achievement, and the school environment. Thirty-seven Maryland public elementary schools from five school districts (rural and suburban) volunteered to participate in the trial. Maryland was selected because of the state's developing infrastructure to support high quality implementation of SWPBIS, the commitment from the community agencies (i.e., Maryland State Department of Education, Sheppard Pratt Health System, and local school systems), and the proximity to the researchers' institution. The schools were matched on select baseline demographics (e.g., percentage of students receiving free or reduced meals), of which 21 schools were randomized to the intervention condition (referred to as *SWPBIS schools*) and 16 were assigned to the comparison condition (referred to as *comparison schools*). A slightly higher proportion of schools were randomized to the SWPBIS condition to increase the statistical power to examine research questions regarding variation in implementation quality (see Bradshaw, Koth, et al., in press; Bradshaw, Reinke, Brown, Bevans, & Leaf, 2008).

Training. Because the design of the study was an effectiveness trial (Flay, 1986), which examines the impact of programs under "real-world conditions," rather than an efficacy trial, which examines "optimal conditions of delivery" (Flay et al., 2005, p. 153), the SWPBIS training and support system were led and coordinated by the state (rather than the researchers) following the state of Maryland's SWPBIS typical training procedures (Barrett et al., 2008). Specifically, each of the 21 schools assigned to receive SWPBIS training formed internal SWPBIS teams comprised of 6 to 10 members (e.g., staff, teachers, administrators), of which 4 to 5 team members (including an administrator) attended an initial 2-day summer training led by Dr. George Sugai, one of the developers of SWPBIS. In addition to reviewing in detail the core features of the SWPBIS model, the staff members were trained to develop an implementation and training plan, which included professional development for the other school staff. Schools were instructed to hold two additional planning and training days at their schools prior to implementing the model to further develop their action

Table 1
School Characteristics for Schoolwide Positive Behavioral Interventions and Supports (SWPBIS) and Comparison Schools at Baseline

	SWPBIS (n = 21 schools)		Comparison (n = 16 schools)	
School Characteristics	*M*	*SD*	*M*	*SD*
School enrollment	471.76	132.78	505.50	188.57
Student-to-teacher ratio	18.48	4.33	18.61	4.69
Free and reduced-price meals (%)	42.93	19.22	36.25	20.93
Special education students (%)	13.24	4.27	15.08	6.66
Caucasian students (%)	53.81	33.16	67.51	28.99
Student mobility (%)	25.88	8.24	20.51	7.19
Suspension (%)	7.73	7.43	5.06	4.73
Math performance (%)[a]	47.20	22.37	46.96	19.05
Reading performance (%)[a]	50.66	19.32	52.94	16.43

Note: An overall MANOVA on the school-level characteristics indicated no significant difference between schools trained in SWPBIS and comparison schools at baseline, Wilks's $\Lambda = .674$, $F(9, 19) = 1.022$, $p = .46$.
a. Percentage of fifth-grade students who scored in the profficent or advanced range on the state's standardized test.

plan for implementation. To ensure and maintain consistently high levels of implementation fidelity, SWPBIS school teams participating in this study (like other SWPBIS schools in the state) attended annual 2-day summer booster training events. All initial training and booster training events were coordinated and led by the SWPBIS Maryland State Leadership Team and were attended by other SWPBIS teams from across the state. The booster training sessions were led by members of the Maryland State Leadership Team and affiliates of the National Positive Behavioral Interventions and Supports Technical Assistance Center, and they covered basic elements of SWPBIS implementation and issues related to sustainability. Additional supports and professional development were provided to the schools' behavior support coaches through state-coordinated day-long training events conducted four times each year. Each coach also participated in an initial day-long coaches training and attended the day-long annual booster SWPBIS training events with their schools' SWPBIS teams. All coaches' training and technical support sessions were led by the State's SWPBIS trainers (e.g., Ms. Susan Barrett, who is affiliated with National Positive Behavioral Interventions and Supports Technical Assistance Center) and were coordinated by the Maryland State Leadership Team. Because of the design of the effectiveness trial, coaches were assigned and supervised by the school districts rather than the researchers. See Barrett et al. (2008) for additional information about the Maryland SWPBIS training and support infrastructure.

Participating schools. The sample of participating elementary schools was diverse and representative of other elementary schools in those districts (Stuart & Leaf, 2007). Specifically, 48% of the participating schools were suburban, 41% were urban fringe, and 49% received Title I support. Additional baseline school-level demographic characteristics are provided in Table 1. The schools were located in five different school districts, which varied with regard to size. The number of participating schools per district ranged from 2 to 13. The 37 schools were enrolled in the project over 2 consecutive years to ease the training, support, and research burden. Identical recruitment, training, and support procedures were used for schools in both cohorts. Dr. Sugai led both training events, which followed the same format and covered the same content.

Efforts to ensure compliance with treatment condition. Procedures were developed to increase the likelihood that schools in the trial would comply with their randomized treatment status. Specifically, upon volunteering to participate in the trial (prior to randomization), the administrator at each school reviewed and signed a written participation form in which they voluntarily agreed that if the school were randomized to the comparison condition that the school would not implement SWPBIS for the duration of the trial. The participating schools' administrators also agreed in writing on this form that if the school were randomized to the SWPBIS condition, the school staff would attend the training and support implementing SWPBIS for the duration of the trial. When there was a change of administrator at any school participating in the trial, the lead investigator or research project coordinator met with the school administrator in person to review the goals of the project,

expectations of the school, and required data elements. In the rare instances where there was a higher than expected score on the implementation quality measure in one of the comparison schools (there were approximately three comparison schools with higher than expected implementation quality scores), the lead researcher on the project contacted the schools' administrator and the school district coordinator for SWPBIS to inquire whether the school had received formal training in SWPBIS; however, there was no evidence of formal training in any instance. To further ensure that schools in the comparison condition did not receive training in SWPBIS, the state's SWPBIS training coordinator received a list of schools in the comparison condition; these schools were not eligible to register for or participate in the state's training events. All schools were informed of these procedures along with the randomization process in writing prior to agreeing to enroll in the trial.

Measures

Implementation fidelity. The *School-Wide Evaluation Tool* (SET; Sugai, Lewis-Palmer, Todd, & Horner, 2001) consists of 29 items organized into the following seven subscales that represent the seven key features of SWPBIS: Expectations Defined, Behavioral Expectations Taught, System for Rewarding Behavioral Expectations, System for Responding to Behavioral Violations, Monitoring and Evaluation, Management, and District-Level Support. Each item of the SET is scored on a 3-point scale (0 = not implemented, 1= partial implementation, and 2 = full implementation) with each subscale represented by a percentage (from 0% to 100%); higher scores indicate greater fidelity of SWPBIS. The seven subscale scores were averaged to yield an overall SET score. The SET was administered by a trained observer who determined the degree to which a school had implemented each of the model's seven critical features by reviewing written materials and established discipline procedures (e.g., discipline handbook, school improvement goals, behavioral incident summaries) as well as noting visual displays of the three to five expected behaviors posted in 10 specified locations throughout the school (e.g., hall, classrooms, cafeteria, library). SET assessors conducted brief interviews with an administrator (lasting approximately 30 minutes each) regarding the school's procedures, policies, and standards for positive behavior and rule infractions. In cases where the school had more than one administrator, the administrator most familiar with the behavior or discipline program at the school participated in the interview. Brief interviews were conducted with at least two students (one boy and one girl) per class per grade level (1 to 3 minutes each) regarding the school's system and standards for positive behavior and rule infractions. It was not feasible to randomly select the students to be interviewed; however, efforts were made by the interviewers to select children who were representative of the student body with regard to ethnicity. At least 12 staff members per school were selected using a random number table and interviewed for approximately 2 to 5 minutes each about school procedures, policies, and standards for positive behavior and rule infractions (see Horner et al., 2004, for additional information on the SET administration and scoring procedures).

Prior research by Horner et al. (2004) indicated that the SET has strong psychometric properties with regard to internal consistency (Cronbach's α = .96), interobserver reliability (range = 98.4% to 100%), and test-retest reliability (range 93% to 100%, mean = 97.3%). For this study, the SET's detailed scoring guide and documentation form were rescored by project staff to ensure accuracy. The interrater reliability of the SET scoring guide was high (r = .99, p < .001), and all scoring discrepancies were resolved through consultation with the local lead trainer. The Cronbach's alpha for the 29-item SETs conducted on this sample of 37 schools was .85 at baseline and over .90 in all subsequent years. Factor analyses on these data also confirmed the original factor structure of the SET, and the alphas for the individual SET subscales ranged from .64 to .96.

The *Effective Behavior Support Survey* (EBS; Sugai, Todd, & Horner, 2000) is a staff-report measure of implementation fidelity. The survey was completed by all staff in both conditions to determine the extent to which the four behavior support systems were considered in place in the school: (a) schoolwide discipline systems (15-item α for this sample = .90; e.g., "Schoolwide expected student behaviors are taught directly" and "Schoolwide expected student behaviors are rewarded regularly"), (b) nonclassroom management systems (such as the cafeteria, hallway, playground; 9-item α = .85; e.g., "Rewards exist for meeting expected schoolwide student behaviors in nonclassroom settings" and "All staff are involved directly or indirectly in management of nonclassroom settings"), (c) classroom management systems (12-item α = .87; e.g., "Expected student behavior and routines in classrooms are stated positively and defined clearly" and "Classroom-based options exist to allow classroom instruction to continue when problem behavior occurs"), and (d) systems for individual students engaging in chronic problem behaviors (8-item α = .87; e.g., "A simple process exists for teachers to request assistance" and "Behavior is monitored and feedback

provided regularly to the behavior support team and relevant staff"). Each item was measured by a 3-point scale (in place, partially in place, or not in place) with higher scores indicating a higher level of SWPBIS implementation. The EBS measure required approximately 5 to 7 minutes to complete.

Student outcomes. Office discipline referral (ODR) data were collected only from the trained schools in Years 1 through 4 using the *School-Wide Information System* (SWIS; May et al., 2003). SWIS is an Internet-based data system used to collect and manage student discipline referral data by the school staff. Both major (e.g., abusive language, fighting, and lying) and minor (e.g., physical contact, disruption, and property misuse) office referrals were examined. An important aspect of the schools' training in SWPBIS focuses on development of an office disciplinary referral form that is consistent with the SWIS. Similarly, the SWPBIS team was charged with training all staff in the SWPBIS schools to complete the referral form to ensure consistent documentation of incidents across staff (see Irvin, Tobin, Sprague, Sugai, & Vincent, 2004). ODR data are not available for the baseline year or from the comparison schools because the use of SWIS was conceptualized as part of the SWPBIS program. Some researchers (e.g., Irvin et al., 2004) have questioned the validity of ODR data when collected without formal training or without a systematic data collection procedure, such as the SWIS; therefore, we only examined the SWIS ODR data from the SWPBIS schools following training in SWPBIS.

School-level suspension rates were obtained from the Maryland State Department of Education (MSDE) for the baseline year through Year 4 of the study. A single suspension rate was calculated for each school year by using the total number of suspension events for a given school year divided by the student enrollment for that year and then multiplied by 100 to yield a percentage score.

School-level scores on the state's standardized academic achievement tests, Maryland School Assessment (MSA), for third and fifth grade math and reading were obtained for the relevant school years from the MSDE. Prior research by the MSDE (2003, 2007) documented the reliability and validity of the reading and math MSAs, as indicated by an expert panel review of test items, high intercorrelations between sections of the tests, strong evidence of unidimensionality from principal component analysis, and high interrater reliability. There also were strong associations between the MSA and the Voluntary State Curriculum. Because the state adopted the MSA after the 1st year of the trial, we do not have baseline data from the participating schools. Furthermore, there was a general positive trend in the MSA scores across schools in the entire state (MSDE, 2007), suggesting that the overall positive trend in MSA scores across schools in both conditions was associated with the historical event of the adoption of the new state test. Consequently, we combined the percentage of students who scored in the advanced and proficient ranges (i.e., those receiving passing scores on the test) on each test across all time points to yield a single gain score for each school for each of the tests across all years of the trial.

Procedure

Administration of the fidelity measures. The SET was administered at schools in both conditions at baseline (May prior to randomization and SWPBIS training) and annually in the spring (May) thereafter. A total of 15 SET assessors were hired by the project, with approximately 10 assessors conducting SETs for the project each year. The SET assessors were primarily master- and doctorate-level professionals (e.g., teachers, special educators, school counselors, educational trainers), most of whom were consulting with school districts (not involved in the study) or had recently retired from full-time work in an educational setting. They were hired by the research project and kept unaware of the primary purpose of the study and the schools' intervention status during all phases of the study; the SET assessors were not involved in any other aspect of the data collection for the project. An initial 2-day didactic group SET training session on interpretation of items, strategies for administering the SET, and the mechanics and scoring of the SET was conducted by Dr. Teri Lewis-Palmer, a developer of the SET. Two additional days of one-on-one training and shadowing occurred in an SWPBIS elementary school (that was not included in the trial) by Dr. Lewis-Palmer and a local lead trainer (Ms. Susan Barrett). All assessors obtained at least 85% interobserver agreement with the trainer prior to conducting a SET in a project school. The average interobserver agreement was 91.62% for the initial SET training for all SET assessors. On an annual basis, a 1-day group refresher training session was led by the local lead trainer for all previously trained SET assessors, during which core aspects of the administration and scoring of the SET were reviewed. All assessors again obtained at least an 85% interobserver agreement level with the lead SET trainer in a nonproject school. For additional information regarding the SET training and administration procedures, see Bradshaw, Reinke, et al. (2008).

All staff at the schools in both conditions completed the EBS. The survey packets were mailed in bulk to the school and distributed via faculty mailboxes to the school staff by an administrator, school psychologist, or administrative assistant. Staff members were informed through the written consent form that their participation was voluntary; their individual responses would not be shared with anyone at the school, district, or state department of education; their names would be removed from the data by the researchers; and their data would only be summarized in aggregate form. Staff members completed the study materials on their own time and returned the materials directly to the researchers through the U.S. mail in the self-addressed, stamped envelopes provided by the researchers. Each staff questionnaire packet included a small incentive with an approximate value of less than one dollar (e.g., disposable ballpoint pen, bookmark). The staff response rate ranged from 80% to 86% across the 5 years and did not differ significantly across the SWPBIS and comparison conditions. Baseline data, along with 4 subsequent years of data were analyzed in this study. The EBS data were aggregated to the school level by averaging all staff members' scores (i.e., percent items marked as "in place") and calculating a single score for each subscale for each school across each of the 5 years of the trial.

Analyses

Given our focus on school-level indicators of implementation fidelity and student outcomes, all analyses were conducted in SPSS 16.0 using the school-level data. We examined the impact of training in SWPBIS on implementation fidelity, as measured by the SET and EBS, by employing repeated measures general linear models (GLM). Specifically, we analyzed the data from baseline through Year 4 by using unadjusted repeated measures GLM and computing the Wilks's Lambda to determine if there were significant intervention condition by time effects on the individual subscales. We reported the partial η^2 (Tabachnick & Fidell, 2001) and Cohen's d (Cohen, 1992; Hedges, 2007) for all statistically significant effects to indicate the strength of the effect of SWPBIS training on the change in outcomes over time.

We subsequently analyzed the effect of SWPBIS on student outcomes such as ODR, suspension rates, and achievement scores. Since the SWIS was only collected from the SWPBIS-trained schools, we could not compare intervention and control schools on ODRs. Therefore, we analyzed the ODR data from SWPBIS schools using repeated measures GLM. Next, we examined whether there was a significant reduction in the suspension rate for the SWPBIS and comparison schools using the Wilcoxon signed rank test. Finally, we examined the improvement in school-level achievement across the four state test subscale scores using t tests.

Results

Implementation Fidelity

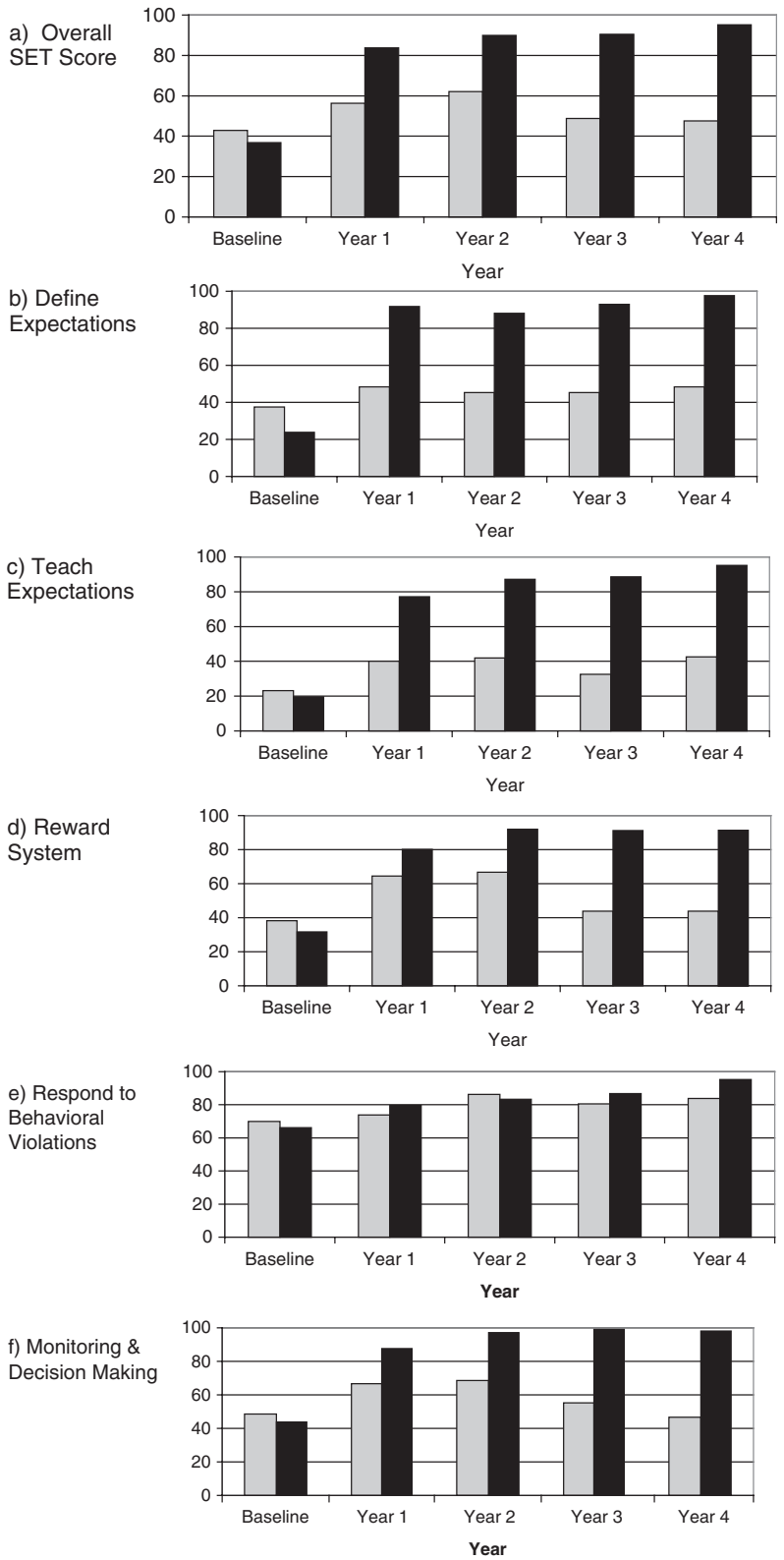
SET. Analysis of the SET data suggested a significant intervention effect (i.e., Intervention Condition × Time Interaction) for the overall SET score, Wilks's $\Lambda = .38$, $F(4, 32) = 13.36$, $p = .001$, $\eta^2 = .63$, $d = 3.22$ (Figure 1a). In addition, there were significant intervention effects on six of the seven subscale scores: define expectations, Wilks's $\Lambda = .55$, $F(4, 32) = 6.64$, $p = .001$, $\eta^2 = .45$, $d = 2.10$ (Figure 1b); teach behavioral expectations, Wilks's $\Lambda = .51$, $F(4, 32) = 7.77$, $p = .001$, $\eta^2 = .49$, $d = 2.63$ (Figure 1c); reward system, Wilks's $\Lambda = .55$, $F(4, 32) = 6.54$, $p = .001$, $\eta^2 = .45$, $d = 1.99$ (Figure 1d); monitoring and decision making, Wilks's $\Lambda = .58$, $F(4, 32) = 5.74$, $p = .001$, $\eta^2 = .42$, $d = 1.94$ (Figure 1f); management, Wilks's $\Lambda = .45$, $F(4, 32) = 9.70$, $p < .001$, $\eta^2 = .55$, $d = 2.40$ (Figure 1g); and district support, Wilks's $\Lambda = .68$, $F(4, 32) = 3.70$, $p = .014$, $\eta^2 = .32$, $d = 1.54$ (Figure 1h). The seventh subscale, behavioral violations, was significant when comparing Year 1 and Year 4 scores, Wilks's $\Lambda = .87$, $F(1, 35) = 5.33$, $p = .027$, $\eta^2 = .13$, $d = .90$, but did not reach statistical significance when including all 4 years of data in the repeated measures GLM, Wilks's $\Lambda = .77$, $F(4, 32) = 2.37$, $p = .073$, $\eta^2 = .23$, $d = .90$ (Figure 1e).

EBS. All repeated measures GLM analyses indicated a statistically significant effect of training in SWPBIS on the four EBS subscale scores. Specifically, there was a significant effect of SWPBIS training on the percentage of staff rating their school as having schoolwide systems in place, Wilks's $\Lambda = .50$, $F(4, 32) = 7.97$, $p < .001$, $\eta^2 = .50$, $d = 1.71$ (Figure 2a); in addition to the ratings for nonclassroom settings, Wilks's $\Lambda = .50$, $F(4, 32) = 7.88$, $p < .001$, $\eta^2 = .50$, $d = 1.47$ (Figure 2b); classroom settings, Wilks's $\Lambda = .67$, $F(4, 32) = 3.95$, $p = .01$, $\eta^2 = .33$, $d = 1.08$ (Figure 2c); and individualized student systems, Wilks's $\Lambda = .58$, $F(4, 32) = 5.75$, $p = .001$, $\eta^2 = .42$, $d = 1.46$ (Figure 2d).

Outcome Measures

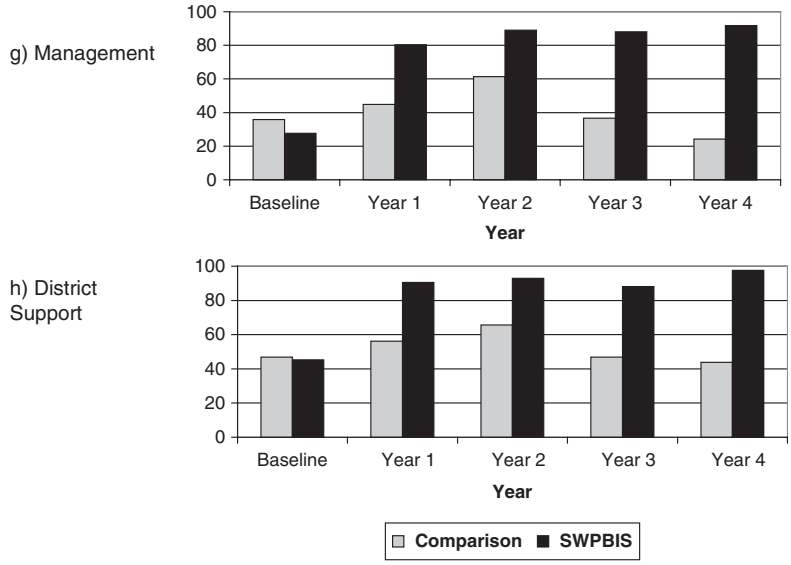
ODRs. Three analyses of ODR data were conducted using repeated measures GLM to determine whether there were significant differences in the rates of ODRs

Figure 1
Unadjusted Mean SET Scores at SWPBIS and Comparison Schools at Baseline and Years 1 Through 4



(continued)

Figure 1 (continued)



across the four posttraining years. Based on prior research indicating that school-level factors are associated with different rates of student behavior problems (Birnbaum et al., 2003), we controlled for school system or district, percentage of students receiving free and reduced meals, and school enrollment in the analysis of the ODR data. We first calculated the number of major office referrals per 100 students per day for each SWPBIS-trained school over the course of the trial. This rate was .201 at the end of the 1st year of the trial and dropped to .159 in the last year of the trial. Across all years of the trial, the rate of major ODRs per 100 students per day remained well below the national SWIS average, which ranged from .34 to .37 for the school years spanning the trial (Spaulding et al., 2008; Figure 3a). The repeated measures GLM for the rates of major ODRs per 100 students per day was nonsignificant (Wilks's $\Lambda = .84$, $F(1, 14) = 2.59$, $p = .13$, $\eta^2 = .16$, $d = .21$ (Figure 3a). However, the percentage of students with a major or minor ODR decreased significantly over the course of the study from 18.8% to 18.1%, Wilks's $\Lambda = .67$, $F(1, 14) = 6.99$, $p = .019$, $\eta^2 = .33$, $d = .08$ (Figure 3b). The number of major and minor ODR events per student also decreased significantly over the course of the trial, Wilks's $\Lambda = .52$, $F(1, 14) = 12.90$, $p = .003$, $\eta^2 = .48$, $d = .12$ (Figure 3c).

Suspensions. The mean suspension rates for each year of the study are reported for SWPBIS and Comparison schools (Figure 4). A Wilcoxon signed ranks test, in which a Z score was computed separately for the two conditions, was nonsignificant for the comparison schools ($Z = -1.54$, $p = .12$) but was statistically significant for the SWPBIS schools ($Z = -2.17$, $p = .03$, $d = .27$). This test indicates that the percentage of students receiving suspensions significantly declined over time for SWPBIS schools but not for comparison schools.

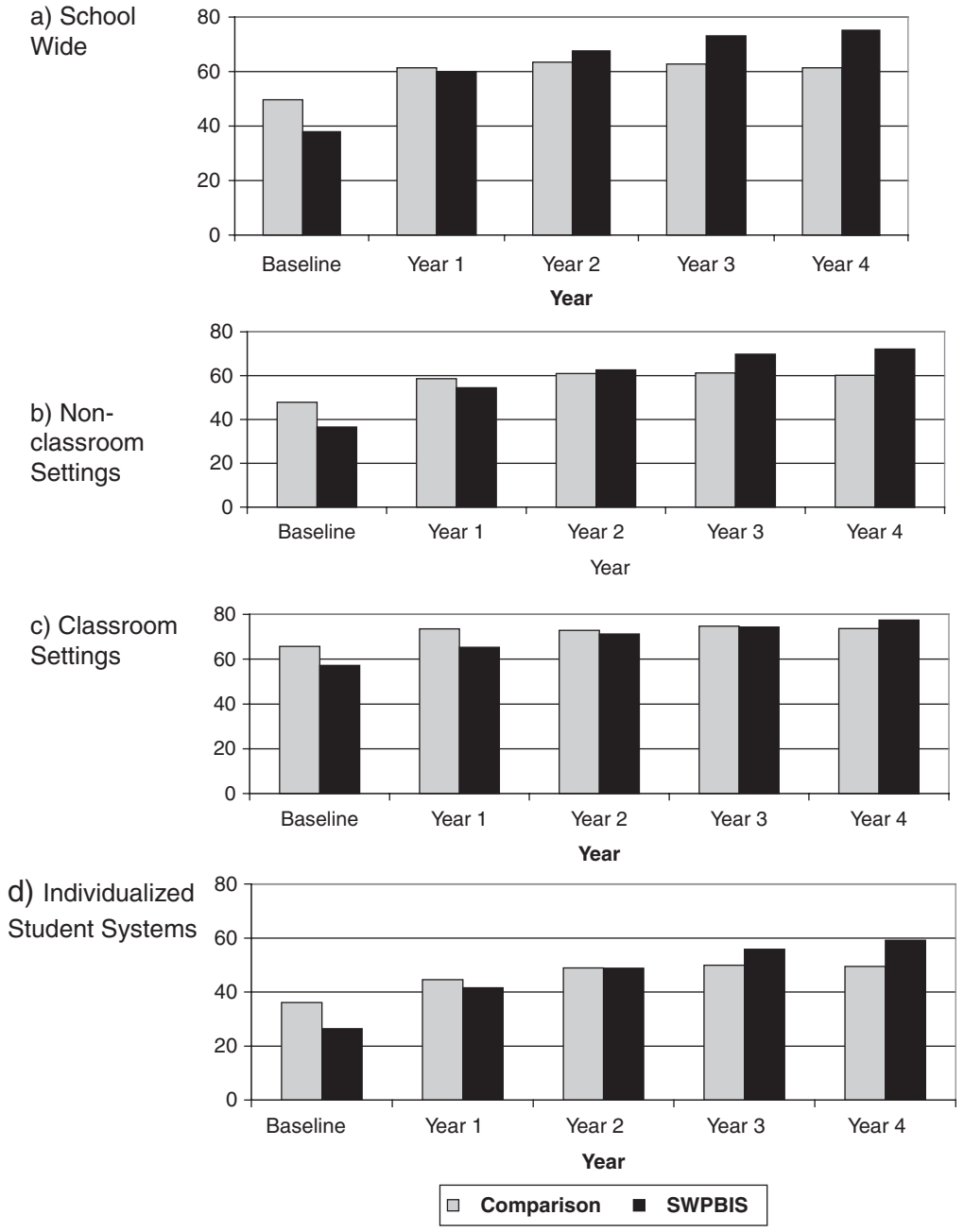
School-level achievement. Figure 5 shows the cumulative increases in the percent of third- and fifth-grade students who scored in the proficient and advanced ranges of the math and reading state standardized achievement tests over the study period. The data suggest a trend, although nonsignificant, for fifth-grade math, such that fifth graders in SWPBIS schools tended to demonstrate greater gains in math scores compared to the gains made by the comparison school students ($t = -1.67$, $df = 35$, $p = .105$, $d = .54$). However, t tests comparing cumulative gains in test scores between students in SWPBIS and comparison schools indicate no difference for third-grade math and reading or fifth-grade reading, although the improvement in scores tended to be greater for SWPBIS schools than for comparison schools on third- and fifth-grade reading.

Discussion

Impact of Training on Implementation Quality and Sustainability

This study used data from a longitudinal randomized controlled trial of SWPBIS to examine the effectiveness

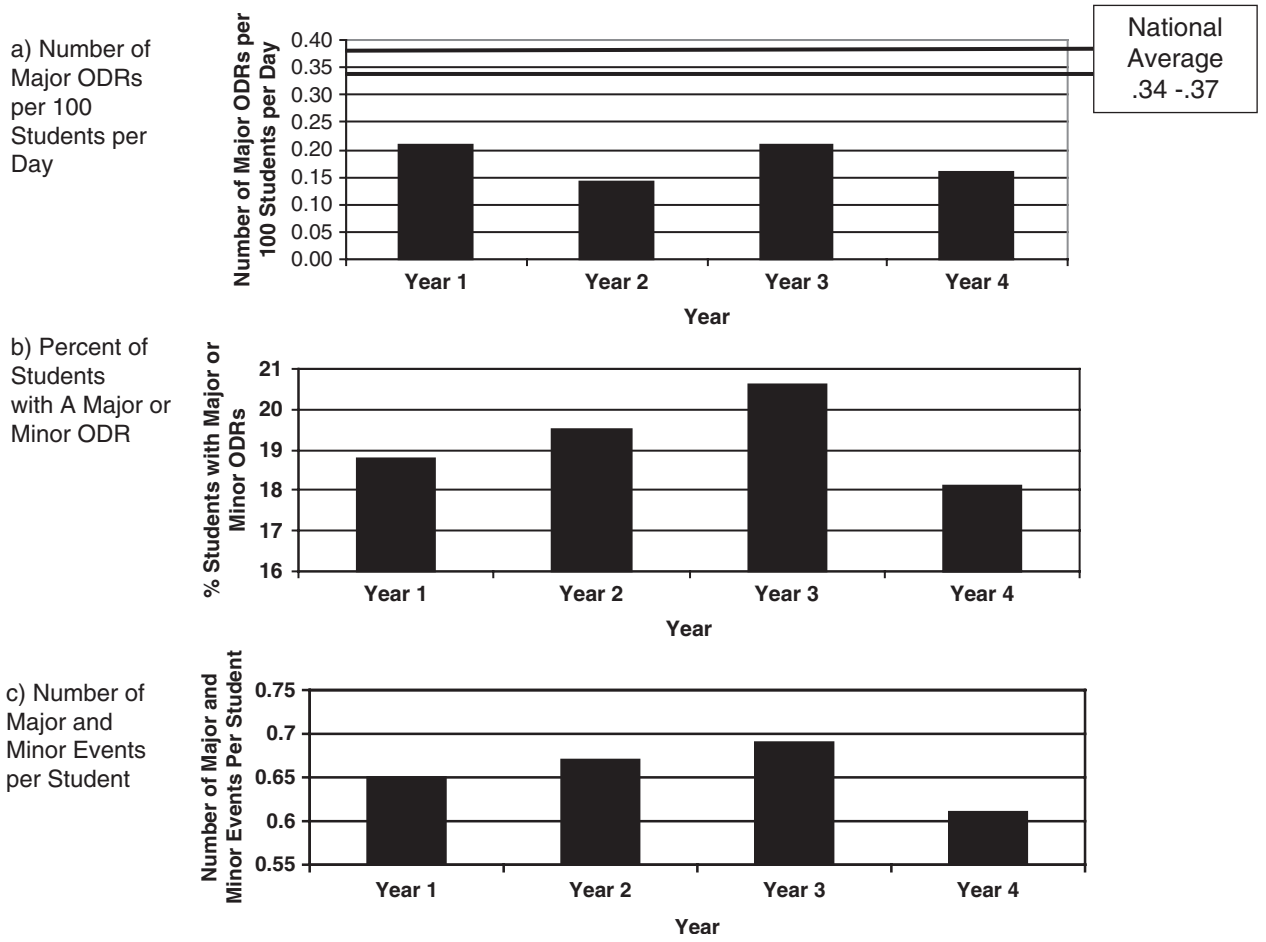
Figure 2
Mean Staff-Reported EBS Scores at SWPBIS and Comparison Schools



of the schoolwide prevention model in elementary schools. The analyses of the school-level data indicated that the schools trained in SWPBIS implemented the program with high fidelity, according to both staff self-reports (i.e., EBS) and assessments (i.e., SET) conducted by outside evaluators who were unaware of the schools' intervention status. Once reaching high fidelity, all 21 trained schools sustained high fidelity implementation of SWPBIS for the duration of the trial, as evidenced by higher EBS and SET scores.

Although the effect sizes (1.08 to 1.71) for the EBS were very large (Cohen, 1992), they were not as large as the effect sizes for the SET subscales (1.54 to 2.63). This suggests that the EBS subscales may not be as sensitive as the SET scores to the changes that occurred as a result of training in SWPBIS. Inspection of the EBS subscale

Figure 3
Office Disciplinary Referrals (ODRs) for SWPBIS Schools



scores suggests that schools in the comparison condition tended to have higher baseline scores on the EBS than did the schools randomized to the SWPBIS condition; however, the SWPBIS schools improved more than the comparison schools did on all four subscale scores by the end of the trial.

Inspection of the effect sizes for the subscale scores of the SET suggests that training in SWPBIS had the strongest effects on the teaching behavioral expectations, management, and defining behavioral expectations subscales. In contrast, the effect on the system for responding to the behavioral violations subscale was only significant when comparing the baseline and Year 4 scores. The effect on this subscale was likely attenuated by a ceiling effect, resulting from relatively high baseline scores on this subscale among schools in both conditions.

Careful review of the SET and EBS data also suggested that some diffusion of particular elements of the SWPBIS model (e.g., discipline system) occurred in the comparison schools. This may have happened in response to the somewhat leading questions on the SET and EBS or perhaps in response to heightened accountability resulting from No Child Left Behind Act. For example, the increasing emphasis on responding to student behavior and discipline problems may have contributed to the increase in scores on the system for responding to behavioral violations subscale within the schools in both conditions. For further discussion of the particular elements of SWPBIS present in schools prior to formal training in SWPBIS, see Bradshaw, Reinke, et al. (2008).

Closer inspection of the pattern of means reported in Figure 1 indicates that the scores on the SET subscales within the comparison schools tended to increase after the first administration of the measure and then decline slightly after the 2nd or 3rd year, whereas they remained consistently high after Year 1 within the SWPBIS schools. This pattern suggests that although the schools in the comparison condition did implement some elements of

Figure 4
Percent of Students With a Suspension Event for SWPBIS and Comparison Schools

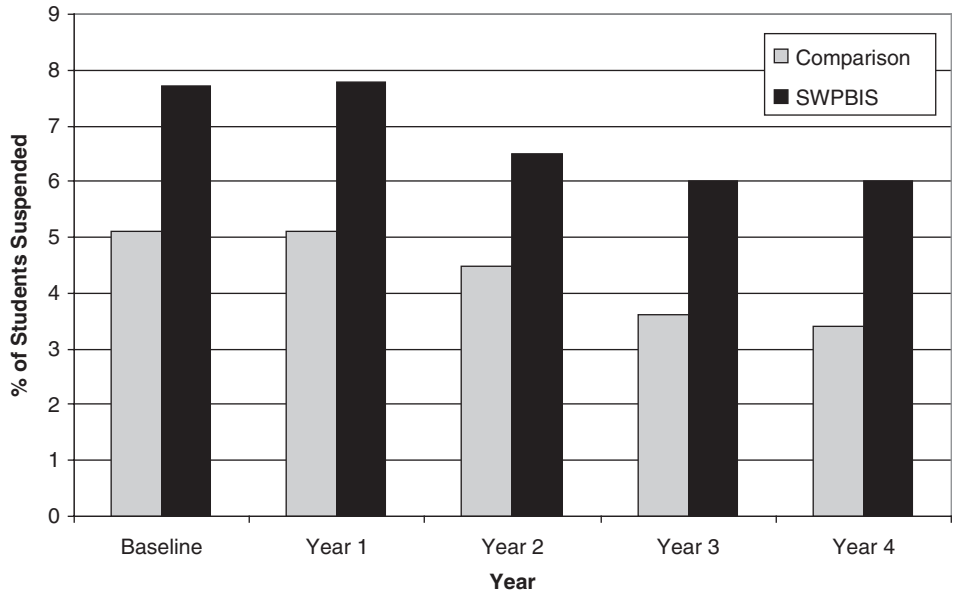
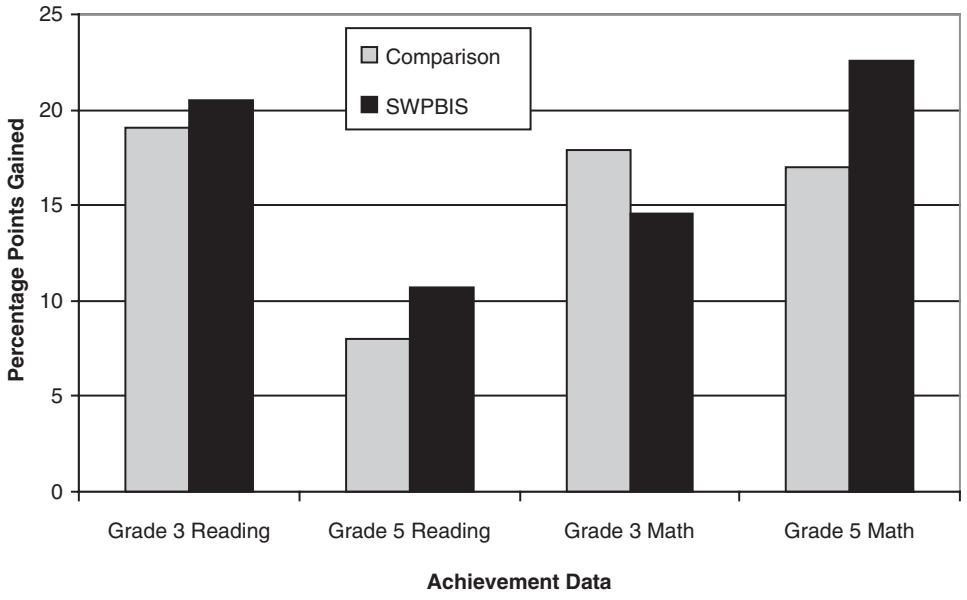


Figure 5
Cumulative Gains in Math and Reading Achievement Data for Third and Fifth Graders at SWPBIS and Comparison Schools



SWPBIS they did not sustain these efforts over the course of the trial. Recent work on sustainability of SWPBIS suggests that strong leadership at the school and district levels, onsite and ongoing coaching to support high quality implementation, evidence that SWPBIS can be implemented and incorporated into everyday practice, and evidence of the impact or efficacy of SWPBIS are critical to the sustainability of SWPBIS (Barrett et al., 2008; Doolittle, 2006; McIntosh et al., 2009). It appears that having some elements of SWPBIS in place within

the comparison schools was not sufficient to translate into sustainable implementation of the model or to influence the outcomes of the trial (Bradshaw, Koth, et al., 2008). In sum, despite the possibility of some contamination of the comparison schools, our analyses indicated that schools assigned to the SWPBIS condition, compared to nontrained schools, implemented the intervention with significantly greater fidelity that was sustained over the course of the trial.

Impact of Training on Student Outcomes

The analyses also revealed some significant effects on student outcomes. With regard to ODRs, the schools trained in SWPBIS reported a significant reduction in both the percentage of children with a major or minor ODR event as well as for the overall rate of major and minor ODR events. Our interpretation of these findings is hindered by the lack of both baseline ODR data and ODR data from the nontrained schools. It is also important to note that the schools trained in SWPBIS on average tended to have fewer major ODRs per 100 students per school day relative to the national SWIS data for elementary schools (Spaulding et al., 2008). The relatively low base rate of these problem behaviors (i.e., about one half to two thirds of the national level) suggests a possible floor effect, whereby this sample of schools had relatively little room for significant reduction in ODRs. Therefore, it is not surprising that the intervention effect sizes for the ODRs were relatively small (Cohen, 1992). Perhaps larger effect sizes would be observed in schools with higher base rates of ODRs. Additional research is needed regarding the reliability and validity of ODR data as an indicator of student behavior problems (see Irvin et al., 2004) and the extent to which these findings generalize to other measures of student behavior problems (e.g., classroom observations, teacher ratings).

With regard to the suspension outcomes, the schools trained in SWPBIS evinced a significant reduction in the rates of suspensions, whereas the rate in the nontrained schools remained unchanged during this time. The general downward trend for the suspension data in both SWPBIS schools and comparison schools, although only statistically significant within the SWPBIS schools, is generally consistent with other nonurban schools within the state during this time. However, this general downward trend in suspensions may also suggest a similar general downward trend for the ODRs in trial schools. We should also note that only the number of suspension events, rather than the number of students with a suspension event, is currently available for analysis. Therefore, it is possible that some schools may have a small group of children who account for the majority of suspension events. Additional research is needed to determine the impact of SWPBIS on the number of suspension events per student.

Finally, we examined the impact of training in SWPBIS on gains in standardized test achievement scores. Although none of the four tests reached statistical significance, the improvements observed in the SWPBIS schools tended to outpace the improvements observed in the nontrained schools on three of the four tests. It is possible that the relatively low power observed in this study reduced our ability to detect significant differences between the groups. Because the state had recently adopted a new standardized test, which was different in format and administration to the previous state test, we were unable to compare pretraining test performance with posttraining standardized test performance. As noted above, the shift in the state test likely contributed to the overall positive trend in test performance for schools in both conditions (MSDE, 2007).

It is important to reiterate that the SWPBIS training focused on developing systems to directly influence behavior management rather than academics; therefore, any such effect on educational outcomes would likely be mediated by changes in student behavior problems or improvement in the school climate and, thus, may take longer to emerge (Sugai & Horner, 2006). Specifically, it is theorized that training in SWPBIS may translate into academic outcomes for students by reducing the rates of behavior problems in the classroom, which could increase opportunities for learning (Scott & Barrett, 2004). The impact of training in SWPBIS on academic outcomes may reach statistical significance when coupled with an enhanced academic program or curriculum (McIntosh, Horner, Chard, Boland, & Good, 2006).

Challenges Encountered in Effectiveness Research

Although randomized trials of this scale are difficult to conduct with high integrity, the partnership between the researchers and the collaborating state and district agencies likely contributed to the feasibility of the study. For example, the district and state partners participated in the recruitment sessions for schools. This type of partnership between researchers and community agencies (i.e., state, districts) also likely increased the schools' compliance with the treatment condition, the submission of data by participating staff, and the retention of the schools in the trial.

However, it is important to keep in mind that the SWPBIS training and support efforts examined in this study were coordinated, facilitated, and funded by the SWPBIS state team, not the researchers. In contrast, the researchers

coordinated, facilitated, and funded all of the data collection and analysis efforts. Whereas these findings suggest that the state's efforts to produce sustainable change in the schools randomized to the SWPBIS condition were successful, it is unclear what level of training and support is necessary to bring about sustainable changes. The lack of a standardized SWPBIS training format, intensity, and duration across states and/or schools likely contributes to variations in both the implementation quality and outcomes achieved through SWPBIS. We were able to ensure that all SWPBIS schools within the trial attended the state's initial training sessions and that these sessions were consistent in format, content, and duration across the two cohorts of trial schools. Because the design was an effectiveness rather than an efficacy trial, we did not have control nor do we have detailed information on the training or planning activities conducted by the SWPBIS teams and staff after they returned to the schools. However, items on the SET do document whether these activities occurred (e.g., creating an action plan, training staff, developing lesson plans to teach the behavioral expectations). Similarly, the behavior support coaches were assigned and supervised by the district staff rather than the research team. Therefore, detailed information on the types and amount of support provided to the coaches and through the coaches to the SWPBIS teams is not available.

We should note that there were some differences in the 5 school districts' resources and infrastructure to support SWPBIS and the coaches; however, we lack sufficient power to test for significant differences in the outcomes by district. The uneven distribution of schools across districts also precludes us from examining district effects and forming strong conclusions (e.g., one district has just two schools participating in the trial, with one school in each condition). However, because the SET does include a measure of district support, we were able to assess districts' levels of resources and infrastructure indirectly. Our analyses indicate that the scores on this subscale increased significantly for the schools randomized to the SWPBIS condition compared to comparison schools. Future researchers should consider using qualitative methods to explore the relationships between school districts' resources and SWPBIS outcomes or consider enrolling a larger number of school districts in research studies to examine the influence of district factors more explicitly.

There is growing interest among researchers in the study of implementation quality, which includes examination of both the extent to which the model is implemented as designed and the type of training, coaching, and technical assistance required to support high quality implementation (Domitrovich et al., 2008). The SWPBIS model provides an ideal framework for examining these issues because of the availability of previously validated measures of implementation quality that include both internal and external assessments (see Bradshaw, Debnam, Koth, & Leaf, in press, for a review) as well as the emphasis on coaching to support high-quality implementation of the model. Additional research is needed to determine the most efficient and effective training and support structure to maximize implementation quality and optimize student and staff outcomes.

This study only examined the impact of training in the universal SWPBIS model. Whereas some schools may have implemented additional supports or programs for students who did not respond to the universal program, these efforts were not systematic across schools. Specifically, during the time in which the trial was conducted, neither the state nor school districts had developed a systematic training or support infrastructure for targeted or indicated preventive interventions.

Nevertheless, it is important to bear in mind that the schools in both conditions likely implemented some other programs to support students, which may have influenced the findings. We attempted to monitor the number and type of different programs implemented in the trial schools. We found that a similar number of programs (other than SWPBIS) were implemented in the trained and nontrained schools and tended to increase over the course of the trial. For example, the comparison schools reported implementing an average of 1.3 programs to prevent behavior and social-emotional problems in the 1st year of the project and 5.8 programs in the final year. Similarly, among the SWPBIS schools, an average of 1.5 additional prevention programs was implemented in the 1st year, whereas an average of 5.1 programs was implemented in the schools in the final year. The most common programs focused on character education and/or development, social-emotional or social skills, bullying prevention, drug prevention (e.g., D.A.R.E.), and conflict resolution and/or peer mediation; however, the implementation quality of these other programs is unknown. Prior research suggests that few programs are implemented with high quality outside of research studies (Gottfredson & Gottfredson, 2002). Anecdotal evidence suggests that the implementation quality of these programs varied considerably across schools. Although it is difficult to identify the cause of this increasing trend in the use of other programs, the growing number of state and national policies emphasizing the use of prevention programs likely contributed to this increasing trend among the schools in the trial. Furthermore, the schools' involvement in the trial and in the data collection efforts may

have increased their interest in and attention to the use of programs to prevent behavior problems and promote a more favorable school climate.

As reported above, prior research by Stuart and Leaf (2007) using propensity scores indicated that the demographic characteristics of the elementary schools participating in the trial were not significantly different from those of the other elementary schools in those districts. Therefore, it seems plausible that the significant intervention effects observed within the trial would generalize to other elementary schools in the state that participated in the same SWPBIS training and support events. Nevertheless, it is important to bear in mind that all the schools in the trial volunteered to participate in the study with the hope that they would be randomized to the trained condition. Schools that are required to implement SWPBIS may not evince the same effects observed in this study. As noted above, schools with more challenging behavior problems, such as schools in urban communities, may experience greater impacts on student outcomes following implementation of SWPBIS. Because these schools were all elementary schools, additional research is also needed to determine the impact of SWPBIS training in secondary schools. The extent to which these findings will generalize to schools in other states, which may have different SWPBIS training or support models, is unknown.

Conclusion

Taken together, the findings of this study suggest that training in SWPBIS is associated with large and sustainable changes in the number and types of schoolwide positive behavior supports provided to students. These results also provide support for the effectiveness of SWPBIS as a method for reducing ODRs and suspensions among elementary school students. Although the effects on student outcomes were comparatively smaller than the effects observed on procedural changes, it is likely that the effects may vary by the school context and/or student-level characteristics. Additional research is needed to identify potential mediators (e.g., teacher efficacy, improved organizational health) (Bradshaw, Koth, et al., 2008) and moderators (e.g., school size, implementation quality) of the effects of SWPBIS on both students and staff. Such work will help researchers determine for whom and under what conditions SWPBIS is most effective. Future research should also explore in greater detail factors contributing to the sustainability of SWPBIS (McIntosh et al., 2009) as well as the systems change that occurs as a result of high quality implementation of SWPBIS (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005).

References

Barrett, S., Bradshaw, C.P., & Lewis-Palmer, T. (2008). Maryland statewide PBIS initiative: Systems, evaluation, and next steps. *Journal of Positive Behavior Interventions*, *10*, 105-114.

Birnbaum, A. S., Lytle, L. A., Hannan, P. J., Murray, D. M., Perry, C. L., & Forester, J. L. (2003). School functioning and violent behavior among young adolescents: A contextual analysis. *Health Education Research*, *18*, 389-403.

Bradshaw, C. P., Debnam, K. J., Koth, C., & Leaf, P. J. (in press). Preliminary validation of the Implementation Phases Inventory for assessing fidelity of schoolwide Positive Behavior Supports. *Journal of Positive Behavior Interventions*.

Bradshaw, C. P., Koth, C. W., Bevans, K. B., Ialongo, N., & Leaf, P. J. (2008). The impact of school-wide Positive Behavioral Interventions and Supports (PBIS) on the organizational health of elementary schools. *School Psychology Quarterly*, *23*, 462-473.

Bradshaw, C. P., Koth, C. W., Thornton, L. A., & Leaf, P. J. (in press). Altering school climate through school-wide Positive Behavioral Interventions and Supports: Findings from a group-randomized effectiveness trial. *Prevention Science*.

Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2008). *The impact of School-wide Positive Behavioral Interventions and Supports (SWPBIS) on children's school-based service use and discipline problems.* Manuscript submitted for publication.

Bradshaw, C. P., Reinke, W. M., Brown, L. D., Bevans, K. B., & Leaf, P. J. (2008). Implementation of School-Wide Positive Behavioral Interventions and Supports (PBIS) in elementary schools: Observations from a randomized trial. *Education & Treatment of Children*, *31*, 1-26.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159.

Domitrovich, C. E., Bradshaw, C. P., Poduska, J., Hoagwood, K., Buckley, J., Olin, S., et al. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion: Training and Practice, Research and Policy*, *1*(3), 6-28.

Doolittle, J. (2006). *Sustainability of positive supports in schools.* Unpublished dissertation, University of Oregon.

Durand, V. M., & Carr, E. G. (1992). An analysis of maintenance following functional communication training. *Journal of Applied Behavior Analysis*, *25*, 777-794.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature* [Florida Mental Health Institute Publication #231]. Tampa: University of South Florida, Louis de la Parte Florida Mental Health Institute, National Implementation Research Network.

Flay, B. R. (1986). Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive Medicine*, 15, 451-474.

Flay, B. R., Biglan, A., Boruch, R. F., González Castro, F., Gottfredson, D., Kellam, S., et al. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science*, *6*, 151-175.

Gottfredson, D. C., & Gottfredson, G. D. (2002). Quality of school-based prevention programs: Results from a national survey. *Journal of Research in Crime and Delinquency*, *39*, 3-35.

Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, *32*, 341-370.

Horner, R. (2009, March 26). *Extending the science, values and vision of Positive Behavior Support*. Speaker at Sixth International Conference on Positive Behavior Support, Jacksonville, FL.

Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A. W., et al. (in press). A randomized, wait-list controlled effectiveness trial assessing School-wide Positive Behavior Support in elementary schools. *Journal of Positive Behavior Interventions*.

Horner, R. H., Sugai, G., Todd, A. W., & Lewis-Palmer, T. (2005). School-wide positive behavior support. In L. Bambara & L. Kern (Eds.), *Individualized supports of students with problem behavior plans* (pp. 359-390). New York: Guilford.

Horner, R., Todd, A., Lewis-Palmer, T., Irvin, L. K., Sugai, G., & Boland, J. (2004). The School-wide Evaluation Tool (SET): A research instrument for assessing school-wide positive behavior supports. *Journal of Positive Behavior Interventions*, *6*, 3-12.

Irvin, L., Tobin, T., Sprague, J., Sugai, G., & Vincent, C. (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.

Knoff, H. M. (2000). Organizational development and strategic planning for the millennium: A blueprint toward effective school discipline, school safety, and crisis prevention. *Psychology in the Schools*, *37*, 17-32.

Lewis, T. J., & Sugai, G. (1999). Effective behavior support: A systems approach to proactive school-wide management. *Focus on Exceptional Children*, *31*, 1-24.

Lindsley, O. R. (1992). Why aren't effective teaching tools widely adopted? *Journal of Applied Behavior Analysis*, *25*, 21-26.

Luiselli, J. K., Putnam, R. F., Handler, M. W., & Feinberg, A. B. (2005). Whole-School Positive Behaviour Support: Effects on student discipline problems and academic performance. *Educational Psychology*, *25*, 183-198.

Maryland State Department of Education. (2003). *Maryland school assessment—Reading: Grades 3*, *5*, *and 8 technical report: 2003 administration*. Baltimore, MD: Author.

Maryland State Department of Education. (2007). *Maryland school assessment—Mathematics: Grades 3 through 8 technical report: 2007 administration*. Baltimore, MD: Author.

May, S., Ard, W., Todd, A., Horner, R., Glasgow, A., Sugai, G., et al. (2003). *Schoolwide information system*. Eugene: University of Oregon, Educational and Community Supports.

McIntosh, K., Horner, R. H., Chard, D., Boland, J., & Good, R. (2006). The use of reading and behavior screening measures to predict non-response to school-wide positive behavior support: A longitudinal analysis. *School Psychology Review*, *35*, 275-291.

McIntosh, K., Horner, R. H., & Sugai, G. (2009). Sustainability of systems-level evidence-based practices in schools: Current knowledge and future directions In W. Sailor, G. Dunlap, G. Sugai, & R. Horner (Eds.), *Handbook of positive behavior support* (pp. 327-352). New York: Springer.

Metzler, C. W., Biglan, A., Rusby, J. C., & Sprague, J. R. (2001). Evaluation of a comprehensive behavior management program to improve school-wide positive behavior support. *Education and Treatment of Children*, *24*, 448-479.

Mrazek, P. J., & Haggerty, R. J. (Eds.). (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington, DC: National Academy Press.

Muscott, H., Mann, E., & LeBrun, M. R. (2008). Positive Behavioral Interventions and Supports in New Hampshire: Effects of large-scale implementation of Schoolwide Positive Behavior Support on student discipline and academic achievement. *Journal of Positive Behavior Interventions*, *10*, 190-205.

National Technical Assistance Center on Positive Behavioral Interventions and Supports. (2007). *Is School-Wide Positive Behavior Support an evidence-based practice?* Eugene, OR: Author. Retrieved December 3, 2008, from http://www.pbis.org/files/101007evidencebase4pbs.pdf

Nersesian, M., Todd, A., Lehmann, J., & Watson, J. (2000). School-wide behavior support through district-level system change. *Journal of Positive Behavior Interventions*, *2*, 244-246.

Scott, T., & Barrett, S. (2004). Using staff and student time engaged in disciplinary procedures to evaluate the impact of school wide PBS. *Journal of Positive Behavior Intervention*, *6*, 21-27.

Spaulding, S., Horner, R., Irvin, L., May, S., Emeldi, M., Tobin, T., et al. (2008, March). *School-wide PBS research and evaluation efforts using large-scale data sources*. Paper presented at the fifth annual International Association for Positive Behavior Supports Convention, Rosemont, IL.

Stuart, E., & Leaf, P. (2007, May). *Moving from efficacy to effectiveness in school-based prevention research: Case study of PBIS*. Poster presented at the Society for Prevention Research, Washington, DC.

Sugai, G., & Horner, R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review*, *35*, 245-259.

Sugai, G., Horner, R., & Gresham, F. (2002). Behaviorally effective school environments. In M. Shinn, H. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems II: Prevention and remedial approaches* (pp. 315-350). Bethesda, MD: National Association of School Psychologists.

Sugai, G., Lewis-Palmer, T., Todd, A., & Horner, R. H. (2001). *School-wide evaluation tool*. Eugene: University of Oregon.

Sugai, G., Todd, A. W., & Horner, R. (2000). *Effective Behavior Support (EBS) Survey: Assessing and planning behavior supports in schools*. Eugene: University of Oregon.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed). Needham Heights, MA: Allyn & Bacon.

Taylor-Greene, S., Brown, D., Nelson, L., Longton, J., Gassman, T., Cohen, J., et al. (1997). School-wide behavioral support: Starting the year off right. *Journal of Behavioral Education*, *7*, 99-112.

Taylor-Greene, S., & Kartub, D. T. (2000). Durable implementation of school-wide behavior support: The high five program. *Journal of Positive Behavior Interventions*, *2*, 233-235.

Catherine P. Bradshaw, PhD, MEd, is an assistant professor in the Department of Mental Health at the Johns Hopkins Bloomberg School of Public Health and the Associate Director for the Johns Hopkins Center for the Prevention of Youth Violence. Her research focuses on the prevention of aggressive and problem behavior in childhood and adolescence and the design and evaluation of school-based prevention and intervention programs.

Mary M. Mitchell, PhD, is an assistant scientist in the Department of Mental Health at the Johns Hopkins Bloomberg School of Public Health. Her research focuses on the evaluation of school-based prevention and intervention programs using latent variable models and hierarchical linear modeling techniques.

Philip J. Leaf, PhD, is a professor in the Department of Mental Health at the Johns Hopkins Bloomberg School of Public Health and the Director for the Johns Hopkins Center for the Prevention of Youth Violence. His research interests include child mental health services, health disparities, school policy and reform, and youth violence prevention.