

Journal of Positive Behavior Interventions

<http://pbi.sagepub.com/>

Illinois Statewide Positive Behavioral Interventions and Supports: Evolution and Impact on Student Outcomes Across Years

Brandi Simonsen, Lucille Eber, Anne C. Black, George Sugai, Holly Lewandowski, Barbara Sims and Diane Myers

Journal of Positive Behavior Interventions 2012 14: 5 originally published online 7 July 2011

DOI: 10.1177/1098300711412601

The online version of this article can be found at:

<http://pbi.sagepub.com/content/14/1/5>

Published by:

Hammill Institute on Disabilities



and



<http://www.sagepublications.com>

Additional services and information for *Journal of Positive Behavior Interventions* can be found at:

Email Alerts: <http://pbi.sagepub.com/cgi/alerts>

Subscriptions: <http://pbi.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

>> [Version of Record](#) - Dec 19, 2011

[OnlineFirst Version of Record](#) - Jul 7, 2011

[What is This?](#)

Illinois Statewide Positive Behavioral Interventions and Supports: Evolution and Impact on Student Outcomes Across Years

Brandi Simonsen¹, Lucille Eber², Anne C. Black³, George Sugai¹, Holly Lewandowski², Barbara Sims⁴, and Diane Myers⁵

Abstract

More than 1,000 Illinois schools are implementing schoolwide positive behavior support (SWPBS) to enhance outcomes for students and staff. Consequently, Illinois established layered support structures to facilitate scaling up SWPBS. This paper describes the development of this infrastructure and presents the results of HLM analyses exploring the effects of implementing SWPBS, with and without fidelity across time, on student behavior and academic outcomes (office discipline referrals, suspensions, and state-wide test scores in reading and math) for a sample of 428 Illinois schools implementing SWPBS. Results indicate that (a) most schools implemented with fidelity and maintained or improved student performance across time and (b) implementation fidelity was associated with improved social outcomes and academic outcomes in math. Study limitations and implications are discussed.

Keywords

schoolwide positive behavior support, scaling up, state-level implementation

Schoolwide positive behavior support (SWPBS) is a systems-level, positive, and preventive approach that results in desired change in student and staff behavior (Bradshaw, Koth, Bevans, Ialongo, & Leaf, 2008; Bradshaw, Mitchell, & Leaf, 2010; Bradshaw, Reinke, Brown, Bevans, & Leaf, 2008; Horner et al., 2009). Schools that implement SWPBS develop a continuum of prevention and intervention strategies, which are typically organized in three tiers (e.g., Walker et al., 1996): universal or primary (Tier 1), targeted-group or secondary (Tier 2), and intensive or tertiary (Tier 3). Descriptions of the three tiers have been published in numerous articles (e.g., Horner & Sugai, 2005; Sugai & Horner, 2002; Walker et al., 1996).

SWPBS relies on district- and state-level support structures to facilitate implementation of SWPBS with fidelity across an increasing number of schools (McIntosh, Horner, & Sugai, 2009; Sugai et al., 2010; Sugai, Horner, & McIntosh, 2008). Implementation fidelity, or the extent to which components of an intervention are implemented as intended (Lane, Bocian, MacMillon, & Gresham, 2004), is critical to achieve desired outcomes. Without fidelity, intervention effectiveness is difficult to judge. In this paper, we (a) describe common district- and state-level features intended to promote implementation fidelity across schools, (b) provide an example of these features in Illinois, (c) discuss gaps in the literature on SWPBS

implementation fidelity, and (d) present the methods and results of a study that aims to document the importance of implementation fidelity as SWPBS is taken to scale.

Common Features of Scaling-Up SWPBS

As of 2010, more than 13,000 schools were implementing SWPBS with support of the Office of Special Education Program's National Technical Assistance Center on Positive Behavior Interventions and Supports (PBIS Center; www.pbis.org). Therefore, issues of scaling up (i.e., increasing the number of schools, districts, and states that are implementing SWPBS with fidelity) are important and timely. In

¹University of Connecticut, Storrs, CT, USA

²Illinois PBIS Network

³Yale University, New Haven, CT, USA

⁴Illinois State Board of Education, Springfield, IL, USA

⁵Assumption College, Worcester, MA, USA

Corresponding Author:

Brandi Simonsen, University of Connecticut, Neag School of Education, 249 Glenbrook Road, Unit 2064, Storrs, CT 06269-2064, USA
Email: brandi.simonsen@uconn.edu

Action Editor: V. Mark Durand

2004, the PBIS Center developed the Schoolwide Positive Behavior Support Implementers' Blueprint and Self-Assessment (subsequently referred to as "SWPBS Blueprint"). The SWPBS Blueprint, which was updated in 2010 (Sugai et al., 2010), guides districts and states in supporting schools as they scale up their implementation.

At the district and state levels, the PBIS Center recommends forming a leadership team to coordinate four primary functions: training, coaching, evaluation, and developing behavioral expertise (Sugai et al., 2010). Initially, the leadership team's primary focus is to create local school/district exemplars. Then, to make sustained implementation possible, the leadership team focuses on securing and maintaining visibility, political support, stable funding, and policy changes that support implementation (see Sugai et al., 2010).

Many states have used this approach to develop district- and state-level structures. As of 2008, 31 states had active state-level leadership teams (Spaulding, Horner, May, & Vincent, 2008). In particular, Maryland, Iowa, and New Hampshire (a) used the SWPBS Blueprint and scaling-up logic to improve behavior supports and student behavior outcomes and (b) published their results in peer-reviewed journals. Participating schools in Maryland (Barrett, Bradshaw, & Lewis-Palmer, 2008), Iowa (Mass-Galloway, Panyan, Smith, & Wessendorf, 2008), and New Hampshire (Muscott, Mann, & LeBrun, 2008) largely implemented SWPBS with fidelity and achieved desired student outcomes (e.g., decreased office discipline referrals, decreased suspensions, increased scores on statewide assessments).

Like these other states, Illinois followed the 2004 SWPBS Blueprint to build state-level support structures. However, the Illinois implementation is unique in that the PBIS initiative was established more than 10 years ago (before the SWPBS blueprint was widely available), and additional layers of support (school, district, region, and state) were established to facilitate implementation across the state. In addition, a larger number of schools were involved in implementation: as of 2008, 804 Illinois schools were implementing SWPBS, whereas other states reported between 0 and 691 schools (Spaulding et al., 2008)

Focus on Illinois PBIS

Illinois has formally supported students' social behavior at the state level for almost 20 years. In 1990, the state established the Emotional and Behavioral Disabilities (EBD) Network. In April 1998, this network expanded its focus from students with EBD to all students by initiating SWPBS training for 23 schools. In 2004, with 520 schools trained and implementing SWPBS, the Network officially changed its name to the Illinois PBIS Network. As of 2009, the Illinois PBIS network was actively involved in training, coaching, and evaluating more than 1,000 schools. The

Illinois PBIS network took the following steps to scale up and promote fidelity of SWPBS implementation across their state.

Infrastructure of Illinois PBIS

Like most states, the Illinois PBIS network provides schools with multiple layers of support. At the school level, a school-based SWPBS team is established; it is composed of a representative group of stakeholders (faculty, staff, parents, and students) and facilitated by a school-based internal coach. Across schools, districtwide teams and external coaches provide additional support to the internal coaches and school-based teams. External coaches typically have more experience and training than internal coaches, and they provide on- and off-site technical assistance. Coaches are funded primarily by school districts and in some cases by intermediate districts, such as Special Education Cooperatives. School-based SWPBS teams, internal coaches, districtwide teams, and external coaches progress through a clear sequence of training activities. These activities focus on the three tiers of prevention and intervention and are delivered across 3 to 5 years (see "Sample Plan for PBIS Implementation" under the "Training" link on www.pbisillinois.org).

At the regional level, PBIS Technical Assistance Coordinators provide direct technical assistance and training to school-based teams, internal coaches, districtwide teams, and external coaches and are supervised by PBIS Technical Assistance Directors (see Illinois PBIS Network, 2009, for full description of the Illinois PBIS infrastructure). At the state level, a statewide PBIS director oversees the regional efforts and coordinates training, coaching, and evaluation activities across regions within the state, with the assistance of her administrative team. To facilitate evaluation activities, the statewide PBIS director accesses the National PBIS Center's database (www.pbseval.org), which is used by coaches and school teams to enter and maintain office discipline referral, self-assessment, and other fidelity data. These data are integrated with a statewide database that stores additional student and school-level data. The combined information is used to develop and disseminate statewide annual evaluation reports as well as brief quarterly reports and newsletters. The statewide PBIS Network leadership participates in and receives direct assistance from the national PBIS Center.

Capacity Building

Across the various infrastructure layers, the programmatic focus is on capacity building (i.e., developing expertise within local and existing personnel). The statewide PBIS director works with regional PBIS leaders to build fluency

with training content among all network staff across the three tiers, including the wraparound process. Specifically, Technical Assistance Directors develop training capacity among regional coordinators, who then co-train with external coaches to build their training capacity, thereby increasing the number of PBIS trainers. External coaches work with school-based internal coaches to build their fluency with providing technical assistance, which increases the number of individuals with the required skills and experience to become external coaches. Over the past ten years, the Illinois PBIS network has used this model to increase the number of schools and internal and external coaches. Given the exponential growth in the number of schools and internal coaches and given that many of those schools have been implementing for multiple years, the Illinois PBIS network presents a unique opportunity to examine the impact of fidelity of SWPBS implementation on student outcomes.

Gaps in the Literature on SWPBS Implementation Fidelity

In Illinois and other states, systems (e.g., coaching, training, and evaluation structures) were put in place to promote and assess the fidelity of SWPBS implementation across schools. Implementation fidelity is critical when evaluating intervention effectiveness. Despite the importance of implementation fidelity, relatively few studies report fidelity data (Lane, Kalberg, Bruhn, Mahoney, & Driscoll, 2008). Researchers conducting large-scale studies on SWPBS have begun to address this gap in the literature by collecting and reporting fidelity data (e.g., Bradshaw et al., 2010; Bradshaw, Koth, et al., 2008; Bradshaw, Reinke, et al., 2008; Horner et al., 2009). However, even fewer studies have directly examined the impact of implementation fidelity on student outcomes. After performing a search of the PsychInfo and ERIC databases for the descriptors of “school wide positive behavior support” and “fidelity,” it appeared that no published peer-reviewed studies have systematically examined the effects of implementing SWPBS with (and without) fidelity over time.

This paper aims to address this gap in the literature. We present a longitudinal analysis of data from Illinois schools that implemented SWPBS with varying levels of fidelity across 7 years (from 2000–2001 through 2007–2008 school years). Specifically, this study was conducted to address the following research question: What is the relationship between SWPBS implementation fidelity and school-level student behavior and academic outcomes (as measured by office discipline referrals, out of school suspensions, total suspensions, and the percentage of students who met or exceeded benchmarks on statewide reading and math tests) across time? In addition, we explored the proportion of schools implementing with fidelity across time.

Method

Sample

The sample consisted of the subset Illinois schools that implemented SWPBS from 2000 to 2008 and entered data into the state-wide data base. Altogether, the sample consisted of 428 schools nested within 125 districts; not all schools had data for all years, and for many schools, data were only available for 1 or 2 years. This sample included 274 elementary (K–6), 46 K–8, 91 middle (6–9), and 17 high (9–12) schools.

The number of schools with available data increased over the study period. Subsets of districts and schools with available *Schoolwide Evaluation Tool* (SET; Sugai, Lewis-Palmer, Todd, & Horner, 2001) data differed for each targeted outcome. For the suspension analyses, which studied days out of school (OSS) and total number of suspensions (TS), 242 schools nested within 81 districts had corresponding SET data for 6 years (2001/2002–2007/2008). For office discipline referrals (ODRs), data for the same 6 years were available from 400 schools nested within 118 districts. Finally, Illinois State Achievement Test data for schools with SET data were available over 7 years (2000/2001–2007/2008) from 324 schools within 101 districts. Table 1 summarizes descriptive data for sample schools across study years (counts, means, and standardized mean-difference effect sizes).

Study Variables

Data were retrieved from existing databases maintained by the Illinois State Board of Education (ISBE) and the Illinois PBIS network, in collaboration with the national PBIS Center (www.pbseval.org). In addition, general information about which schools were implementing SWPBS (number of schools implementing, name of schools, etc.) was maintained by the Illinois PBIS network on a locally managed data system. The SET (Sugai et al., 2001) was used to document fidelity of SWPBS implementation (independent variable), and the following measures were used as indicators of student social behavior and academic outcomes (dependent variables): office discipline referrals, OSS, TS, and *Illinois State Achievement Test* (ISAT) scores in reading and math.

Schoolwide Evaluation Tool (SET; Sugai et al., 2001). The SET is a reliable and valid measure of the first tier of SWPBS implementation. The SET comprises seven subscales: behavior expectations defined, behavioral expectations taught, reward system, violation system, monitoring and evaluation, management, and district support. Indices of internal consistency ($\alpha = .96$), test–retest reliability (average of 97.3% agreement), and interobserver agreement (average of 99.0% agreement) were high; indices of construct

Table 1. Mean Outcome Scores and Standardized Group Mean Difference Effect Sizes for Each Study Variable by SET Status and Year (2000–2008)

	2000–2001	2001–2002	2002–2003	2003–2004	2004–2005	2005–2006	2006–2007	2007–2008	Overall
ODRs per 100 students									
SET = N		1.06	1.08	0.75	0.79	0.80	0.91	1.03	0.89
<i>n</i>		9	21	21	42	37	37	28	
SET = Y		0.88	0.54	0.67	0.53	0.71	0.70	0.48	0.60
<i>n</i>		5	24	36	32	54	83	104	
Effect size		0.14	0.50	0.12	0.39	0.11	0.22	0.51	0.32
Total suspensions (No. of suspension events)									
SET = N		64.22	114.10	86.38	170.84	135.04	77.63	221.50	132.26
<i>n</i>		9	21	21	25	24	19	22	
SET = Y		90.80	77.67	68.44	51.00	47.03	68.02	54.16	61.31
<i>n</i>		5	24	36	18	36	58	68	
Effect size		–0.41	0.23	0.12	0.39	0.33	0.07	0.54	0.31
Out-of-school suspensions (total no. of days suspended out of school)									
SET = N		191.38	298.75	224.42	495.48	408.95	202.06	665.95	373.15
<i>n</i>		8	20	19	21	20	18	19	
SET = Y		161.4	197.14	194.20	123.07	112.72	154.65	108.31	142.96
<i>n</i>		5	22	30	14	32	55	64	
Effect size		0.14	0.22	0.09	0.46	0.39	0.13	0.66	0.37
ISAT Math (% students MOE grade-level mastery criteria)									
SET = N	71.33	61.07	63.61	75.07	64.82	78.26	76.85	77.58	72.89
<i>n</i>	6	27	20	33	71	74	85	60	
SET = Y	58.00	67.91	67.53	74.56	76.45	83.00	83.94	81.11	80.63
<i>n</i>	2	6	18	33	66	96	159	162	
Effect size	–1.82	0.38	0.21	–0.04	0.59	.36	0.41	0.21	0.43
ISAT Reading (% students MOE grade-level mastery criteria)									
SET = N	66.00	56.80	58.49	70.45	64.82	68.47	67.51	69.49	66.05
<i>n</i>	6	27	20	33	71	74	85	60	
SET = Y	53.25	60.59	54.71	67.11	67.85	71.33	71.64	72.15	70.25
<i>n</i>	2	6	18	33	66	96	159	162	
Effect size	–1.40	0.25	–0.27	–0.28	0.35	0.20	0.25	0.17	0.27

Note. Because variables were not normally distributed, means and mean differences should be interpreted with caution; group differences may be underrepresented. ISAT = Illinois State Achievement Test; SET = Schoolwide Evaluation Tool; MOE = meeting or exceeding.

validity (the SET correlated with the Effective Behavior Support: Self-Assessment Survey, $r = 0.75$, $p \leq .01$) were acceptable; and the SET was demonstrated to be sensitive to change over time (statistically significant differences were observed for schools across time; Horner et al., 2004). Although the SET is the standard metric for evaluating fidelity of Tier 1 implementation for research (Algozzine et al., 2010), new analyses suggest that reliability and validity indices of the SET may vary by school level (Vincent, Spaulding, & Tobin, 2010).

The SET was completed in schools participating in the Illinois PBIS network by trained outside evaluators (i.e., regional Technical Assistance Coordinators, external coaches, or retired administrators who were not directly involved with the school). Evaluators participated in a 1-day SET training and shadowed another trained evaluator before conducting SETs independently. SETs were completed for each school after initial training in PBIS and then once a year thereafter, typically in the spring. Overall and subscale SET scores were entered into the SWPBS evaluation

site (pbseval.org) for each school. As with any applied large-scale evaluation, there were factors that interfered with SET completion (e.g., administrative changes, scheduling conflicts); therefore, some schools do not have consistent SET scores. Previous research demonstrated that schools that meet 80% of fidelity criteria on the overall scale (Horner et al., 2004) and 80% of fidelity criteria on the subscale related to teaching expectations have the most positive student outcomes (Doolittle, 2006). Therefore, the “80/80” criterion was used to sort schools into two groups for each year: those that did (met or exceeded 80/80) and those that did not (i.e., lower than 80/80) implement with fidelity.

Office Discipline Referral (ODR). ODR data were entered into the schoolwide information system (www.swis.org) by each school and were accessed by researchers through an interface with the www.pbseval.org site. ODR data were stripped of identifiers and reported by school. ODR data have been demonstrated to be reliable and valid when used for within-school comparisons across time (Irvin et al., 2006; Irvin, Tobin, Sprague, Sugai, & Vincent, 2004). ODR values represent the number of ODRs per 100 students. The distribution of ODR scores was positively skewed (skewness/SE ratio ranged by year from 5.28 to 31.05), with low frequencies of events occurring for most schools. To more closely meet the assumptions of the analytical model for normally distributed dependent variables, the square roots of ODR scores were used (Tabachnick & Fidell, 2001).

Total Number of Suspensions and Out-of-School Suspension Days (TS and OSS). TS and OSS data were reported by each school. School-level data were entered into the www.pbseval.org database. TS reflects the number of suspension events whereas OSS reflects the number of days suspended. Both variables represent count data, and both TS and OS were positively skewed (skewness/SE ratios ranged by year from 5.56 to 17.55, and 4.66 to 14.93, respectively), with low frequencies of events occurring for most schools. HLM software (version 6.06; Raudenbush, Bryk, & Congdon, 2004) allows specifying count data (Poisson distribution); thus data were not transformed.¹

Illinois State Achievement Test (ISAT). The IL State Board of Education maintained ISAT scores, and we used school-level aggregate data as the intervention was administered at the school level. In particular, we examined the percentage of students meeting or exceeding (MOE) grade-level mastery criteria within each school. ISAT scores for reading and math are posted for Grades 3, 5, and 8 as academic benchmark years. For schools with multiple grades of ISAT (e.g., K–8 grade schools that have data for ISAT in Grades 3, 5, and 8), the mean percentage MOE across grades was used. High school data represent the percentage of students MOE Grade 11 mastery criteria. Reading and math percentage MOE scores were negatively skewed (skewness/SE

ratio ranged by year from –3.38 to –8.21, and –4.72 to –15.37, respectively), with higher percentages occurring most frequently. To normalize the distribution to meet model assumptions, these percentage scores were squared.

Procedures for Data Analysis

With fidelity of SWPBS implementation as a primary variable of interest, rates of fidelity were analyzed as the proportion of schools meeting 80/80 criteria on the SET across years of the study. Because subsets of schools and districts differed according to outcome variables, and schools had different identifiers in each data set, rates of implementation were calculated separately by outcome category. Proportions of implementation fidelity were calculated separately for ODR, suspension (OSS and TS), and academic (reading and math) variables. Yearly proportions were then averaged across outcomes (accounting for different numbers of schools by outcome) to provide a single proportion of implementation for each year.²

We modeled each outcome variable over time using hierarchical linear modeling (HLM) and nonlinear modeling. These analytical techniques allowed for the estimation of models, accounting for the nonindependence of repeated measures within schools and schools within districts. HLM software (Version 6.06; Raudenbush, Bryk, & Congdon, 1994–2008) was used to estimate three-level hierarchical models, representing observation years (Level 1 scores) nested within schools (Level 2 scores), and schools nested within districts (Level 3 scores). First, an empty model (no predictors) was specified for each variable to determine the intraclass correlation coefficients (ICCs), or the proportions of variance in scores explained by schools and districts. (All ICCs were greater than 0, indicating that data were not independent.) Then, each analytical model included year (centered at the first observation year, such that the intercept for each model represented Year 1 scores) and SET as a time-varying covariate (a variable whose value could differ at each time point for a given school). Because of the coding of variables, the Level 1 intercept estimate represented the estimated score in the first year for a school below the SET threshold (not implementing SWPBS with sufficient fidelity). Given constraints of limited data, intercepts were only estimated at Levels 2 (school level) and 3 (district level), and no school- or district-level predictors were used. We estimated the variance around each of the Level 2 and 3 intercept estimates (see models in appendix). Power was not sufficient to estimate other random effects (e.g., variances around the estimates for effects of Year and SET); other variance components were fixed at zero. For transformed ODR and ISAT variables, the distribution was specified as normal in the model. Distributions of OSS and TS were specified as Poisson with over-dispersion (i.e., count

data with variances larger than their respective means). Full maximum likelihood estimation (the default estimation procedure in HLM 6.06 with three-level models) was used to estimate all models.

Results

Model results were favorable overall, revealing maintenance or improvement in outcomes over time for all schools in the fidelity of SWPBS implementation, and demonstrating in most cases an association between SWPBS implementation fidelity and significantly better outcomes for those schools across years (see Figure 1).

Solid line represents schools not meeting 80/80 fidelity criteria, and broken line represents schools meeting 80/80 criteria.

Fidelity

Percentages of schools meeting or exceeding SET criteria (averaged across outcome variables) increased over the course of the study from 36% of schools in 2000–2001 to 78% of schools in 2007–2008. Fidelity differed by school type. Elementary (K–6) schools, the most highly represented school type in the sample, implemented with the highest rate of fidelity at 81% by 2007–2008. That is, by the end of the study, 8 of every 10 elementary schools in the sample met 80/80 criteria on the SET. High schools (Grades 9–12) had the lowest rates of implementation at 31% in 2007–2008. However the number of high schools was remarkably low; fewer than 10 high schools were represented in any of the data sets in any given year. Notably, 73% of all middle schools (Grades 6–9) implemented with fidelity by the final year of the study.

Office Discipline Referrals

The proportion of variance in ODR scores explained by schools at Level 2 was .23 and by districts at Level 3 was .56; that is, approximately 23% and 56% of the total variance in ODR data was accounted for in school and district differences, respectively.

Because the square root of scores was used, interpretation of the intercept in this model was not meaningful. However, model results indicated that ODR decreased significantly over time (see Figure 1[a]); the effect of years was statistically significant and negative ($\gamma_{200} = -0.04$, $p = .001$; Table 2). In addition, we noted a statistical trend that ODR levels across years were lower for schools that implemented SWPBS with fidelity ($\gamma_{100} = -0.06$, $p = .086$).

Suspension Data

The proportion of variance accounted for by school and district grouping was less than 1% for both OSS and TS.

Frequencies of suspensions were distributed differently by year, with greater variability in the number of suspensions for some years than others. As a result, most of the variability in scores was accounted for by school year, rather than by school or district.

Out-of-School Suspensions (OSS). The average rate of OSS did not change significantly over time ($\gamma_{200} = 0.01$, $p = .771$), despite fluctuation in distribution of frequencies by year. However, schools implementing SWPBS with fidelity had significantly lower rates of OSS ($\gamma_{100} = -0.19$, $p = .002$) than other schools in the sample (see Figure 1[b]). The number of OSS days estimated by the statistical model for schools below SET fidelity criteria in 2001–2002 was 68, and for schools implementing with fidelity, was 57. This represents a difference in event rate of 17% (Table 2).

Total suspensions. The results of the TS model were similar to those for OSS. The rate of yearly TS did not change significantly over time ($\gamma_{200} = 0.02$, $p = .330$), but schools implementing SWPBS with fidelity had significantly lower rates of suspensions ($\gamma_{100} = -0.15$, $p = .007$) than schools not implementing with fidelity (see Figure 1[c]). The estimated number of suspension events for schools below SET criteria in 2001–2002 was 27, and for schools meeting SET criteria it was 24, a 12% difference (Table 2).

ISAT

The proportions of variance in reading scores accounted for by schools and districts were .33 and .41, respectively. That is, school differences explained 33% of the variability in reading scores, whereas district differences explained 41% of score variability. Proportions differed slightly for math, with 30% and 34% of the variance in math scores explained by the school and district, respectively.

Reading. For all schools in the sample, the average percentage of students MOE grade-level mastery criteria for ISAT Reading tests increased significantly over the course of the study ($\gamma_{200} = 281.64$, $p < .001$). There was no statistically significant difference on this outcome between schools that met SET fidelity criteria and those that did not ($\gamma_{100} = 61.85$, $p = .411$; see Table 2 and Figure 1[d]).

Math. Across all schools in the sample, the average percentage of students in Grades 3, 5, and 8 who MOE grade-level mastery criteria on the ISAT Math test increased significantly over time ($\gamma_{200} = 402.58$, $p < .001$). However, consistent with the behavioral outcome variables, schools that implemented SWPBS with fidelity had significantly higher percentages of students overall who mastered the Math test ($\gamma_{100} = 244.65$, $p = .009$; see Figure 1[e]; also see Table 2 and observed differences in percentages by year in Table 1).

Discussion

The large scale of SWPBS implementation in Illinois presented a unique opportunity to explore the effects of implementing

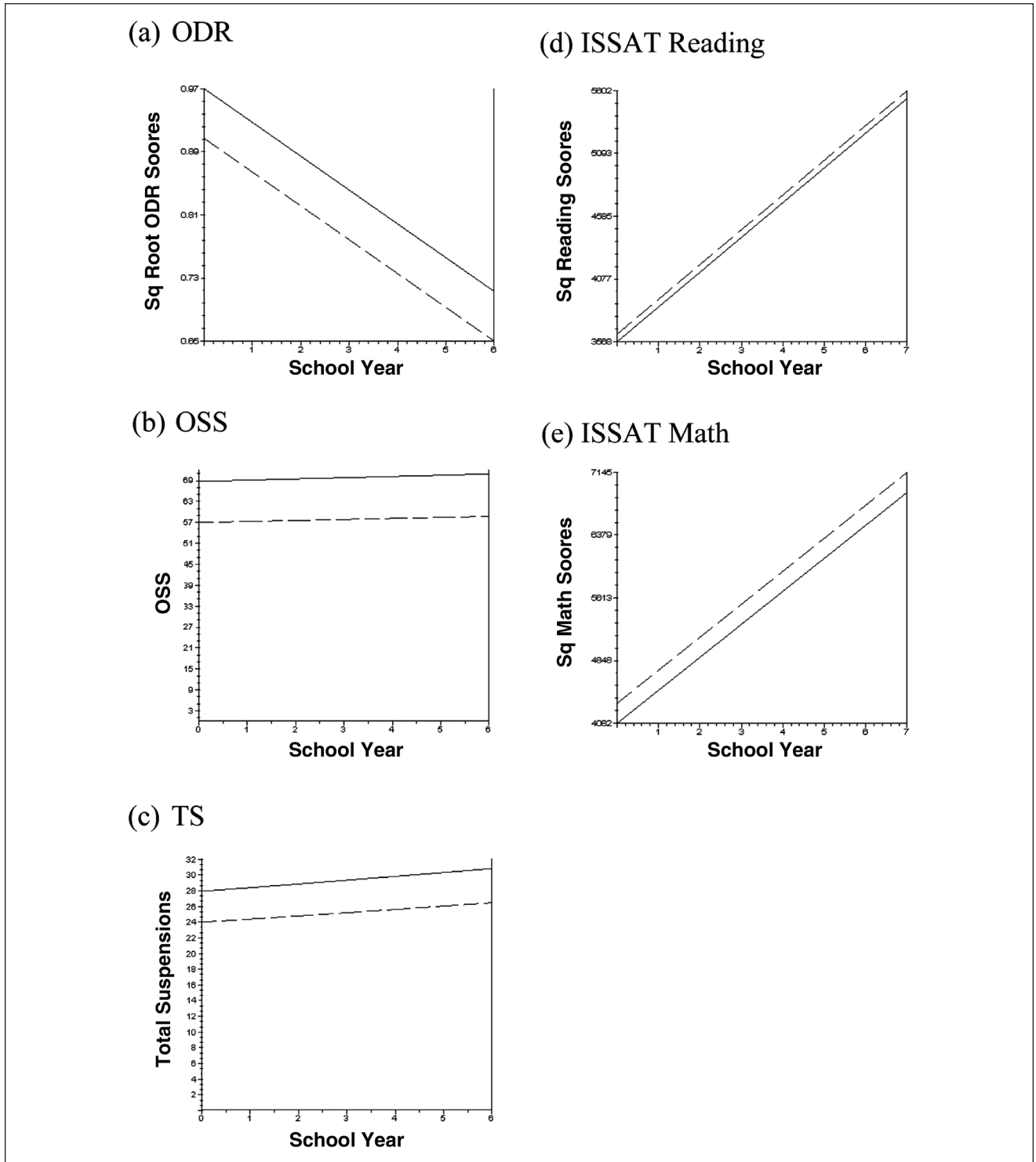


Figure 1. Effects of fidelity and time for each study variable

Note. Solid line represents schools not meeting 80/80 fidelity criteria, and broken line represents schools meeting 80/80 criteria.

SWPBS with and without fidelity. In this article, we used HLM to examine the effects of implementing SWPBS with and without fidelity across time for 428 Illinois schools that implemented SWPBS and reported data. In this section, we discuss study results, limitations, and implications.

Interpretation of Study Results

All schools in the Illinois study sample were implementing SWPBS, and the percentage of schools implementing with fidelity increased throughout the study (from 36% to 78%).

Table 2. Model Estimates for all Dependent Variables

Effect	ODR ^a		OSS ^b		TS ^c		ISAT Reading ^d		ISAT Math ^e	
	df	Estimate (SE)	df	Estimate (SE)	df	Estimate (SE)	df	Estimate (SE)	df	Estimate (SE)
Intercept γ_{000}	117	0.974 (0.05) ***	80	4.23 (0.17) ***	80	3.33 (0.17) ***	100	3568.48 (175.46) ***	100	4081.87 (191.85) ***
SET γ_{100}	305	-0.06 (0.04)*	344	-0.19 (0.06) ***	383	-0.15 (0.06)**	887	61.85 (75.24)	887	244.65 (92.33) ***
School year γ_{200}	305	-0.04 (0.01) ***	344	0.01 (0.02)	383	0.02 (0.02)	887	281.64 (20.99) ***	887	402.58 (25.84) ***
Level 1 variance σ^2		0.037						633421.85		990632.92
Level 2 variance $\tau\pi$	189	0.064 ***	111	2.03 ***	118	1.77 ***	212	1138284.27 ***	212	1284177.60 ***
Level 3 variance $\tau\beta$	82	0.031 ***	51	0.28*	54	0.40 ***	99	1237007.20 ***	99	1140951.62 ***

Note. ODR = office discipline referral; OSS = out-of-school suspension; TS = total number of suspensions; ISAT = Illinois State Achievement Test; SET = Schoolwide Evaluation Tool.

a. Number of schools entered in the model = 400; districts = 118.

b. Number of schools entered in the model = 242; districts = 81.

c. Number of schools entered in the model = 242; districts = 81.

d. Number of schools entered in the model = 324; districts = 101.

e. Number of schools entered in the model = 324; districts = 101.

* $p < .10$. ** $p < .05$. *** $p < .01$.

Schools demonstrated a statistically significant improvement on all social behavior and academic outcome measures. Across measures, statistically significant improvements were associated with (a) time (ISAT Reading), (b) implementation fidelity (OSS and TS), or (c) both time and fidelity (ODRs and ISAT Math). Implementing SWPBS with fidelity was related to better rates of all social behavior outcomes and higher average percentages of students MOE mastery criteria for math, but was not related to the average percentage of students MOE mastery levels in reading.

SWPBS primarily supports programming for improvement in social behavior (i.e., prompting, teaching, reinforcement); therefore, it was not surprising to find that schools that implemented SWPBS with fidelity also experienced lower rates of ODRs, OSS, and TS. Similar results have been reported in randomized controlled trials (Bradshaw et al., 2010; Bradshaw, Koth, et al., 2008; Bradshaw, Reinke, et al., 2008; Horner et al., 2009) and other descriptive studies of state implementation (e.g., Barrett et al., 2008; Mass-Galloway et al., 2008; Muscott et al., 2008). In addition, regardless of fidelity, schools experienced decreases in ODRs over time, but time alone did not lead to overall decreases in OSS or TS. One explanation may be that student behaviors that resulted in an ODR may be sensitive to change and responsive to SWPBS implementation, regardless of fidelity. In contrast, student behaviors that resulted in suspension remained intractable with lower SWPBS fidelity and only appeared to decrease with higher levels of fidelity.

Schools that implemented SWPBS with fidelity also had a larger portion of students who MOE benchmarks on the ISAT test in math than schools that did not implement with fidelity. As student behavior improved, teachers may have been more able to reallocate their resources (time, attention, planning) from managing problem behavior to teaching academic content. For example, Scott and Barrett (2004) performed a cost-benefit analysis to explore the amount of

instructional time saved by teachers at an elementary school as a result of implementing SWPBS and decreasing ODRs. They assumed that (a) each ODR corresponded to approximately 20 min of lost instructional time for a student and (b) each student experienced a 6-hr (360-min) school day. They found, on average, 79.5 days of instructional time were saved per year as a result of implementing SWPBS.

In contrast, schools that implemented SWPBS, regardless of fidelity, increased the proportion of students who MOE benchmarks on the ISAT test in reading over time. In other words, fidelity of implementation did not affect ISAT scores on reading in a statistically significant manner. This is consistent with earlier research showing improvements in math but not reading. Muscott et al. (2008) found improvements in math (i.e., the percentage of students achieving proficiency [basic or above]) for 73% of schools implementing with fidelity; however, they only found improvements in reading for 41% of schools implementing with fidelity. In contrast, Horner and colleagues (2009) reported a tentative finding that SWPBS implementation was related to an increase in the proportion of third-graders who met or exceeded the state standard in reading. Specifically, they noted statistically significant improvement (conditional effects) of (a) treatment schools over time and (b) treatment schools over control schools at one time point. However, the overall (interaction) effects of treatment and time were not statistically significant, and they suggest readers should interpret conditional effects "with caution" (p. 140). Further research is needed to understand the relationship between the fidelity of SWPBS implementation and students' academic outcomes.

Limitations

The results from this study should be interpreted in light of the following limitations. First, the number of schools varied

across outcome measures, which may have resulted in less statistical power to detect statistically significant effects for some outcome measures (e.g., OSS and TS). As a result, we acknowledged in this study all effects that were significant at $\alpha = .10$ and below. In addition, smaller sample sizes limited our ability to model random effects (e.g., slopes for year and SET). The appropriateness of assuming the same effect of year and SET fidelity on outcome measures for every school is unclear, and future studies should assess whether and how these effects may vary across schools. Specifically, it is possible that schools changed at different rates over time, and benefited differently from fidelity to SWPBS. Models with sufficient power to estimate school-specific effects should be tested. Furthermore, maximum likelihood estimation is a large-sample technique. With relatively small and unbalanced numbers of observations at Levels 1 and 2, standard errors for fixed effects estimates may have been underestimated (Raudenbush & Bryk, 2002), potentially inflating the Type I error rate.

Second, although the sample sizes utilized in these analyses enabled statements about general patterns, statements about how individual schools responded to SWPBS implementation should be made with caution. For example, we reported the mean percent of students MOE grade-level mastery criteria on the ISAT across grades for schools with multiple grades participating in the ISAT. This procedure may have obscured grade-level differences, and findings should not be generalized to all schools or all grade levels. In addition, the number of implementing high schools meeting inclusion criteria was relatively low, and broad statements about the relationship between SWPBS implementation and student outcomes in high schools are therefore more difficult to make. The general direction of results suggests that SWPBS is a viable means for improving disciplinary climate and enhancing academic outcomes.

Third, because levels of implementation and specific features of the SWPBS approach were contextualized to Illinois policy, resources, training curricula, needs, etc., generalizations of the findings from the Illinois experience to other state implementation efforts should be made with careful consideration of similarities and differences across schools, districts, and states.

Fourth, only 428 of approximately 1,000 schools met criteria for inclusion in this analysis. Thus, interpretations of the results from this study should not necessarily be extended to all IL PBIS schools. That is, outcomes and interpretations may have been different if data from excluded schools had been included, and it is unclear whether and how schools without data differ from the schools included in these analyses. However, the patterns represented by this sample of 428 schools suggest the importance of implementation fidelity for academic and

social behavior outcomes, especially in elementary and middle schools.

Implications

The results from this study suggest the importance of implementing SWPBS with fidelity. Specifically, implementing SWPBS with fidelity was associated with increases in ISAT scores in math and decreases in ODRs, OSS, and TS. Therefore, our results have a number of important implications for research, practice, and policy.

Implications for research. First, our findings extend the literature base on the importance of collecting and reporting data on the fidelity of SWPBS implementation, as implementing with lower levels of fidelity is frequently related to differential outcomes over time. Therefore, researchers should ensure that fidelity data are included in their research plans and publication reports.

Second, this article provides a springboard for additional research on the effects of implementing SWPBS with fidelity on academic outcomes. In this study, fidelity of SWPBS implementation was related to larger proportions of students who met or exceeded state benchmarks in math, but not in reading. However, as discussed, research in this area is still emerging, and deserves further investigation. Future research should take curricular and instructional variables into consideration when evaluating the impact of implementing SWPBS with fidelity. Similarly, future research efforts may collect academic data that are more closely aligned with the daily, weekly, and monthly curriculum. For example, at the elementary level, curriculum-based assessment or measurement (CBA or CBM, respectively) information on literacy and numeracy performance progress could be collected to determine if there is a stronger relationship between SWPBS fidelity and these academic outcomes than annual state test scores.

Third, in this study, typical school-level disciplinary data were collected (i.e., ODR, OSS, TS), which often required an administrator's response or action for data to be recorded (ODR) or generated (OSS and TS). In other words, an administrator decides whether to enter an ODR into the database and issue a suspension; therefore, these metrics may be underestimates of the actual levels of problem behavior in a school. To provide a greater depth of understanding of the specific effects of SWPBS implementation, future research should collect and examine information related to more immediate behavioral and social climate variables (e.g., academic engagement, opportunities to respond, scheduled/allocated academic time, disruptive behavior).

Fourth, to make generalizable statements about the usefulness of the Illinois implementation design and outcomes, future research should examine similar longitudinal efforts

in other states and ensure sufficiently large and balanced samples at all levels of the hierarchical model. These more ideal data conditions would provide a platform for replication of the current analyses and would allow for the estimation of more complex models, including random effects and school- and district-level predictors of outcomes. In addition, more definitive statements about causal directions and relationships will require researchers to implement different research designs (e.g., randomized controlled trials that include/model potential mediator and moderator variables) to enable more direct statements about causal influences.

Fifth, the Illinois PBIS network used the elements (e.g., leadership team, coordination, training, coaching, evaluation) of the SWPBS Implementation Blueprint to guide its statewide efforts. The inference in this study is that these systemic elements are important, related to implementation fidelity, and equally influential. As these inferences have not been systematically studied, future studies may examine the state level organizational factors that contribute to effects at the student and school levels. In addition, the Illinois PBIS network adapted application of these supports to fit the regional, district, and school characteristics of the Illinois school system; therefore, adaptations to these organizational factors may also be considered in future studies.

Implications for practice and policy. Results from this study offer important considerations for implementers and policy makers at the local, state, and national levels who are working to scale up SWPBS implementation with fidelity. Given the improved outcomes associated with accurate and consistent implementation, implementers should attend to factors that affect implementation fidelity and policy makers should establish guidelines, contingencies, and rewards for high levels of implementation fidelity. For example, in Illinois, a comprehensive and dedicated investment in data collection and evaluation enables formative and immediate programmatic enhancements, visible demonstrations of accomplishments, and means for acknowledging successful SWPBS school and district implementers. Results from the Illinois implementation are evidence of improved student outcomes when implementation is accurate, consistent, and sustained.

Conclusion

The Illinois SWPBS implementation presented an exciting and informative opportunity for a longitudinal state-level evaluation. In general, the results of HLM analyses document that implementing SWPBS with fidelity is associated with better student social behavior (office discipline referrals, suspensions) and academic (statewide test scores in math) outcomes for Illinois schools. Although preliminary, the results from this research suggest that other system-level implementers of SWPBS should invest in structures and capacities that maximize adoption of

evidence-based practices and promote implementation fidelity.

Appendix

Example Three-Level Model

$$\text{Level 1: } Y_{ijk} = \pi_{0jk} + \pi_{1jk}(SET_{ijk}) + \pi_{2jk}(YEAR_{ijk}) + e_{ijk}$$

$$\text{Level 2: } \pi_{0jk} = \beta_{0ok} + r_{0jk}$$

$$\pi_{1jk} = \beta_{1ok}$$

$$\pi_{2jk} = \beta_{2ok}$$

$$\text{Level 3: } \pi_{0ok} = \beta_{00D} + r_{00k} + u_{0Dk}$$

$$\beta_{1Dk} = r_{10D}$$

$$\beta_{2Dk} = r_{20D}$$

$$\text{Mixed Model: } Y_{ijk} = \gamma_{000} + \gamma_{100}(SET_{ijk}) + \gamma_{200}(YEAR_{ijk}) + r_{0jk} + u_{00k} + e_{ijk}$$

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The development of this article was supported in part by Grant H029D40055 from the Office of Special Education Programs, U.S. Department of Education for OSEP Center on Positive Behavioral Interventions and Supports (www.pbis.org). Opinions expressed herein are the author's and do not necessarily reflect the position of the U.S. Department of Education, and such endorsements should not be inferred.

Notes

1. Because ODR data were analyzed as number of events per 100 students (and scores represented proportion, not count/integer data), the Poisson distribution did not apply to this variable. Instead, we transformed ODR scores to approximate a normal distribution. ISAT Math and Reading percent scores were also transformed for this reason.
2. To illustrate the calculation of average proportion of schools meeting or exceeding SET fidelity thresholds, consider data from the 2007–2008 school year. The ODR database had 86 K-6 schools with 82% meeting fidelity criteria; the ISAT database had 55 K-6 schools with 80% meeting fidelity criteria; and the Suspension database had 114 K-6 schools with 81% meeting fidelity criteria. The average percentage of K-6 schools meeting fidelity criteria was calculated as $[(86 \times 82\%) + (55 \times 80\%) + (114 \times 81\%)] / 255 \text{ schools total} = 81\%$.

References

- Algozzine, B., Horner, R. H., Sugai, G., Barrett, S., Dickey, S. R., Eber, L., . . . Tobin, T. (2010). Evaluation blueprint for school-wide positive behavior support. Eugene, OR: National Technical

- Assistance Center on Positive Behavior Interventions and Support. Retrieved from www.pbis.org
- Barrett, S. B., Bradshaw, C. P., & Lewis-Palmer, T. (2008). Maryland statewide PBIS initiative; Systems, evaluation, and next steps. *Journal of Positive Behavior Interventions, 10*, 105–114.
- Bradshaw, C., Koth, C., Bevans, K., Jalongo, N., & Leaf, P. (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., Mitchell, M., & Leaf, P. (2010). Examining the effects of school-wide positive behavioral interventions and supports on student outcomes: Results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions, 12*, 133–148.
- Bradshaw, C., Reinke, W., Brown, L., Bevans, K., & Leaf, P. (2008). Implementation of school-wide positive behavioral interventions and supports (PBIS) in elementary schools: Observations from a randomized trial. *Education and Treatment of Children, 31*, 1–26.
- Doolittle, J. (2006). Sustainability of positive supports in schools. Unpublished dissertation, University of Oregon, Eugene.
- Horner, R. H., & Sugai, G. (2005). School-wide positive behavior support: An alternative approach to discipline in schools. In L. Bambara & L. Kern (Eds.), *Positive behavior support* (pp. 359–390). New York, NY: Guilford.
- Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A., & 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.
- Horner, R. H., Todd, A. W., Lewis-Palmer, T., Irvin, L. K., Sugai, G., & Boland, J. B. (2004). The School-Wide Evaluation Tool (SET): A research instrument for assessing school-wide positive behavior support. *Journal of Positive Behavior Interventions, 6*, 3–12.
- Illinois PBIS Network. (2009). Chapter 2: PBIS implementation in Illinois. In *Illinois Positive Behavioral Interventions and Supports (ILPBIS) Network: 2008–2009 Progress Report* (pp. 5–16). Illinois PBIS Network. Retrieved from http://www.pbisillinois.org/Online_Library
- Irvin, L. K., Horner, R. H., Ingram, K., Todd, A. W., Sugai, G., Sampson, N. K., & Boland, J. (2006). Using office discipline referral data for decision making about student behavior in elementary and middle school: An empirical evaluation of validity. *Journal of Positive Behavior Interventions, 8*, 10–23.
- Irvin, L. K., Tobin, T. J., Sprague, J. R., Sugai, G., & Vincent, C. G. (2004). Validity of office discipline referral measures as indices of school-wide behavioral status and effects of school-wide behavioral interventions. *Journal of Positive Behavioral Interventions, 6*, 131–147.
- Lane, K. L., Bocian, K. M., MacMillan, D. L., & Gresham, F. M. (2004). Treatment integrity: An essential-but often forgotten-component of school-based interventions. *Preventing School Failure, 48*, 36–43.
- Mass-Galloway, R. L., Panyan, M. V., Smith, C. R., & Wessendorf, S. (2008). Systems change within school-wide positive behavior supports; Iowa's work in progress. *Journal of Positive Behavior Interventions, 10*, 129–135.
- 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, R. Horner, & G. Sugai (Eds.), *Handbook of positive behavior support* (Chap. 14, pp. 327–352). New York, NY: Springer.
- Muscott, H. S., Mann, E. L., & LeBrun, M. R. (2008). Positive Behavioral Interventions and Supports in New Hampshire; Effects of large-scale implementation of school wide positive behavior support on student discipline and academic achievement. *Journal of Positive Behavior Interventions, 10*, 190–205.
- Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2004). *Hierarchical Linear and Nonlinear Modeling (HLM) for Windows (Version 6.06)* [Computer Software]. Lincolnwood, IL: Scientific Software International.
- Scott, T. M., & Barrett, S. B. (2004). Using staff and student time engaged in disciplinary procedures to evaluate the impact of school-wide PBS. *Journal of Positive Behavior Interventions, 6*, 21–27.
- Spaulding, S. A., Horner, R. H., May, S. L., & Vincent, C. G. (2008). Evaluation brief: Implementation of school-wide PBS across the United States. OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. Retrieved from http://pbis.org/evaluation/evaluation_briefs/default.aspx
- Sugai, G., & Horner, R. H. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Child and Family Behavior Therapy, 24*, 23–50.
- Sugai, G., Horner, R. H., Algozzine, R., Barrett, S., Lewis, T., Anderson, C., . . . Simonsen, B. (2010). *School-wide positive behavior support: Implementers' blueprint and self-assessment*. Eugene: University of Oregon.
- Sugai, G., Horner, R. H., & McIntosh, K. (2008). Best practices in developing a broad-scale system of support for school-wide positive behavior support. In A. Thomas & J. P. Grimes (Eds.), *Best practices in school psychology V* (Vol. 3, pp. 765–780). Bethesda, MD: National Association of School Psychologists.
- Sugai, G., Lewis-Palmer, T., Todd, A., & Horner, R. H. (2001). *Schoolwide evaluation tool*. Eugene: University of Oregon.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn & Bacon.
- Vincent, C., Spaulding, S., & Tobin, T. J. (2010). A reexamination of the psychometric properties of the school-wide evaluation tool (SET). *Journal of Positive Behavior Interventions, 12*, 161–179. doi:10.1177/1098300709332345
- Walker, H. M., Horner, R. H., Sugai, G., Bullis, M., Sprague, J. R., Bricker, D., & Kaufman, M. J. (1996). Integrated approaches to preventing antisocial behavior patterns among school-age children and youth. *Journal of Emotional and Behavioral Disorders, 4*, 194–209.

About the Authors

Brandi Simonsen, PhD, is an associate professor of Special Education at the Neag School of Education at the University of Connecticut. Her current interests include schoolwide positive behavior support in alternative settings, classwide positive behavior support, and secondary/tertiary supports for students with more intensive needs.

Lucille Eber, EdD, is the Statewide Director of the Illinois PBIS Network. Her current interests include scaling up schoolwide positive behavior support, developing state-wide data and evaluation systems to support implementation, and integrating wraparound process throughout the SWPBS framework

Anne C. Black, PhD, is an associate research scientist in the Department of Psychiatry at Yale University. Her recent work has focused on risk taking and impulsivity, and on methods for modeling incomplete data.

George Sugai, PhD, is the Carole J. Neag Endowed Professor in the Neag School of Education at the University of Connecticut.

His interests include schoolwide positive behavior supports, function-based behavior supports, systems implementation improvement, and educating students with behavior disorders.

Holly Lewandowski, MA, is a program evaluation consultant and former evaluation coordinator for the Illinois PBIS Network. Her current interests include using utilization-focused evaluation and developmental evaluation to evaluate programs for nonprofits, universities, and state-run agencies.

Barbara Sims, MA, is an associate director with the National Implementation Research Network in the Frank Porter Graham Child Development Institute at the University of North Carolina at Chapel Hill. She is currently concentrating on the application of implementation research, especially in building the capacity of state education systems to implement and scale up effective education innovations

Diane Myers, PhD, is an assistant professor of Special Education at Assumption College. Her interests include positive behavior support, teacher education and professional development, and classroom management.