A Randomized, Wait-List Controlled Effectiveness Trial Assessing School-Wide Positive Behavior Support in Elementary Schools

Robert H. Horner  
*University of Oregon, Eugene*

George Sugai  
*University of Connecticut, Storrs*

Keith Smolkowski  
*University of Oregon, Eugene*

Lucille Eber  
*Illinois State Board of Education, LaGrange Park*

Jean Nakasato  
*Hawaii State Board of Education, Honolulu*

Anne W. Todd  

Jody Esperanza  
*University of Oregon, Eugene*

We report a randomized, wait-list controlled trial assessing the effects of school-wide positive behavior support (SWPBS). An effectiveness analysis was conducted with elementary schools in Hawaii and Illinois where training and technical assistance in SWPBS was provided by regular state personnel over a 3-year period. Results document that the training and technical assistance were functionally related to improved implementation of universal-level SWPBS practices. Improved use of SWPBS was functionally related to improvements in the perceived safety of the school setting and the proportion of third graders meeting or exceeding state reading assessment standards. Results also document that levels of office discipline referrals were comparatively low, but the absence of experimental control for this variable precludes inference about the impact of SWPBS. Implications for future research directions are offered.

**Keywords:** school-wide positive behavior support; randomized control trial, effectiveness trial

School-wide positive behavior support (SWPBS) is a systems approach to establishing both the overall social culture and intensive behavior supports needed to achieve academic and social success for all students. The SWPBS approach emerged from (a) the three-tier community health model that promotes primary, secondary, and tertiary prevention (Larson, 1994; National Research Council & Institute of Medicine, 1999; Shonkoff & Phillips, 2000); (b) recommendations to apply whole-school (and system-wide) interventions (Adelman & Taylor, 2006; Biglan, 1995; Mayer, 1995; Mayer & Butterworth, 1979; Walker et al., 1996); and (c) early demonstrations that behavioral interventions implemented at the whole-school level were linked to improved social

Authors’ Note: The authors acknowledge the assistance of Dr. Claudia Vincent in the final preparation of this manuscript. This research was supported in part by U.S. Department of Education Grant No. H326S980003. Opinions expressed herein do not necessarily reflect the policy of the Department of Education, and no official endorsement by the Department should be inferred.

Editor’s Note: The action editor for this article was Robert Koegel.

SWPBS is an application of evidence-based behavioral practices within organizational systems that are designed to enhance the fidelity of implementation and sustainability of effects (Doolittle, 2006; Sugai & Horner, 2006). The primary prevention tier of SWPBS involves defining, teaching, monitoring, and rewarding a small set of behavioral expectations for all students across non-classroom and classroom settings. In addition, a clearly defined and consistently implemented continuum of consequences and supports for problem behaviors are established, and the faculty adopt a process of continuously measuring the social behavior of students in the school and using those data for active decision making (Tobin, Sugai, & Colvin, 2000). The goal of this focus on primary prevention is to establish a social culture in which students expect and support appropriate behavior from each other and opportunities for teaching and learning can be maximized. Students should experience the school context as socially predictable, consistent, safe, and positive.

Although the process for investing in primary tier prevention varies across elementary, middle, junior high, and high schools, the basic features and outcomes are consistent (Barrett, Bradshaw, & Lewis-Palmer, 2008; Flannery, Sugai, & Anderson, in press; Mass-Galloway, Panyon, Smith, & Wessendorf, 2008; Kincaid, Childs, Wallace, & Blase, 2007; Muscott, Mann, & LeBrun, 2008). School-wide behavioral expectations are defined, taught, and rewarded within a management system that also includes a continuum of consequences for behavioral errors, and continuous collection and use of data for decision making. As with the Response to Intervention approaches to preventing literacy failure (Brown-Chidsey & Steece, 2005; D. Fuchs, Mock, Morgan, & Young, 2003; L. S. Fuchs & Fuchs, 1998; Ikeda et al., 2002; MacMillan & Speece, 1999), SWPBS presumes that active investment in prevention of problem behavior will occur with ongoing monitoring of student behavior and early intervention when problems are identified. Within the SWPBS approach problems may occur with the behavior of individual students and require individualized interventions, or problems may be more context specific, involving groups of students, and require organizational or structural changes (e.g. schedule changes, altering supervision patterns, modifying group consequences for lunch periods). To meet these more intense behavior support needs, schools adopting SWPBS are expected to include secondary and tertiary tier behavior support strategies with their primary tier prevention efforts. Secondary tier behavior supports are designed for students “at risk” for problem behaviors, who benefit from low-intensity interventions that can be administered with high efficiency (Crone, Horner, & Hawken, 2004; Walker, Golly, McLane, & Kimmich, 2005). Tertiary tier behavior supports involve highly individualized interventions that are based on functional behavioral assessment, often include family or community collaboration, and involve high investment to prevent the emergence or continuation of higher intensity problem behaviors (Fairbanks, Sugai, Guardino, & Lathrop, 2007; Ingram, Lewis-Palmer, & Sugai, 2005; Kincaid et al., 2007; Scott & Eber, 2003). Adoption of SWPBS typically involves a 2- to 3-year process of professional development and systems change, with training support provided by district, state, or national trainers (Blonigen et al., 2008; Center on Positive Behavioral Interventions and Supports, 2005).

More than 7,000 schools across the United States are currently in varying stages of adopting SWPBS (Bradley, Doolittle, Lopez, Smith, & Sugai, 2007). Evaluation reports document that the SWPBS practices can be adopted with fidelity (Barrett et al., 2008; Eber 2006; Muscott et al., 2008) and that adoption of SWPBS is associated with reduction in both observed rates of problem behavior (Horner, Sugai, Todd, & Lewis-Palmer, 2005) and reported office discipline referrals (ODRs; Barrett et al., 2008; Eber, 2006; Horner et al., 2005; Lohrman-O’Rourke et al., 2000; Luiselli, Putnam, & Sunderland, 2002; Olmstead v. L.C., 1999; Taylor-Greene et al., 1997; Taylor-Greene & Kartub, 2000). Evaluation reports also provide examples of SWPBS sustaining for nearly a decade (Doolittle, 2006; Horner et al., 2005), with benefits extending to academic gains as well as improved social behavior (Eber, 2006; Gottfredson, Gottfredson, & Hybl, 1993; Kellam, Rebok, Ialongo, & Mayer, 1994; McIntosh, Chard, Boland, & Horner, 2006; J. R. Nelson et al., 2002; Putnam, Horner, & Algozzine, 2006).

To date, however, few experimental assessments of SWPBS have been conducted (C. M. Nelson & Rutherford, 1988), and the need to document that valued outcomes can be achieved by typical interventionists (e.g., effectiveness analyses) continues to grow. Effectiveness analyses are of particular importance for examining the approach under typical educational conditions (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005; Greenwood, Delquadri, & Bulgren, 1993; Kutash, Duchnowski, & Lynn, 2006). Research examining the efficacy of an intervention traditionally focuses on highly controlled applications of the intervention by those individuals who developed the
intervention. This research is of particular value in determining if a particular intervention might work but always leaves open the question of whether the intervention could be applied with fidelity and effect by typical personnel under typical conditions. An effectiveness analysis occurs under regular conditions with regular personnel and is of particular value for educators looking for practical educational reforms. The purpose of the present analysis was to conduct a randomized, wait-list controlled, effectiveness analysis of SWPBS. Four research questions were addressed: Does a functional relationship exist between delivery of the SWPBS implementation procedures by regular, state personnel and (a) fidelity of SWPBS primary prevention practices used within elementary schools, (b) improved levels of perceived safety (risk and protective factors) in the schools, (c) reduced levels of reported ODRs, and (d) the proportion of third graders who meet or exceed the state reading achievement standard?

**Method**

**Participants**

The research was conducted between 2002 and 2006 with elementary schools (K–5) in Illinois and Hawaii. Both states had invested more than 5 years in development of state-level capacity for implementing SWPBS prior to the study, each had at least 100 schools implementing SWPBS, and each state had state-level personnel experienced in the training and support practices associated with the approach (Eber, 2006; Nakasato, 2000). The schools were selected based on (a) state capacity to provide whole-school, team-training in SWPBS; (b) self-nomination by school building administrators; and (c) absence of prior direct training for school building personnel in SWPBS. The first 30 schools in each state meeting these requirements within the time period of implementation were included in the analysis. Within each state, the 30 selected schools were randomly assigned to either a “Treatment” or a “Control/Delayed” group. The combined Treatment group (N = 30; 15 from Illinois and 15 from Hawaii) was designated to receive SWPBS training at Time 1 (T1) of the study, and the combined Control/Delay group (N = 30) was designated to receive SWPBS training 1 year later at Time 2 (T2) of the study.

During the 1st year of the analysis design adjustments were required because of three Illinois Treatment schools that left the study due to school consolidations/closures or administrator changes, and seven Illinois Control/Delay schools that found an alternative way to obtain access to training in SWPBS. Because these seven Illinois Control/Delay schools no longer met the criterion for “control,” they were deleted from the analysis. To compensate for the loss of Illinois schools an additional 13 schools from Chicago Public Schools were selected and randomly assigned (6 to Treatment and 7 to Control/Delay) resulting in a combined group of 33 schools in the Treatment group and 30 in the Control/Delay group. Additional attrition over the 4 years of the project left 30 schools in the Treatment group and 23 in the Control/Delay group available at the final assessment. To reduce bias due to attrition and missing data, however, all schools were included in the analysis (Nich & Carroll, 1997), even if they provided data at only a single time point. Table 1 provides a summary of basic demographic features for the 33 Treatment and 30 Control/Delay schools. Schools participating in the study reported enrollment levels averaging 471 (range = 131–969), proportion of students from non-White ethnicities averaging 61% (range = 2–100%), proportion of students qualifying for free or reduced lunch averaging 51% (range = 0–99%), and proportion of students on IEPs averaging 9% (range = 6%–74%). The Control/Delay schools had larger enrollment than the Treatment schools, but all other demographic variables were not statistically significantly different.

**Design**

The study was designed as a randomized, wait-list control effectiveness trial with groups of schools from Illinois and Hawaii measured repeatedly. Within each cluster of schools (Illinois, Hawaii) measurement occurred prior to any training (T1). A second measurement was taken after the Treatment half of each group had received approximately 1 year of SWPBS training and technical support (T2), and ongoing measurement occurred after initiation of the training with the Control/Delay group (T3+). The actual dates of implementation varied with the Illinois schools starting in 2003 and the Hawaii and Chicago schools beginning the process in 2004. Although actual years varied, T1 was always the initial measure when no training had been provided, T2 was when training had been initiated with Treatment but not Control/Delay schools, and T3 was when all schools had initiated training in SWPBS. A diagram of the basic design for the 33 Treatment and 30 Delay/Control schools is provided in Table 2.

**Measures**

Four measures were included in the analysis. The measures examined the extent to which schools were successful at adopting SWPBS when they received training from state trainers, the impact of SWPBS on perceived school...
safety, the reported levels of ODRs, and the impact of SWPBS on the proportion of third graders meeting the state reading achievement standard.

**Implementation of SWPBS: School-wide Evaluation Tool (SET).** The SET (Sugai, Lewis-Palmer, Todd, & Horner, 2001) is a direct observation index of the extent to which a school is implementing SWPBS primary tier prevention practices. The instrument contains 28 items that assess seven core features of SWPBS: (a) behavioral expectations are defined, (b) behavioral expectations are taught to students, (c) rewards are delivered for appropriate social behavior, (d) predictable consequences are delivered for inappropriate behavior, (e) formal systems are used to collect data and use data for decision making, (f) an administrator is supportive and actively involved in improving student social behavior, and (g) district support exists for improving student social behavior. An evaluator not employed by the school and trained to an 85% interobserver agreement standard by the authors (Todd, Lewis-Palmer, Horner, Sugai, & Phillips, 2002) spends approximately 2 hr in a school examining permanent products and interviewing an administrator, students, and faculty/staff. The instrument produces an overall “Total” score between 0 and 100%, and subscale scores between 0 and 100% for each of the seven core features. When the study began, a school was considered as implementing SWPBS at criterion when the SET Total score was at least 80%. The internal consistency of the SET has been documented with an alpha of .96, with test–retest reliability reported at .97, construct validity compared with Effective Behavior Support Self Assessment Survey $r = .75$, and interscorer agreement of .99 (Horner et al., 2004).

**Perceived school safety: School Safety Survey (SSS).** The SSS (Sprague, Colvin, & Irvin, 1996) provides a summary rating of at least five individuals serving four different roles within a school (an administrator, a supervisory staff member, a classified staff member, and at least one teacher). The SSS produces two scores: A Risk Factor score (with higher scores indicating higher level of behavioral risk) is based on 17 questions examining (a) design of space; (b) crowding; (c) perceived caring; (d) perceived sensitivity to cultural differences; (e) student bonding with school; (f) the quality of student–adult interactions, perceived fairness of school rules; and (g) level of adult supervision; the Protective Factor score (with higher scores indicating higher protection from

---

**Table 1**

**Average Enrollment and Average Percentage of Students in Racial–Ethnic, Socioeconomic, and Instructional Categories at Time 1 by Condition**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Condition</th>
<th>M</th>
<th>SD</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrollment*</td>
<td>Treatment</td>
<td>440.3</td>
<td>169.0</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Control/Delay</td>
<td>547.8</td>
<td>242.4</td>
<td>28</td>
</tr>
<tr>
<td>% American Indian</td>
<td>Treatment</td>
<td>0.2</td>
<td>0.3</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Control/Delay</td>
<td>0.4</td>
<td>0.6</td>
<td>28</td>
</tr>
<tr>
<td>% Asian or Pacific Islander</td>
<td>Treatment</td>
<td>38.1</td>
<td>40.2</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Control/Delay</td>
<td>34.2</td>
<td>38.8</td>
<td>28</td>
</tr>
<tr>
<td>% Black</td>
<td>Treatment</td>
<td>24.1</td>
<td>32.9</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Control/Delay</td>
<td>34.1</td>
<td>42.7</td>
<td>28</td>
</tr>
<tr>
<td>% Hispanic</td>
<td>Treatment</td>
<td>9.0</td>
<td>22.4</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Control/Delay</td>
<td>13.1</td>
<td>27.0</td>
<td>28</td>
</tr>
<tr>
<td>% White</td>
<td>Treatment</td>
<td>27.6</td>
<td>25.6</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Control/Delay</td>
<td>18.2</td>
<td>24.9</td>
<td>28</td>
</tr>
<tr>
<td>% free or reduced lunch</td>
<td>Treatment</td>
<td>50.0</td>
<td>31.9</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Control/Delay</td>
<td>58.6</td>
<td>31.7</td>
<td>28</td>
</tr>
<tr>
<td>% students with IEPs</td>
<td>Treatment</td>
<td>12.5</td>
<td>13.0</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Control/Delay</td>
<td>12.8</td>
<td>6.8</td>
<td>27</td>
</tr>
</tbody>
</table>

Note: Statistics include the mean, standard deviation, and sample size. IEP = Individualized Education Program.

a. Enrollment differed significantly between conditions ($t = −2.03$, $p = .047$).

---

**Table 2**

**Diagram of Study Design**

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment*</td>
<td>O</td>
<td>X</td>
<td>O</td>
</tr>
<tr>
<td>Control*</td>
<td>O</td>
<td>O</td>
<td>X</td>
</tr>
</tbody>
</table>

Note: O = Observation; X = Treatment.

a. $N = 33$.
b. $N = 30$. 

---
behavioral risk factors) is based on 16 questions examining (a) school climate, (b) clarity of behavioral expectations, (c) perception that all students are included in the school, (d) student perception of identification with the school, (e) student participation, (f) opportunities for student skill acquisition, and (g) formal and predictable systems for conflict resolution. The SSS has a reported internal consistency with alpha equaling .90 and has been used as an index of basic school safety (Laxton & Sprague, 2005; Sprague et al., 1996; Sprague & Walker, 2005).

**Level of problem behavior: Office discipline referrals as measured by the School-wide Information System (SWIS).** To index overall level of social behavior within each school, ODR data were collected by school staff and reported using the SWIS (May et al., 2000). The SWIS is a Web-based discipline data system currently used by more than 3,000 schools nationally. A district-level “SWIS Facilitator” provides school personnel with (a) technical support in the design of discipline codes, office referral forms, and discipline procedures; (b) direct training on use of the SWIS computer application to enter data and instantly recover graphic and tabular data summaries; and (c) direct training in the use of discipline data for decision making. Irvin, Tobin, Sprague, Sugai, and Vincent (2004) reviewed the conditions in which ODR data may or may not provide valid information, and Tobin and Sugai (1999) outlined the value of using ODR data for action planning at the whole school level. Irvin et al. (2006) recently documented that SWIS meets basic validity criteria. One metric provided by SWIS is the rate of ODR per 100 enrolled students per school day. Of 1,010 elementary schools (grade ranges K–6) using SWIS during the full 2005–06 academic year the mean ODR per 100 enrolled students per school day was .37 (SD = .50; Bradley et al., 2007).

**Academic achievement: State standardized tests.** For the purposes of this analysis, the proportion of third graders meeting the state reading assessment was used as a dependent variable assessing SWPBS impact on academic achievement. In Illinois, the Illinois State Achievement Test was administered each spring. In Hawaii, the Stanford Achievement Test (version 9) was used to assess third-grade reading competence. Data from reading performance collected and reported by state educators were used to document the percentage of third graders in each school each year meeting or exceeding the state reading standard for third grade. These data were retrieved from respective state departments of education.

**Procedures**

Consistent with an effectiveness analysis, all direct intervention with the participating schools was conducted by typical state personnel with typical state resources. The researchers were involved in data collection when data collection exceeded typical state requirements (e.g., SET). In addition, research faculty had conducted direct training of the state SWPBS trainers in Hawaii and Illinois as part of earlier SWPBS implementation in each state.

At T1 each school was monitored for current level of SWPBS implementation, school safety, and academic achievement. Prior to intervention, none of the participating schools used ODR data collection systems that met the validity standards outlined by Irvin et al. (2004). Those schools assigned to the Treatment group then received training and technical support in SWPBS. Training involved teams from each school attending three to four training events (1–2 days each) per year over a 2-year period. Training was delivered by regular state personnel formally trained in SWPBS practices (cf. http://www.pbismanual.uoe.cs.org for a summary of training content). No funding or personnel were provided by the researchers to deliver training for school teams. Training and technical assistance by state trainers during the 1st year of training focused on (a) establishing a faculty-wide commitment to build a positive, school-wide social culture; (b) developing the team structure and management systems for implementing educational reform; (c) implementing a process for defining, teaching, monitoring, and acknowledging three to five school-wide behavioral expectations; (d) establishing a continuum of consistent consequences for inappropriate behavior; and (e) implementing the SWIS as a process for collecting, summarizing, and using data for active decision making. The state trainers met with school teams over the 2-year period of implementation and used SWPBS training materials to establish local “coaches” in each district/region (cf. http://www.pbismanual.uoe.cs.org for coaches training materials). SWPBS coaches met approximately monthly with each team and provided logistical and coordination support. A central part of both team training events and monthly coaching support was the review and use of fidelity (implementation self-assessment checklists) and outcome (office discipline referral) data for planning and decision making.

Because of policy changes in Hawaii, the training provided to Hawaii Control/Delay schools between T2
and T3 was reduced from systematic delivery by state trainers, to 2 to 3 days of technical support by district personnel. This level of training and technical assistance did not allow delivery of the full training content that was available to other schools in the analysis. Note that the focus of the present analysis was on the fidelity and impact of training content provided during the 1st year of training (e.g., the primary prevention component of SWPBS). We report the addition of training in secondary/tertiary practices here to be procedurally complete.

### Results

#### Number of Schools per Time Period

School attrition over the 4-year process occurred due to (a) schools from the Illinois Control/Delay group gaining access to SWPBS training, (b) closure of schools, and (c) change in building administrator. These factors resulted in the addition of the Chicago school cluster.

### Implementation of SWPBS

The initial focus of the analysis was on the extent to which regular educational personnel could use SWPBS training materials to implement SWPBS practices in schools. The primary dependent variable was the SET Total score. The mean SET Total scores for each group at each time period are provided in Table 3.

Two analyses were performed with the SET Total score. The first was an unadjusted Time × Condition (group) analysis with data from only T1 and T2. This analysis is described in Murray (1998, pp. 180–184, 296–303). This model has the advantage of including all data for estimated differences, whether or not a school provided data at both time points (T1 & T2), and accounts for the autocorrelation among assessments within the same school. Murray discussed the strengths and weaknesses of this model in more detail (p. 183).

We also conducted a random coefficients analysis (Murray, 1998, pp. 216–220, 340–345; Singer & Willett, 2003) that incorporated data from all available assessment times. Because of the wait-list design, this model is

### Table 3

**Outcome Measures by Condition and Time Period**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Condition</th>
<th>Statistic</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>School-wide Evaluation Tool total score</td>
<td>Treatment</td>
<td>$M$</td>
<td>.381</td>
<td>.785</td>
<td>.823</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$SD$</td>
<td>.159</td>
<td>.120</td>
<td>.103</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$n$</td>
<td>33</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Control/Delay</td>
<td>$M$</td>
<td>.388</td>
<td>.459</td>
<td>.640</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$SD$</td>
<td>.146</td>
<td>.189</td>
<td>.214</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$n$</td>
<td>26</td>
<td>27</td>
<td>23</td>
</tr>
<tr>
<td>School Safety Survey risk factor score</td>
<td>Treatment</td>
<td>$M$</td>
<td>.370</td>
<td>.344</td>
<td>.343</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$SD$</td>
<td>.123</td>
<td>.124</td>
<td>.143</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$n$</td>
<td>24</td>
<td>29</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Control/Delay</td>
<td>$M$</td>
<td>.387</td>
<td>.415</td>
<td>.358</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$SD$</td>
<td>.142</td>
<td>.149</td>
<td>.154</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$n$</td>
<td>19</td>
<td>24</td>
<td>20</td>
</tr>
<tr>
<td>Office Discipline Referrals per 100 students per school day</td>
<td>Treatment</td>
<td>$M$</td>
<td>.390</td>
<td>.360</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$SD$</td>
<td>.230</td>
<td>.350</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$n$</td>
<td>25</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control/Delay</td>
<td>$M$</td>
<td></td>
<td>.340</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$SD$</td>
<td>.290</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$n$</td>
<td></td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Proportion of third graders meeting or exceeding the state reading standard</td>
<td>Treatment</td>
<td>$M$</td>
<td>.455</td>
<td>.529</td>
<td>.536</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$SD$</td>
<td>.216</td>
<td>.197</td>
<td>.178</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$n$</td>
<td>33</td>
<td>31</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Control/Delay</td>
<td>$M$</td>
<td>.380</td>
<td>.402</td>
<td>.436</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$SD$</td>
<td>.188</td>
<td>.186</td>
<td>.203</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$n$</td>
<td>28</td>
<td>27</td>
<td>23</td>
</tr>
</tbody>
</table>

Note: Bold indicates posttraining. Lower scores in school safety index improvements. Office Discipline Referrals reported only for schools with a full year of data. Statistics include the mean, standard deviation, and sample size.
more complicated and difficult to provide a single test of intervention. The analysis included effects for time, group (Treatment or Control/Delay), training status (pretest or posttest), Training × Group, time posttraining, and Group × Time Posttraining. The training term was added to account for the wait-list design of the study; the time posttraining (posttraining slope), and its interaction with group tested the potential increase or decline of SET Total scores after training. Finally, a post hoc pre–post comparison of change from before to after training was estimated independently for both groups.

The first analysis was intended to provide a clear description of the immediate effects of intervention and the associated tests. It compares means in Table 3 from T1 and T2 only. The second analysis provides long-term results. Key comparisons include only pre- to posttraining change for each group and the increase or decline in slope after training. Both sets of analyses were performed with SAS PROC MIXED (Littell, Milliken, Stroup, & Wolfinger, 1996; SAS Institute, 2002). Both statistical models also employed maximum likelihood estimation and used all available data from all assessments. Such an analysis can provide unbiased conclusions even in the face of substantial attrition provided the data were missing at random (Laird, 1988; Nich & Carroll, 1997; Schafer & Graham, 2002), that is, that the missing data do not depend on unobserved determinants of the outcome of interest (Little & Rubin, 2002). We do not believe that missing data in this study represents a serious departure from the missing at random assumption.

With the first analysis, the Time × Condition (group) effect was statistically significant (.328), t(53) = 6.48, p < .0001. Because the Time × Condition analysis lends itself to easy interpretation, we present the raw model estimates for this analysis (e.g., .328). On average, then, the analysis indicates that Treatment schools improved .328 more than the Control/Delay schools from T1 to T2. A partial correlation coefficient (r) was computed to estimate the effect size, and for Time × Condition the partial r was .67 and Cohen’s d was 1.78. Subsequent conditional effects (Jaccard & Turrisi, 2003) showed that (a) the Control/Delay group did not differ significantly from the Treatment group at T1 (–.001), t(59) = –.03, p = .9765; (b) for the Treatment group, T2 differed significantly from T1 (.405), t(53) = 11.84, p < .0001; and that (c) at T2, the Control/Delay group differed significantly from the Treatment group (.327), t(53) = 7.88, p < .0001. The Control/Delay group improved only slightly from T1 to T2 (.077), t(53) = 2.06, p = .0441. These results showed a significant effect for SET Total scores both between groups (Treatment vs. Control) and within groups (pre–post).

The second set of analyses demonstrated a statistically significant increase in SET Total scores immediately after training for the Treatment group, t(76) = 7.11, p < .0001, but not the Control/Delay group, t(76) = 1.79, p = .0779. Both these effects were adjusted for the pretraining slope. The Control/Delay group continued to increase in their SET Total score from T3 through T5, t(76) = 3.14, p = .0024, but the slope after training, from T2 to T5, was not statistically significant for the Treatment group, t(76) = 1.45, p = .1524. The longitudinal analysis confirmed that the Control/Delay group did not differ significantly from the Treatment group at T1, t(59) = –.039, p = .9687. The Control/Delay schools for Illinois and Chicago averaged 81% and 79% Total SET scores at T3, respectively, but the lower fidelity implementation in Hawaii (1–2 days of district training rather than 5–6 days of state trainer training) resulted in a mean Total SET of 59% for the Hawaii Control/Delay schools at T3.

**Perceived School Safety**

If SWPBS is effective at altering the social culture of the school, perception of the safety and quality of the social environment should improve. The SSS was used as an index of the extent to which the overall social culture was perceived as a safer and more socially supportive environment. The mean SSS Risk Factor scores per school per time period are provided in Table 3. The Protective Factor scores for Hawaii and Illinois schools were consistently high and precluded the option for assessing change.

The analysis for SSS Risk Factor employed the same strategies as described for the SET Total score but with a different dependent variable. Neither analysis resulted in statistically significant pretest differences. The Time × Condition (group) analysis with data from T1 and T2 provided a statistically significant Time × Condition interaction (–.064), t(35) = –2.55, p = .0154. A partial correlation coefficient (r) was computed to estimate the effect size, and for Time × Condition, the partial r was –.40 and Cohen’s d was –.86. The conditional effects showed a statistically significant difference between Treatment and Control/Delay groups at T2 (–.078), t(35) = –2.03, p = .0499, although the decline after training between T1 and T2 for Treatment schools did not reach statistical significance (.026), t(35) = –1.54, p = .1318. Control/Delay schools, however, in the absence of training, increased in risk between T1 and T2 (.039), t(35) = 2.03, p = .0496. The random coefficients analysis showed a statistically
significant decrease in risk immediately after training for both the Treatment, \( t(37) = -2.29, p = .0278 \), and Control/Delay, \( t(37) = -2.69, p = .0107 \), groups.

Office Discipline Referrals

Because the typical ODR data collected by schools prior to intervention did not meet Irvin et al. (2004) standards we do not have the Pre-SWPBS data needed to experimentally examine the effects of SWPBS on ODR rates. Descriptively, Table 3 provides mean ODR rates per 100 students per school day for schools with a full year of data in each group by time. It is worthy of note that the mean ODR rate per 100 students per school day for 1,010 elementary schools as reported by the SWIS national database (http://www.swis.org; August 15, 2006) was .37 (SD = .50). Following training and technical assistance in SWPBS the elementary schools in this study were reporting comparatively low rates of ODRs. This finding cannot, however, be associated with implementation of SWPBS, given the absence of Pre-SWPBS data.

Academic Outcomes

SWPBS is intended to improve the overall effectiveness of schools as learning environments by increasing (a) the amount of time students are in school, (b) the proportion of minutes that classrooms are engaged in instruction, and (c) the level of student engagement during instruction. Focusing on social behavior is not expected to improve academic outcomes, but improving the social behavior of students combined with effective curriculum and instruction is expected to result in better academic outcomes (Kellam, Mayer, Rebok, & Hawkins, 1998; Putnam et al., 2006). To measure the overall impact of SWPBS on academic outcomes, the percentage of third graders meeting or exceeding the state reading standard was assessed annually. The mean percentage of third graders meeting the state reading standard is provided in Table 3. These results are encouraging but remain preliminary.

The Time \( \times \) Condition effect was not statistically significant for the proportion of third graders who met or exceeded the state reading standard (.036), \( t(57) = 1.21, p = .2307 \). The analysis, however, showed a statistically significant differences between T1 and T2 for the Treatment group (.056), \( t(57) = 2.75, p = .0080 \), and between Treatment and Control/Delay at T2 (.111), \( t(57) = 2.20, p = .0320 \), with a partial correlation of \( r = .28 \) and Cohen’s \( d = .58 \) for the latter comparison. Because the Time \( \times \) Condition effect was not statistically significant, the conditional effects should be interpreted with caution and perhaps held tentative until further study can be conducted. Most likely, the interaction was not statistically significant because of the combination of a small but statistically nonsignificant increase between T1 and T2 for Control/Delay schools (.020), \( t(57) = 0.93, p = .3546 \), and a higher T1 score for the Treatment group compared to the Control/Delay group. The Treatment and Control/Delay groups, however, did not differ statistically at T1 (.075), \( t(59) = 1.44, p = .1551 \).

The random coefficients analysis did not estimate statistically significant differences associated with training or at T1, nor did the analysis demonstrate statistically significant improvement (slopes) after training.

Discussion

A randomized, wait-list controlled, effectiveness analysis was conducted with elementary schools in Hawaii and Illinois where training and technical assistance in school-wide positive behavior support was provided by regular state personnel over a 3-year period. Analysis of the results from this randomized control trial provides useful but preliminary messages. The T1 SET data document that the schools were not currently using SWPBS practices, and the design provides experimentally rigorous documentation that state personnel operating within typical resources provided the training and technical assistance needed to implement SWPBS practices and systems with fidelity. In addition the results provide statistically significant documentation that schools implementing SWPBS were perceived as safer environments. Finally, the results provide preliminary indications that schools implementing SWPBS were associated with increased third-grade reading performance; however, this finding needs elaboration and replication. The post-only information about ODR rates indicates that the schools were reporting lower than average levels of problem behavior, but the absence of pretraining information about ODRs prevents documentation of a causal association between SWPBS and lowered discipline referral levels.

Limitations of this study are important for careful interpretation of the results. Interpretation of the extent to which schools were able to adopt SWPBS practices shows strong effects except for the T3 data when Control/Delay schools in Hawaii received reduced fidelity of SWPBS training. These schools received a lower dose of the intervention (1–2 days of training instead of 5–6) from less experienced trainers (district trainers rather than state-trainers). Although the relative contribution of these factors on the impact of SWPBS implementation remains unclear, they suggest considerations for future research and remind us of the importance of treatment fidelity and the unpredictability of conducting research in
applied settings with real implementers. This information will be especially useful as researchers shift their attention from efficacy to effectiveness (Glasgow, Lichtenstein, & Marcus, 2003).

A second limitation lies in the absence of experimental data to assess ODR rates. Previous evaluation and experimental analyses have reported reductions in problem behavior and ODRs when SWPBS practices were implemented (Barrett et al., 2008; Eber, 2006; Kincaid et al., 2007; Muscott et al., 2008). The strategies used by schools to collect ODR information prior to SWPBS implementation unfortunately did not provide interpretable counts, and implementation of SWIS prior to intervention would have constituted partial implementation of SWPBS. The net finding is that although the results document comparatively low rates of ODRs in schools implementing SWPBS, no association between these rates and implementation of SWPBS can be inferred. Future research is needed to compensate for the high variability found in typical school procedures for collecting and managing information on discipline status. In addition, other measures may be needed to corroborate and/or substitute for extant school discipline information.

A final limitation lies in the analysis of the link between SWPBS and improved academic gains. In this study we used the proportion of third graders meeting state standards as a general index of the impact of SWPBS on academic outcomes. Although preliminary, our results generally support the hypothesis that implementation of school-wide systems of discipline interact with effective instruction to improve academic outcomes. We do not believe that teaching students the skills associated with being respectful and responsible will lead directly to mastery of core literacy competencies. Acquisition of reading skills is related to a complex interaction of effective curriculum materials, unambiguous instruction, and adequate intensity (time in instruction). However, to maximize academic outcomes children must be present, attentive, and engaged. The present results encourage investment in future research that measures directly the hypothesized interaction between effective teaching and effective behavior support. Recent single-case analyses examining the link between behavioral and literacy interventions have documented encouraging experimental results (McKenna, 2005; Preciado, 2006; Sanford, 2006). Larger scale assessments are needed to both examine these effects and document the functional mechanisms.

The results from this research provide a randomized, control trial demonstration that SWPBS can be implemented with fidelity with benefit to the social climate and academic outcomes of students. The results suggest that evidence-based practices, implemented systemically at the whole-school level is feasible and useful (Fixsen et al., 2005). The findings also define future needs related to (a) documentation of the basic mechanisms that are responsible for these effects, (b) extending the SWPBS analysis from the primary prevention tier to a more complete analysis of how schools apply secondary and tertiary tiers of the approach, (c) linking behavioral and academic support systems for efficient school-wide implementation, and (d) examining how effective practices can be implemented on scales of social importance.

Every school day, more than 100,000 schools in the United States open their doors to students and teachers with the important responsibility of preparing a competent and caring citizenry. The findings from this report and previous evaluation reports (Barrett et al., 2008; Colvin et al., 1993; Eber, 2006; Gottfredson et al., 1993; Kincaid et al., 2007; Muscott et al., 2008) suggest that a significant potential exists for improving the school-wide social culture in these schools. Investing in the school-wide social culture and behavior supports in schools has the potential to improve (a) the social competence of the students, (b) the amount of time and resources needed to address behavior problems, and potentially (c) the academic outcomes achieved by students. We believe that these findings are useful as schools consider the expectations and requirements of No Child Left Behind and reauthorization of the Individuals with Disabilities Education Act. The specific challenges associated with maintaining safe schools, closing the achievement gap, and better accommodating the range of learners in classrooms require learning environments that are predictable, consistent, positive, and safe. SWPBS may be one approach for helping schools become more effective learning environments.

References


Laxton, T., & Sprague, J., (2005). Refining the construct of school safety: An exploration of the correlates and construct validity of school safety measures. Unpublished dissertation. <AQ> For Laxton and Sprague, list if this is a doctoral dissertation and provide the university.</AQ>


**Robert H. Horner**, PhD, is a professor of special education at the University of Oregon and codirector of the National Technical Assistance Center on Positive Behavioral Interventions and Supports. His interests include systems of positive behavior support, functional assessment, generalization and maintenance, and instructional technology.

**George Sugai**, PhD, is Carole J. Neag Endowed Chair and professor of special education in the Neag School of Education at the University of Connecticut. He also is codirector of the OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports, and director of the Center for Behavioral Education and Research.

**Keith Smolkowski**, PhD, is an associate research scientist and research analyst at Oregon Research Institute. His interests include the design and analysis of complex research projects focused on topics such as early literacy and child social behavior in the areas of education and public health.

**Lucille Eber**, EdD, is the state director of the Illinois Positive Behavioral Interventions and Supports (PBIS) Network that coordinates technical assistance and evaluation related to school-wide PBIS in more than 900 Illinois schools. Dr. Eber also facilitates PBIS implementation and training plans for states and school districts across the country. Her areas of focus include integration of wraparound, interagency systems of care and school-based mental health into school-wide systems of positive behavior supports.

**Jean Nakasato** is an educational specialist with the Hawaii Department of Education. She has been the statewide coordinator for Positive Behavior Supports with the Hawaii Department of Education since 1998.

**Anne W. Todd**, MsW, is a senior research assistant at the University of Oregon, holds elementary education and severely handicapped Oregon Teaching certificates, and is a member of the Positive Behavior Support Research work group. Ms. Todd’s work has contributed to development of tools being used in schools across the country including the School Wide Information System (swis.org) and the School-wide Evaluation Tool.

**Jody Esperanza** is a research assistant at the University of Oregon, College of Education (Educational and Community Supports).