When Prevention is Not Enough: School-Wide Positive Behavior Interventions and Supports on

Students' Perception of Bullying

Abstract

Bullying continues to be a major concern in U.S. schools and is the focus of myriad prevention and intervention efforts. Researchers have recently cited school-wide positive behavior interventions and supports (SWPBIS) as a prevention framework for reducing school-based bullying. Therefore, we examined the effect of universal SWPBIS implemented with fidelity on students' self-report of bullying victimization. We used school-level propensity score matching to compare 76,248 students' self-report of bullying victimization in 118 schools that implemented SWPBIS with fidelity and 118 matched comparison schools. Random-effects regression models found no statistically significant difference between treatment groups on students' self-report of bullying victimization. Recommendations and limitations are discussed.

When Prevention is Not Enough: School-Wide Positive Behavior Interventions and Supports on

Students' Perception of Bullying

Bullying continues to be a major concern in U.S. schools and is the focus of myriad

prevention and intervention efforts (Espelage, Rose, & Polanin, 2015; 2016; Ttofi & Farrington,

2011). Recent national data suggests that more than 1 in 5 (20.8%) school-aged youth reported

being bullied by peers in school within a 12-month period (Musu-Gillette et al., 2018), and that

bullying can have significant negative short and long-term effects on victims (Swearer, Espelage,

Vaillancourt, & Hymel, 2010). In response, all U.S. states have passed anti-bullying legislation

(Yell, Katsiyannis, Rose, & Houchins, 2016). It is clear that schools have a legal and ethical

obligation to address bully victimization in their schools, but it remains unclear as to what

practices, programs, or approaches should be implemented. Researchers have recently cited

school-wide positive behavior interventions and supports (SWPBIS) as a prevention framework

for reducing school-based bullying (Bradshaw, 2013; 2015). Therefore, we examined

Bradshaw's recommendation by evaluating the relation between SWPBIS implemented with

fidelity and students' self-report of bullying victimization.

Bullying Defined

Bullying is a subcategory of interpersonal aggression, including physical (e.g., hitting,

kicking, pushing), verbal (e.g., calling names, threats), social (e.g., rumor spreading, group

exclusion), and electronic (i.e., cyberbullying) behaviors, that is defined as "any unwanted

aggressive behavior(s) by another youth or group of youths who are not siblings or current dating

partners that involves an observed or perceived power imbalance and is repeated or likely to be

repeated" (Gladden, Vivolo-Kantor, Hamburger, & Lumpkin, 2014, p. 7). While the central

tenets of bullying is *intentionality*, *repetition*, and *imbalance of power* (Hymel & Swearer,

2015; Olweus, 1993), the key distinction between bullying and other forms of peer aggression is the abuse of power (Vaillancourt, Hymel, & McDougall, 2003).

Bullying is also grounded in social interactions, where involvement is based on the relationship and associations between an individual and complex social systems (i.e., family, peers, school, community, society) in which the individual is situated (Hong & Espelage, 2012). Navigation within and between these social systems is multifaceted, which may place some subgroups of students at escalated risk for involvement (Musu-Gillette et al., 2018; Rose, Nickerson, & Stormont, 2015). For example, youth with disabilities are victimized at disproportional rates when compared to youth without disabilities (Blake, Lund, Zhou, Kwok, & Benz, 2012; Rose & Gage, 2016; Rose et al., 2015). While the intersection of several predictive factors (e.g., disability status, special education services, severity of disability) are associated with the disproportionate representation of youth with disabilities within the bullying dynamic (Rose et al., 2015), the most notable predictors include externalizing behaviors and social and communication skill deficits (McLaughlin, Byers, & Vaughn, 2010; Rose, Monda-Amaya, & Espelage, 2011). For example, Rose and Espelage (2012) found that students with behavior disorders engaged in significantly higher levels of bullying and fighting behaviors than their peers with and without disabilities. Similarly, Swearer, Wang, Maag, Siebecker, and Frerichs (2012) determined that students with behavior oriented disabilities, including behavior disorders, received more office disciplinary referrals, had lower prosocial behaviors, engaged in higher levels of bully perpetration, and experienced higher rates of victimization than their peers without disabilities.

In addition to disability status, prosocial skills, and externalizing behaviors, individual factors such as race, ethnicity, and gender have been evaluated as predictors of bullying

involvement. However, extant literature on gender, race, and ethnicity has been conflicting, warranting further investigation (National Academies of Sciences, Engineering, & Medicine [NASEM], 2016). For example, some studies have determined that African American youth are victimized less frequently than their White and Latino/a peers (Nansel et al., 2001; Spriggs, Iannotti, Nansel, & Haynie, 2007), where others reported that Latino/a youth are victimized less than their White and African American peers (Hanish & Guerra, 2000; Musu et al., 2018). Similarly, a conflicting gender discrepancy seems to exist, where females are more likely to report being bullied than males (NASEM, 2016), while males are more likely to engage in direct forms of bullying behaviors (Card, Stucky, Sawalani, & Little, 2008; Cook et al., 2010). To compound the issue, some have argued that females are more likely to experience and engage in indirect aggression (Sullivan & Stoner, 2012), while Card and colleagues' (2008) argued that the difference between males and females on indirect bullying was nonsignificant. Given the conflicting data, it has become increasingly more important to examine the role of race and ethnicity in bullying involvement.

School-Wide Positive Behavior Interventions and Support

SWPBIS is a systematic, multi-tiered framework for identifying students in need of behavior support and delivering universal, secondary, and tertiary social-behavioral interventions to increase the likelihood of pro-social behavior acquisition for all students (Sugai & Horner, 2009). SWPBIS is not a curriculum, strategy, intervention, or program but, instead, a process of building a school's capacity to (a) implement effective and preventive behavioral practices with integrity, (b) make data-based and team-based decisions, and (c) build a positive school climate and culture leading to school improvement and success (Gage, Whitford, & Katsiyannis, 2018; Horner et al., 2010). Universal prevention supports are designed to create a safe, predictable

environment for all students by establishing a common set of school-wide expectations, teaching those expectations, and reinforcing students for demonstrating those expectations (Lewis, Mitchell, Trussell, & Newcomer, 2014).

Universal implementation is conducted (a) school-wide, (b) in non-classroom settings, and (b) in classrooms. Behavioral expectations are taught and reinforced at the school level, often using a school-wide token economy system and paired with active supervision and precorrections in non-classroom settings, such as the cafeteria and playground. At the classroom level, universal classroom management is delivered, including high rates of behavior specific praise, opportunities to respond, behavior prompting, and reinforcing behavioral expectations. Targeted interventions are then implemented for students that do not respond to universal prevention efforts and often include evidence-based mentoring programs, such as Check-In Check-Out (Crone, Hawken, & Horner, 2010), or small group social skills lessons (Mitchell, Stormont, & Gage, 2011). Students that continue to exhibit elevated levels of problem behaviors following targeted intervention are referred for intensive, tertiary supports. Tertiary supports typically involve a functional behavior assessment (FBA) and a subsequent individualized behavior intervention plan (BIP).

Research suggests that SWPBIS has positive effects on a number of important school- and student-level outcomes. For example, a series of studies have documented a direct effect of SWPBIS on disciplinary exclusions, including office discipline referrals and in- and out-of-school suspensions (e.g., Childs, Kincaid, George, & Gage, 2016; Gage, Grasley-Boy, George, Childs, & Kincaid, 2018; Gage, Lee, Grasley-Boy, & George, 2018; Simonsen, Eber, Black, Sugai, Lewandowski, Sims, & Meyers, 2012), and student attendance (Freeman, Simonsen, McCoach, Sugai, Lombardi, & Horner, 2016). Research has also found positive effects on adult

perceptions of school climate (Bradshaw, Koth, Thornton, & Leaf, 2009), school organizational health (Bradshaw, Koth, Bevans, Ialongo, & Leaf, 2008) and school safety (Horner, Sugai, Smolkowski, Eber, Nakasato, Todd, & Esperanza, 2009). Recently, a longitudinal state-wide analysis found implementation of universal SWPBIS with fidelity also had a significant and meaningful effect on the percentage of students at or above state benchmarks in reading and mathematics (Gage, Leite, Childs, & Kincaid, 2017).

SWPBIS and Bullying Victimization

In addition to direct effects on students' disciplinary and academic outcomes, there is evidence to suggest that SWPBIS may also have an effect on bullying. A series of studies have evaluated the impact of bully prevention in positive behavior support (BP-PBS; Ross & Horner, 2009; 2014). BP-PBS is designed to (a) define and teach the concept of ''being respectful'' to all students in a school, (b) teach all students a three-step response (stop, walk, talk) that minimizes potential social reinforcement when they encounter disrespectful behavior, (c) precorrect the three-step response prior to entering activities likely to include problematic behavior, (d) teach an appropriate reply when the three-step response is used, and (e) train staff on a universal strategy for responding when students report incidents of problem behavior (Ross & Horner, 2009). The BP-PBS intervention has also been adapted and expanded into the *Bullying and Harassment Prevention in Positive Behavior Support: Expect Respect* intervention and evaluated using single-case design methods (Nese, Horner, Dickey, Stiller, & Tomanovich, 2014). However, BP-PBS and *Expect Respect* are additional interventions above and beyond implementation of SWPBIS. Therefore, although they are designed to be implemented alongside SWPBIS, they operate as separate interventions like the myriad of other school-based bullying

interventions (e.g., *Olweus Bullying Prevention Program*; Olweus, Limber, Flerx, Mullin, Riese, & Snyder, 2007).

A question remains as to whether or not SWPBIS implemented with fidelity may alone reduce the likelihood of bullying in schools without the addition of a stand-alone or specifically integrated bullying program. Only a few studies have examined the effect of universal SWPBIS on bullying in schools, two of which used randomized group experimental designs. Waasdorp, Bradshaw, and Leaf (2012) examined data from a randomized controlled trial evaluating the effectiveness of universal SWPIS implementation. Thirty-seven elementary schools were randomly assigned, with 21 in the treatment group and 16 in the control group. The authors used the Teacher Observation of Classroom Adaption-Checklist (TOCA-C) to evaluate students' risk of bullying perpetration. Specifically, teachers rated students about their likelihood to (a) tease classmates, (b) yell at others, (c) harm others, and (d) get into fights and then the items were averaged to create a composite score. Results from a longitudinal mixed effects model found a significant effect for SWPBIS on the slope value, indicating that teachers in schools implementing SWPBIS rated their students as exhibiting fewer bullying behaviors across time than in control schools. The authors note that the "findings suggest that a universal SWPBIS model is a promising approach for preventing bullying" (p. 155). However, it's worth noting that the measure used was teacher report of bullying perpetration, not students' self-report of bullying victimization, which could capture whether or not peer-to-peer bullying is impacted by SWPBIS.

Ward and Gersten (2013) examined the effect of the Safe and Civils Schools model, a proprietary version of SWPBIS, on staff and student report of bullying victimization in 32 elementary schools, with 17 schools in the treatment group and 15 in the control group. The authors then examined the impact after one year of implementation. Difference-in-difference

models found no significant effect on either staff or student reports of bullying after one year of

implementation. It is worth noting that implementation of Safe and Civil Schools was evaluated

using the Benchmarks of Quality (BoQ), a fidelity measure of SWPBIS and that the

implementation for the treatment schools was, on average, below the 70% threshold for

implementing SWPBIS with fidelity (Cohen, Kincaid, & Childs, 2007).

Purpose

There is limited research evaluating the effect of SWPBIS on bullying victimization.

Prior research has found that adult perceptions of bullying perpetration reduced across time in

schools implementing SWPBIS (Wassdorp et al., 2012). However, the only study that included

student responses (a) did not implement SWPBIS, but instead the proprietary Safe and Civil

Schools, (b) did not implement with fidelity in all schools, and (c) found no impact. Therefore,

research has not established whether universal SWPBIS implemented with fidelity has a distal

effect on bullying involvement. We believe that the effect is distal because SWPBIS alone is not

designed specifically to address bullying, but to decrease problem behavior and increase pro-

social behavior generally. The guiding research question for this study was:

Does Implementation of SWPBIS with fidelity have an effect on students' self-report of bullying

victimization?

Method

Sample

During the 2015-2016 school year, 376,958 elementary students completed a school

climate survey in Georgia. The de-identified dataset included student responses on the measure,

the school and district each student attended, and their grade. Students attended 1,285 schools

distributed across 186 school districts. The sample included 128,419 3rd graders (34%), 124,865

4th graders (33%), and 123,674 fifth graders (33%). We then collected demographic data for all public schools in Georgia for the 2015-2016 school year and discipline data for all public schools for the 2013-2014 school year (the earliest year indicated on their website that Georgia began statewide SWPBIS implementation). We restricted the data to only public elementary and intermediate schools, excluding alternative schools, vocational/technical schools, middle schools, and high schools because the school climate survey was completed by elementary students only and SWPBIS was implemented with fidelity in public elementary and intermediate schools.

Implementation of SWPBIS in Georgia is supported by the state department of education (DOE). The DOE facilitates district-level planning, provides school team training, technical assistance, and ongoing coaching to SWPBIS district. The DOE website reports the names of schools receiving SWPBIS training and support, and three levels of fidelity of implementation: Installing, defined primarily as fidelity below 70% on the Benchmarks of Quality (BoQ); Emerging, defined as fidelity between 70% and 85% on the BoQ; and Operational, defined as BoQ above 85%. We focused exclusively on schools implementing SWPBIS with fidelity (BoQ > 70%; Emerging and Operational) to ensure that the treatment schools implemented the intervention as designed and excluded schools ($n = 218$) that were Installing (i.e. not implementing Tier 1 SWPBIS with fidelity but had received training). A total of 119 schools implemented SWPBIS with fidelity during the 2015-2016 school year. We removed the 218 schools that received SWPBIS training but did not implement with fidelity to ensure that comparisons were only made between schools implementing SWPBIS with fidelity and schools never trained. Next, we reduced the sample of schools if they had any missing demographic or behavior data ($n = 44$), leaving a final sample of 905 possible comparison schools and 118 treatment schools (i.e., implemented SWPBIS with fidelity).

The average school enrollment for the remaining 1,023 schools was 611.0 students ($SD =$ 235.0) and slightly more White students (41.7%) than Black students (36.5%) students were enrolled, followed by Hispanic students (14.5%). Sixty-seven percent of the students were considered economically disadvantaged, 11.1% of students received special education services, and 11.2% were classified as limited English proficient (LEP). Fifty-six percent of schools were located in an urban setting. Across schools, approximately 36% of students performed at or above state benchmarks in reading and 39% in math. On average, schools reported a rate of 0.05 in-school suspensions per student, 0.06 out-of-school suspensions per student, and 0.19 ODR per student. Demographic characteristics by treatment group are presented in Table 1. Overall, there are few differences between the treatment schools and all possible comparison schools on most of the available school-level demographic characteristics, including urbanicity, with 55% of schools in urban settings. At the student-level, there was final sample of 39,187 students with school climate survey data available in the 118 treatment schools. Of those, 34.1% were in 3rd grade, 33.6% were in 4th grade, and 32.4% were in 5th grade.

Measures

SWPBIS fidelity of implementation.

Benchmarks of Quality (BoQ). The BoQ (Cohen, Kincaid, & Childs, 2007; Kincaid, Childs, & George, 2005, 2010; Childs, Kincaid & George, 2011) is a self-report measure used to assess the implementation fidelity of SWPBIS at the tier 1/universal level. The BoQ consists of 53 items rated on a three-point Likert scale (i.e., In Place, Needs Improvement, and Not in Place). Prior psychometric evidence suggests that the BoQ demonstrates strong internal consistency (overall $\alpha = .96$), interrater reliability ($r = .87$), and test retest reliability ($r = .94$). The 53 items are organized under 10 subscales reflecting the essential components of tier 1

implementation (e.g., faculty commitment, expectations and rule developed, classroom systems). Scores for the BoQ are scaled as the percentage of points earned out of the total possible points (107), with scores 70% or above considered implementing with fidelity (Cohen et al., 2007). The BoQ scores are collected each year in April and May for all schools implementing SWPBIS as part of the state-wide implementation project. The BoQ is collected by the school implementation teams in collaboration with the state DOE.

School-level characteristics.

Student demographics. We included 11 student demographic characteristics. First, we captured the total student enrollment for each school, and the percentage of students categorized as White, Black, Hispanic, Asian, Native American, and multi-racial. We also included the percentage of students in each school considered economically disadvantaged, defined as the percentage of students receiving free or reduced lunch. Last, we included the percentage of students receiving special education services (SWD), the percentage of students receiving limited English proficient (LEP) services by their school, and the percentage of students considered migrants by their school.

State Academic Assessment System. The Georgia Student Assessment System is designed to measures how well students have learned the knowledge and skills outlined in the state-adopted content standards in English Language Arts and Mathematics in grades 3 through 5. Students receive a scale score and an Achievement Level designation based on their total test performance in each content area. Achievement levels are as follows: Beginning Learner, Developing Learner, Proficient Learner, and Distinguished Learner. We included the percentage of students at the Proficient Learner and above levels for each school across the four subject tests because this level indicates grade-level performance.

Student Discipline. The state collects and reports on a number of student discipline outcomes. Office discipline referrals (ODR) are the results of 36 different behavioral incidents operationally defined by the state in a discipline matrix designed to provide guidance for schools. During the 2015-2016 school year, 33% of ODR were for incivility, defined as insubordination or disrespect to staff members or other students; includes but is not limited to refusal to follow school staff member instructions, use of vulgar or inappropriate language, and misrepresentation of the truth; and 21% were for disorderly conduct, defined as any act that substantially disrupts the orderly conduct of a school function, substantially disrupts the orderly learning environment, or poses a threat to the health, safety, and/or welfare of students, staff, or others (includes disruptive behaviors on school buses). In addition, we captured the number of in-school suspensions (ISS), out-of-school suspensions (OSS), detentions, bus suspensions for 10 or fewer days, bus suspension for more than 10 days, physical restraints, corporal punishment incidents, juvenile or court referrals, the number of students assigned to an alternative school, other discipline outcome (e.g., call home, school community service), and expulsions. We converted all discipline outcomes to a rate per student by dividing the count for each discipline outcome by the total number of students in each school. As demonstrated in Table 1, most of the severe behavioral outcomes (e.g., corporal punishment, court referral) were infrequent in elementary schools.

Outcome Variables.

Student Perceptions of Bullying. The state conducts an annual school climate survey of all students in grades 3-5. Four items on the survey specifically target students' perception of bullying in their school. Students are asked to respond to the following questions: How often in the past couple of months have older, bigger, more popular, or more powerful kids picked on you

by…(a) hitting or kicking you; (b) spreading rumors about you; (c); threatening you; and (d) picked on you by leaving you out. All four questions are scores in a four-point Likert scale from (0) never, (1) once or twice, (2) a few times, (3) many times, and (4) every day. The item responses were converted to numerical values and summed across the four items, for a possible score of 0 representing never experiencing bullying behaviors to 16 representing experiencing all four types of bullying behaviors every day. We calculated Cronbach's alpha (α) using the full sample of 376,958 students to evaluate the internal consistency of the Bully score. We found $\alpha =$ .827, a mean of 3.29 ($SD = 3.77$) and a range of 0 to 16 for the full sample of students.

Data Analysis

To address the research question, we conducted a quasi-experimental design comparing schools implementing SWPBIS with fidelity (i.e., treatment) to propensity score-matched comparison schools not implementing SWPBIS. This approach allows for estimation of unbiased treatment effects by establishing equivalence on theoretically relevant characteristics (Forston, Verbitsky-Savitz, Kopa, & Gleason, 2012). PSM was conducted at the school-level, then student responses were merged with the PSM dataset.

Propensity score matching. Propensity score matching (PSM) methods are designed to reduce bias in treatment effect estimates in experimental design studies that do not have random assignment of participants to conditions (Leite, 2017). A propensity score is defined as the conditional probability of treatment assignment based on all available covariates (Rosenbaum & Rubin, 1983) and can be used for one-to-one matching treatment to comparison schools. PSM allows for the identification of a covariate equivalent comparison group matched to a treatment group, meeting established standards for high quality QED research, such as those proposed by the What Works Clearinghouse (WWC) evidence standards (2014), and that treatment estimates

have been found to be as accurate as those from randomized controlled trial studies (Forston et al., 2012)

Following procedures outlined by Leite (2017), we estimated propensity scores using logistic regression and 25 school-level covariates. In addition to the 24 characteristics described in Table 1, we included a categorical variable that described the grades each school served (e.g., PK, K, 1st, 2nd, 3rd) to ensure matched schools had the same grade structure. For the logistic regression, we created a dichotomous variable for all schools, where schools that implemented SWPBIS with fidelity were coded as 1, while all other schools were coded as a 0. Then we estimated the predicted probability (p), or propensity score, that a school was in the treatment or control group based on the included covariates ($\log[p/(1-p)]$). Next, we used the estimated propensity scores to match schools using the one-to-one optimal matching method (Rosenbaum, 1989), which minimizes global propensity score distance between treatment and comparison schools. The one-to-one matching procedure identifies a perfect match school for each treatment school so that the treatment and comparison schools are equivalent on all 25 covariates. The one-to-one optimal matching algorithm was conducted using the *matchit* (Ho, Imai, King, Stuart, & Whitworth, 2017) and *optmatch* (Hansen, Fredrickson, Fredrickson, Rcpp, & Rcpp, 2016) packages in R (R Core Team, 2013). To confirm covariate equivalence, we calculated standardized mean difference effect sizes (g), where equivalence is defined as $g < .25$ standard deviations (WWC, 2014).

Mixed-Effects Modeling.

We estimated a mixed-effects model (Bates, Machler, Bolker, & Walker, 2015), also referred to as multilevel models or hierarchical linear models (Raudenbush & Bryk, 2002), to evaluate the effect of SWPBIS implemented with fidelity on students' perceptions of bullying.

We treated both school and district as random effects to account for the nesting of students in schools and districts (i.e., three-level model). First, we estimated an unconditional (null) model for the intra-class correlation coefficient (ICC) to identify how much of the variance in students' perceptions of bullying is attributed to schools and districts. Next, we estimated the effect of SWPBIS implemented with fidelity on the Bullying scale. The final model included covariates of substantive interest, including student's grade level, the percentage of White students, the percentage of economically disadvantaged students, the percentage of students with disabilities, and out-of-school suspensions. In addition, we included interaction terms for each of the four covariates and the treatment indicator. The full mixed-effects model was as follows:

$$y_{ijt} = \gamma_0 + \gamma_1 Z_{ijt} + \sum_{c=1}^{19} \gamma_c X_{cijt} + u_{0j} + u_{0t} + u_{1t} Z_{it} + \varepsilon_{ijt}$$

$$\varepsilon_{ij} \sim N(0, \sigma^2),$$

$$u_{0j} \sim N(0, \tau), u_{0t},$$

$$u_{1t} \sim N(0, \Phi)$$

where y_{ijt} is perceptions of bullying for student i, in school j, in district t, γ_0 is the intercept, Z_{ijt} is a dummy indicator of whether the student attended a school that implemented SWPBIS with fidelity, and γ_1 is the treatment effect. The model included 19 school level covariates that are represented in the model above by X_{cijt} which are related to the outcome through the γ_c coefficients (see covariate names in Table 3). The model has four random effects: u_{0j} is the random intercept of school j with variance τ, u_{0t} is the random intercept and u_{1t} is the random slope of the treatment effect in district t, and ε_{ij} is an individual-level residual with variance σ^2.

The covariance matrix between u_{0t} and u_{1t} is Φ. All mixed-effects models were estimated in lmer4 (Bates et al., 2015) in R 3.1.1. (R Core Team, 2014). All models were estimated using Restricted Maximum Likelihood (REML).

Results

Establishing Equivalence

The 118 schools that implemented SWPBIS with fidelity were propensity score matched to 118 comparison schools using 26 school-level covariates. We used WWC standards, which establish baseline equivalence between a treatment and control group if the difference is less than .25 standard deviation units on all available and conceptually relevant sample characteristics (WWC, 2014). We included current school demographic characteristics and school-level behavioral outcomes, including the rates of ISS, OSS, and ODR per student, from three years prior to the intervention. We matched at the school-level because (a) student-level demographic data was not available and (b) the treatment was school-wide, thus at the school-level. Standardize mean difference effect sizes are presented in Table 1 comparing the treatment and PSM comparison schools. All differences between groups are less than .25, establishing baseline equivalence. Eight of the covariates standardized mean difference was greater than .05, therefore all models should include the covariates per WWC standards.

Treatment Effects

The primary research question in this study was to evaluate the effect of SWPBIS implemented with fidelity on students' perceptions of bullying. Table 2 includes the means and standard deviations for the Bullying full scale score and each individual item. Overall, students in both treatment and comparison schools reported that bullying was, on average, infrequent.

However, the standard deviations suggest that there is variability in student scores that may be accounted for by modeling predictors of bullying.

We estimated a series of mixed effects models to identify a treatment effect, controlling for nesting of students in schools and districts, as well as the eight covariates with equivalence statistics greater than .05 standard deviation units. All models included 77,315 students in 236 schools in 73 school districts. First, we calculated ICC for schools and districts with a fully unconditional model. We found an ICC of 0.04 for between group variance at the school-level, or that 4% of the variance in Bullying is attributable to school-level differences, and an ICC of 0.01 for district-level differences. Results suggest that the majority of variance in the Bullying variable are attributable to individual differences within schools. The ICC results suggest that nesting of students in schools and districts may not impact modeling results. However, we chose to retain the mixed-effect model to account for any potential impact nesting could have on the treatment effect.

As noted, to ensure all models met WWC standards, we included the eight school characteristics with equivalence statistics greater than .05 in all models. The first model examined the treatment effect, controlling for the eight covariates (see Table 3). There was not a significant effect for treatment on students' perceptions of bullying. There was a significant negative effect on bullying for the percentage of students reading at or above benchmark, which suggests that students report less bullying in schools with more students reading at or above benchmark. Next, we modeled the treatment effect, the eight equivalence covariates, and five predictors of substantive interest. Specifically, we examined whether or not there was a relationship between students' perceptions of bullying and the percentage of White students, the percentage of students identified as economically disadvantaged, the percentage of students with

disabilities, the rate of OSS per student, and student grade-level. We also examined the

interaction effects for each of the five substantive characteristics to evaluate whether or not there

was an interaction with each of the variables and implementing SWPBIS with fidelity. Model 2

results, also presented in Table 3, suggests that there was no treatment effect and no interaction

effects for the treatment condition and each of the five substantive characteristics. The significant

relationship between bullying and the percentage of students reading at benchmark remained,

while student grade-level was also significant and negative. Specifically, as students' grade-level

increased, their perceptions of bullying decreased. Lastly, to improve interpretation, we

estimated the Bullying Full Scale marginal means for the treatment and control schools using the

eight covariates with equivalence greater than .05 and calculated the standardized mean

difference (Hedges' g) between the groups. Overall, we found covariate adjusted $g = -.02$.

Discussion

Recently, SWPBIS was considered a promising approach for preventing bullying in

schools (Bradshaw, 2013; 2015). However, very little rigorous empirical research has

specifically examined this relationship. This study was designed to evaluate whether or not

universal SWPBIS has a distal effect on student perceptions of bullying victimization. Using a

quasi-experimental design that established baseline equivalence using propensity score matching,

we found no significant effect, replicating and extending the results of Ward and Gersten (2013).

Although we found a null result, the implications are important and need to be addressed in the

research literature, policy domain, and practical application.

Perhaps the most important finding and implication of this research is the confirmation

that bully prevention programs should be integrated into SWPBIS implementation. Research

demonstrates that SWPBIS implemented with fidelity has positive effects on ODR, suspension,

and academic achievement (Gage, Grasley-Boy, et al., 2018; Gage, Lee, et al., 2018; Gage et al., 2017). Research also suggests that bullying victimization may be reduced when bully prevention programs are integrated within SWPBIS (e.g. Ross & Horner, 2009). Bradshaw (2013) provides a cogent argument for the need to integrate bullying programs into SWPBIS systems, particular because SWPBIS systems focus teaching pro-social behaviors, using data to make decisions, and reinforcing appropriate behaviors (see Good, McIntosh, & Gietz, 2011 for practical advice for integrating bullying programs with SWPBIS).

For example, school-wide social and emotional learning (SEL) programs, coupled with targeted interventions, with a specific focus on bully prevention, could be seamlessly integrated into the SWPBIS framework (Preast, Bowman, & Rose, 2017). At the universal level, SEL has demonstrated decreases in bullying involvement and delinquent behaviors over time (Espelage, Low, Van Ryzin, & Polanin, 2015), including for youth with behavioral disorders and other disabilities (Espelage, Rose, et al., 2015, 2016). At the secondary level, Preast and colleagues (2017) recommend conducting school-wide behavioral screeners to identify students who are at-risk for social and communication skill deficits, which place them at greater risk for bullying involvement (Rose et al., 2011), and providing them with targeted social skills training grounded in critical response approaches (e.g., conversation skills, group task completion, resolving interpersonal conflicts). For non-responders (e.g., students with behavioral disorders), tertiary approaches are more intensive supports and training that meet the individual needs of the student, which may include an FBA and BIP (NASEM, 2016; Rose & Monda-Amaya, 2012).

Limitations

There are a number of limitations that necessitate mention. First, the de-identified dataset contained only student grade and no other student-level characteristic. This does not invalidate

the results as the intervention was delivered at the school-level. However, student-level

characteristics would have allowed us to evaluate whether or not SWPBIS implemented with

fidelity had differential effects on certain students' perceptions of bullying victimization. We

used school-level characteristics as a proxy, but future research should attempt to include

student-level characters to evaluate if there are differential effects at the student level. Second,

we do not have any information about bullying interventions that may or may not have been

implemented in either the treatment or control schools. Third, we do not have an indicator for

how long each of the treatment schools had been implementing SWPBIS or each school's

specific BoQ score. Future research could examine differences in bullying by fidelity, modeling

fidelity as a continuous variable. It is worth noting that we did model the different

implementation levels, comparing schools with higher implementation (Operational) to those

with lower implementation (Emerging). No differences were found. Fourth, the BoQ scores are

based on self-report of the implementation team in collaboration with the DOE. Although we

believe the fidelity scores are accurate, we cannot independently confirm this assumption. Future

research should consider other measures completed by independent observers. Fifth, we do not

have a measure of bullying perpetration to directly compare the results with the Waasdorp et al.

(2012) study results. Last, data for this study come from a single state, therefore, the results may

not generalize. Future research should include a nationally representative sample of students.

Conclusion

SWPBIS is an evidence-based framework for reducing ODR and school suspensions but

may not alone be effective at addressing bullying in schools. Unlike problem behaviors that

typically lead to ODR and suspensions, such as noncompliance or disruptive behavior (Gage,

Lee, et al., 2018), bullying is much more dynamic and involves intentionality, repetition, and an

imbalance of power (Gladden et al., 2014). SWPBIS alone does not address the critical features of bullying and, based on our findings, does not have a distal effect. However, research supports and we believe that integrating evidence-based bully prevention and intervention programs with SWPBIS has the potential to increase positive outcomes for both problem behavior generally and bullying specifically.

References

Blake, J. J., Lund, E. M., Zhou, Q., Kwok, O., & Benz, M. R. (2012). National prevalence rates of bully victimization among students with disabilities in the United States. *School Psychology Quarterly,* 27, 210-222. doi: 10.1037/spq0000008

Bradshaw, C. P. (2015). Translating research to practice in bullying prevention. *American Psychologist*, *70*, 322-332. http://dx.doi.org/10.1037/a0039114

Bradshaw, C. P. (2013). Preventing bullying through positive behavioral interventions and supports (PBIS): A multitiered approach to prevention and integration. *Theory Into Practice*, *52*, 288-295. Doi: 10.1080/00405841.2013.829732

Bradshaw, C. P., Koth, C., Bevans, K., Ialongo, 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. P., Koth, C. W., Thornton, L. A., & Leaf, P. J. (2009). Altering school climate through school-wide positive behavioral interventions and supports: Findings from a group-randomized effectiveness trial. *Prevention science*, *10*(2), 100.

Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2010). Examining the effects of schoolwide

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. doi:10.1177/1098300709334798

Card, N. A., Stucky, B. D., Sawalani, G. M., & Little, T. D. (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. *Child Development, 79,* 1185-1229. doi:10.1111/j.1467-8624.2008.01184.x.

Childs, K. E., Kincaid, D., George, H. P., & Gage, N. A. (2016). The relationship between school-wide implementation of positive behavior intervention and supports and student discipline outcomes. *Journal of Positive Behavior Interventions*, *18*, 89-99.

Cohen, R., Kincaid, D., & Childs, K. (2007). Measuring School-Wide Positive Behavior Support Implementation: Development and Validation of the Benchmarks of Quality (BoQ). *Journal of Positive Behavior Interventions, 9*(4), 203-213. doi:10.1177/10983007070090040301

Cook, C.R., Williams, K.R., Guerra, N.G., and Kim, T. (2010). Variability in the prevalence of bullying and victimization. In S.R. Jimerson, S.M. Swearer, and D.L. Espelage (Eds.), *Handbook of bullying in schools: An international perspective* (pp. 347-362). New York, NY: Routledge.

Dulmus, C., Sowers, K., & Theriot, M. (2006). Prevalence and bullying experiences of victims and victims who become bullies (bully-victims) at rural schools. *Victims & Offenders*, *1*, 15–31. http://dx.doi.org/10.1080/15564880500498945

Espelage, D.L., Low, S., Van Ryzin, M.J., & Polanin, J.R. (2015). Clinical trail of Second Step Middle School Program: Impact on bullying, cyberbullying, homophobic teasing, and sexual harassment perpetration. *School Psychology Review, 44*, 464-479.

Espelage, D. L., Rose, C. A., & Polanin, J. R. (2015). Social-emotional learning program to

      reduce bullying, fighting, and victimization among middle school students with

      disabilities. Remedial and special education, 36(5), 299-311.

Espelage, D. L., Rose, C. A., & Polanin, J. R. (2016). Social-emotional learning program to

      promote prosocial and academic skills among middle school students with

      disabilities. Remedial and Special Education, 37(6), 323-332.

Freeman, J., Simonsen, B., McCoach, D. B., Sugai, G., Lombardino, A., & Horner, R. (2016).

      Relationship between school-wide positive behavior interventions and supports and

      academic, attendance, and behavior outcomes in high schools. *Journal of Positive*

      Behavior Interventions, *18*, 41-51. doi: 10.1177/1098300715580992

Gage, N. A., Grasley-Boy, N., George, H. P., Childs, K., & Kincaid, D. (2018). A quasi-

      experimental design analysis of the effects of school-wide positive behavior interventions

      and supports on discipline in Florida. *Journal of Positive Behavior Interventions*. doi:

      10.1177/1098300718768208

Gage, N. A., Lee, A., Grasley-Boy, N., & George, H. P. (2018). The impact of school-wide

      positive Behavior interventions and supports on school suspensions: A state-wide quasi-

      experimental analysis. *Journal of Positive Behavior Interventions.* doi:

      10.1177/1098300718768204

Gage, N. A., Leite, W., Childs, K., & Kincaid, D. (2017). Average Treatment Effect of School-

      Wide Positive Behavioral Interventions and Supports on School-Level Academic

      Achievement in Florida. *Journal of Positive Behavior Interventions*, *19*, 158-167.

      https://doi.org/10.1177/1098300717693556

Gage, N. A., Whitford, D. K., & Katsiyannis, A. (2018). A review of school-wide positive

behavior interventions and supports as a framework for reducing disciplinary
exclusions. *Journal of Special Education*. doi: 10.1177/0022466918767847

Hanish, L.D., & Guerra, N.G. (2000). The roles of ethnicity and school context in predicting
children's victimization by peers. *American Journal of Community Psychology, 28*, 201-
223.

Hong, J. S., & Espelage, D. L. (2012). A review of research on bullying and peer victimization in
school: An ecological system analysis. *Aggressive and Violent Behavior, 17*, 311-322.
doi:10.1016/j.avb.2012.03.003

Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A., & Esperanza, J.
(2010). A randomized, wait-list controlled effectiveness trial assessing school-wide
positive behavior support in elementary schools. *Journal of Positive Behavior
Interventions, 11*, 133-144. doi: 10.1177/1098300709332067

Gladden, R. M., Vivolo-Kantor, A. M., Hamburger, M. E., & Lumpkin, C. D. (2014). *Bullying
surveillance among youths: Uniform definitions for public health and recommended data
elements, Version 1.0*. Atlanta, GA; National Center for Injury Prevention and Control,
Centers for Disease Control and Prevention, and U. S. Department of Education.

Good, C. P., McIntosh, K., & Carmen, G. (2011). Integrating bullying prevention into
schoolwide positive behavior support. *TEACHING Exceptional Children*, *44*, 58-56.

Hymel, S., & Swearer, S. M. (2015). Four decades of research on school bullying: An
introduction. *American Psychologist, 70*, 293-299. http://dx.doi.org/10.1037/a0038928

Leite, W. (2017). *Propensity score methods using R*. Los Angeles, CA: Sage.

McLaughlin, C., Byers, R., & Vaughn, R. P. (2010). *Responding to bullying among children with
special educational needs and/or disabilities.* London: Anti-Bullying Alliance.

Musu-Gillette, L., Zhang, A., Wang, K., Zhang, J., Kemp, J., Diliberti, M., and Oudekerk, B.A.

(2018). Indicators of School Crime and Safety: 2017 (NCES 2018-036/NCJ 251413).

National Center for Education Statistics, U.S. Department of Education, and Bureau of

Justice Statistics, Office of Justice Programs, U.S. Department of Justice. Washington,

DC.

Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B., & Scheidt, P. (2001).

Bullying behaviors among US youth: Prevalence and association with psychosocial

adjustment. *JAMA: Journal of the American Medical Association*, *285*, 2094–2100.

http://dx.doi.org/10.1001/jama.285.16.2094

National Academies of Sciences, Engineering, and Medicine. (2016). *Preventing Bullying*

Through Science, Policy, and Practice. Washington, DC: The National Academies Press.

doi: 10.17226/23482.

Nese, R. N. T., Horner, R. H., Dickey, C. R., Stiller, B., & Tomlanovich, A. (2014). Decreasing

bullying behaviors in middle school: Expect Respect. *School Psychology Quarterly*, *29*,

272-286. http://dx.doi.org/10.1037/spq0000070

Olweus, D. (1993). *Bullying at school: What we know and what we can do*. Oxford: Blackwell.

Olweus, D., Limber, S. P., Flerx, V. C., Mullin, N., Riese, J., & Snyder, M. (2007). *Olweus*

Bullying Prevention Program: Schoolwide guide. Center City, MN: Hazelden.

Preast, J., Bowman, N., & Rose, C.A. (2017). Creating Inclusive Classroom Communities

through Social and Emotional Learning to reduce social marginalization among students

with or at-risk for disability identification. In C.M. Curran & A.J. Peterson (Eds),

Handbook of Research on Classroom Diversity and Inclusive Education Practice (183-

200). Hershey, PA: IGI Global.

Rose, C. A., & Espelage, D. L. (2012). Risk and protective factors associated with the bullying

involvement of students with emotional and behavioral disorders. *Behavioral Disorders*, *37*(3) 133 – 148.

Rose, C. A., & Gage, N. A. (2017). Exploring the involvement in bullying among individuals with and without disabilities over time. *Exceptional Children, 83*, 298-314.

Rose, C. A., & Monda-Amaya, L. E. (2012). Bullying and victimization among students with disabilities: Effective strategies for classroom teachers. *Intervention in School and Clinic*, *48,* 99 – 107. doi: 10.1177/1053451211430119.

Rose, C. A., Monda-Amaya, L. E., & Espelage, D. L. (2011). Bully perpetration and victimization in special education: A review of the literature. *Remedial and Special Education, 32*, 14-130. doi:10.1177/0741932510361247

Rose, C.A., Nickerson, A.B., & Stormont, M. (2015). Advancing bullying research from a social-ecological lens: An introduction to the special issue. *School Psychology Review, 44*, 339-352.

Ross, S. W., & Horner, R. H. (2009). Bully prevention in positive behavior support. *Journal of Applied Behavior Analysis*, *42*, 747-759.

Ross, S. W., & Horner, R. H., (2014). Bully prevention in positive behavior support: Preliminary evaluation of third-, fourth-, and fifth grade attitudes toward bullying. *Journal of Emotional and Behavioral Disorders, 22*, 225-236. Doi: 10.1177/1063426613491429

Spriggs, A.L., Iannotti, R.J., Nansel, T.R., & Haynie, D.L. (2007). Adolescent bullying involvement and perceived family, peer, and school relations: Commonalities and differences across race/ethnicity. *Journal of Adolescent Health*, 41(3), 283-293.

Sullivan, R. B., & Stoner, G. (2012). Developmental and gender differences in elementary school children's recognition of bullying. *Pastoral Care in Education, 30*(2), 113-125.

Swearer, S. M., Espelage, D. L., Vaillancourt, T., & Hymel, S. (2010). What can be done about

school bullying? Linking research to educational practice. *Educational Researcher, 39,*

38 – 47. http://dx.doi.org/ 10.3102/0013189X09357622

Swearer, S. M., Wang, C., Maag, J. W., Siebecker, A. B., & Frerichs, L. J. (2012).

Understanding the bullying dynamic among students in special and general education.

Journal of School Psychology, 50, 503-520. doi:10.1016/j.jsp.2012.04.001

Ttofi, M. M., & Farrington, D. P. (2011). Effectiveness of school-based programs to reduce

bullying: A systematic and meta-analytic review. *Journal of Experimental Criminology*,

7, 27-56.

Vaillancourt, T., Hymel, S., & McDougall, P. (2003). Bullying is power: Implications for school-

based intervention strategies. *Journal of Applied School Psychology*, *19*, 157–176.

http://dx.doi.org/10.1300/J008v19n02_10

Waasdorp, T. E., Bradshaw, C. P., & Leaf, P. J. (2012). The impact of schoolwide positive

behavioral interventions and supports on bullying and peer rejection: A randomized

controlled effectiveness trial. *Arch Pediatr Med.*, *166*, 149-156.

Ward, B., & Gersten, R. (2013). A randomized evaluation of the Safe and Civil Schools model

for positive behavioral interventions and supports at elementary schools in a large urban

school district. *School Psychology Review*, *42*, 317-333.

Yell, M. L., Katsiyannis, A., Rose, C. A., & Houchins, D. E. (2016). Bullying and harassment of

students with disabilities in schools: Legal considerations and policy formation. Remedial

and Special Education, 37(5), 274-284.

Table 1.
Demographic Characteristics of All Treatment Groups

Demographic	All Possible Comparison Schools (*n* = 905)		PSM Comparison Schools (*n* = 118)		Treatment Schools (*n* = 118)		Equivalence
	M	*SD*	*M*	*SD*	*M*	*SD*	
Total Enrollment	605.98	231.03	648.47	214.18	649.91	261.48	0.01
% White	41.44	29.50	45.33	28.42	44.04	29.70	-0.04
% Black	36.99	30.85	33.26	27.80	32.74	27.21	-0.02
% Hispanic	14.31	16.54	14.40	17.40	16.20	15.90	0.11
% Asian	3.43	7.10	2.68	5.73	2.66	4.17	0.00
% Native American	0.14	0.36	0.03	0.18	0.07	0.25	0.18
% Multi Racial	3.59	1.99	4.07	2.15	4.09	2.47	0.01
% ED	67.11	29.24	69.09	26.89	69.32	23.26	0.01
% SWD	11.13	4.88	11.23	3.39	11.25	3.05	0.01
% LEP	11.08	15.26	11.22	16.92	12.31	15.37	0.07
% Migrants	0.34	1.49	0.30	1.06	0.25	1.06	-0.05
% At or Above Benchmark: Reading	36.12	18.11	35.01	16.11	35.80	13.63	0.05
% At or Above Benchmark: Math	38.53	18.93	38.50	16.14	39.13	14.80	0.04
ODR	0.19	0.29	0.19	0.17	0.19	0.23	0.00
ISS	0.05	0.17	0.05	0.09	0.04	0.10	-0.10
OSS	0.07	0.11	0.06	0.07	0.06	0.06	0.00
Detention	0.00	0.02	0.00	0.02	0.00	0.02	0.00
Bus Suspension	0.03	0.07	0.04	0.06	0.03	0.04	-0.20
Physical Restraint	0.00	0.02	0.00	0.01	0.00	0.01	0.00
Corporal Punishment	0.01	0.04	0.00	0.02	0.01	0.02	0.06
Juvenile or Court Referral	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Assigned to Alternative School	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Other Disciplinary Actions	0.03	0.07	0.03	0.05	0.04	0.11	0.12
Expulsions	0.00	0.00	0.00	0.00	0.00	0.00	0.01

Note. PSM is propensity score matched, equivalence was calculated for the PSM comparison and treatment schools and is the standardized mean difference and established if less than .25 standard deviation units. ED is economically disadvantaged, SWD is students with disabilities, LEP is limited English proficient, ODR is office discipline referral, ISS is in-school suspension, OSS is out-of-school suspension.

Table 2.

Means and Standard Deviations for all the Bullying Scale and Individual Bullying Items

Bullying Item	Treatment Schools (n = 39,187)		PSM Comparison Schools (n = 37,061)	
	M	SD	M	SD
Bullying Full Scale	3.27*	3.76	3.34*	3.78
How often in the past couple of months have older, bigger, more popular, or more powerful kids picked on you by:				
(a) Hitting or kicking you	0.68	1.09	0.71	1.11
(b) Spreading rumors about you	0.87	1.18	0.86	1.17
(c) Threatening you	0.66	1.08	0.67	1.10
(d) Picked on you by leaving you out	1.06	1.26	1.07	1.26

Note: The range of scores for the Bullying Full Scale measure is 0 to 16, the range for the three individual items is 0 to 4. Sample sizes are students attending the 188 treatment schools and 118 PSM comparison schools. *The means are covariate adjusted for the eight covariates with equivalence > 0.05.

Table 3.
Three-Level Random Effects Model of SWPBIS Implemented with Fidelity Predicting Student Perceptions of Bullying

Parameters	Model 1		Model 2	
	Estimate	s.e.	Estimate	s.e.
Fixed Effects:				
Intercept	4.56***	0.16	5.12***	0.52
SWPBIS	-0.03	0.10	0.72	0.52
% White			0.00	0.00
% ED			0.00	0.00
% SWD			0.01	0.02
OSS			0.00	0.00
4th Grade			-0.27***	0.05
5th Grade			-0.71***	0.05
SWPBIS x % White			-0.01	0.00
SWPBIS x % ED			0.00	0.00
SWPBIS x % SWD			0.00	0.03
SWPBIS x OSS			0.00	0.00
SWPBIS x 4th Grade			-0.07	0.06
SWPBIS x 5th Grade			-0.01	0.07
% Hispanic	-0.01	0.01	-0.02*	0.01
% Native American	-0.07	0.19	-0.05	0.19
% LEP	0.00	0.01	0.00	0.01
% Migrant	-0.10*	0.05	-0.07	0.05
% Reading	-0.03***	0.00	-0.03***	0.01
ISS	-0.42	0.52	-0.61	0.52
Bus Suspension	0.55	0.98	0.98	0.97
Corporal Punishment	-2.42	2.26	-1.77	2.28
Other Discipline	1.03	0.57	0.59	0.63
Random Effects:				
School	0.34		0.34	
District	0.05		0.03	
Residual	13.74		13.66	

Note. Significant estimates are in bold case with *p*< .05 *, *p*<.01 **, *p*<.001 ***; 77,315 students, 236 schools, 73 school districts. ED is economically disadvantaged, SWD is students with disabilities, LEP is limited English proficient, ISS is in-school suspension, OSS is out-of-school suspension, SWPBIS is school-wide positive behaver interventions and supports. The reference group for grade-level is 3rd grade.