A Cluster Randomized Controlled Trial of the Safe Public Spaces in Schools Program: Final Summary Overview

NIJ Grant 2015-CK-BX-0012

SEPTEMBER 2019

Kimberly Kendziora | Juliette Berg | Jesse Levin | Lynn Hu | Amanda Danks | David Osher, *AIR*

Sarah Klevan, *Research Alliance for New York City Schools* | Louis Tuthill, *Azusa Pacific University* | Nancy Guerra, *University of California Irvine*

MAKING RESEARCH
RELEVANT

A Cluster Randomized Controlled Trial of the Safe Public Spaces in Schools Program: Final Summary Overview

NIJ Grant 2015-CK-BX-0012

SEPTEMBER 2019

Kimberly Kendziora | Juliette Berg | Jesse Levin | Lynn Hu | Amanda Danks | David Osher
AIR

Sarah Klevan
Research Alliance for New York City Schools

Louis Tuthill
Azusa Pacific University

Nancy Guerra
University of California Irvine

9685_09/19

Contents

Tables

Figures

Purpose

The Comprehensive School Safety Initiative aims to improve the safety of schools and students nationwide by developing, supporting, and rigorously evaluating school safety programs, practices, and strategies. This study supports these aims by testing the efficacy of an intervention focused on the school environment. Although theories of school violence attribute a substantial part of the causes of violence to environmental factors, relatively few interventions have targeted public spaces in schools (Lipsey, Howell, Kelly, Chapman, & Carver, 2010), even though bullying and violent behavior are more common in out-of-classroom spaces than inside the classroom (Musu, Zhang, Wang, Zhang, & Oudekerk, 2019; Perkins, Perkins, & Craig, 2014; Vaillancourt et al., 2010). Researchers have observed that rule violations occur more frequently in areas of high student density (Cash, Bradshaw, & Leaf, 2015), and fewer negative behaviors occur in in the presence of adult supervision (Cash, Debnam, Waasdorp, Wahl, & Bradshaw 2019). This study examined the quality of implementation, cost, and impact of the Safe Public Spaces program (SPS), as well as its effect on the displacement of crime or diffusion of benefits to the area immediately outside the school (Johnson, Guerette, & Bowers, 2014).

Intervention

SPS, developed and delivered by Engaging Schools, establishes a safe and supportive whole-school environment by equipping all staff with the knowledge and skills they need to prevent, analyze, interpret, manage, and respond to a wide range of disciplinary infractions at their school. In 2016 and 2017, **SPS schools** participated in half-day summer retreats for administrators and the Student Support Team to review diagnostic data and develop a plan for

Safe Public Spaces in the following school year. In August and September of 2016, Engaging Schools' consultants facilitated separate training sessions customized to the roles of attendees (e.g., teachers, cafeteria workers, school safety agents). The sessions ranged from 2 to 4 hours in length and included distribution of the program manual. In the fall of 2017, Engaging Schools trained new staff in the SPS program, and returning staff reviewed data from the previous school year and participated in a refresher course. Throughout the 2016–17 and 2017–18 school years, Engaging Schools' consultants facilitated data-informed group training sessions to support ongoing capacity development of staff to implement promotion, prevention, and intervention strategies. Engaging Schools' consultants observed, shadowed, and coached staff to support safe public spaces and increase the use of promotion, prevention, and intervention strategies. Engaging Schools' consultants worked with administrators to model the practice of coaching and gradually release these coaching responsibilities to administrators.

At the end of Year 2 (spring 2018), **control schools** received a half-day on-site consultation with school leadership teams regarding making public spaces safer, a diagnostic tool, and a protocol for assessing public spaces; a 2-day institute for a leadership team from the school composed of administrative leaders, student support staff, teacher leaders, and school safety agents; a half-day follow-up consultation visit; downloadable copies of the *Safe Public Spaces Program Guide* for the entire staff, and three copies of *Shifting Gears: Recalibrating Schoolwide Discipline and Student Support.*

Setting

This study was carried out in 24 New York City (NYC) middle schools that were recruited in cooperation with the NYC Department of Education's (DOE's) Office of Safety and Youth

Development and the Mayoral Leadership Team on School Climate and Discipline. In March of 2016, the CEO of that office sent a letter to middle schools with higher than average levels of discipline incidents describing the study and inviting interested schools to contact the study team for more information. Once a pool of interested schools was identified, schools were statistically matched into pairs on the basis of characteristics including grade levels offered, school enrollment, racial composition, the percentage of students qualifying for free or reduced-price meals, English language learner programs, and programs for students with disabilities. One school in each pair was randomly assigned to SPS and the other school to control, for a total of 12 in each condition. Schools implemented the SPS program in the 2016–17 and 2017–18 school years.

There were several changes to the study sample because of matched-pair attrition and replacement and, in one case, a school no longer meeting eligibility criteria for the study (this school transitioned from a K–8 school to a K–5 school). In addition to these changes, two schools declined to participate in the SPS training beginning in Year 2, although they agreed to continue participation in both the impact study and the process evaluation. This study used an intent-to-treat approach and included all 24 randomized schools in the analysis.

Design, Methods, and Analysis

The study comprises four components: an implementation evaluation, a cost study, an impact study, and a community crime study. Following a description of the intervention, each of these components is discussed in turn.

Implementation Evaluation

The implementation evaluation was designed to provide deeper understanding of how well the SPS program was administered, how school staff experienced the program, and whether there were meaningful differences in safety practices between schools participating in SPS and control schools. The implementation evaluation addressed three research questions:

1. To what extent is implementation consistent with the design of SPS as specified by Engaging Schools?

2. How did school staff perceive the successes and challenges associated with implementing SPS?

3. To what extent do SPS and control schools differ with respect to safety approaches and practices that are essential to the SPS model?

Implementation Evaluation Methods

The implementation evaluation consisted of two forms of data collection: structured observations of schools' public spaces and interviews with school administrators. We conducted **observations** in the fall and spring of the first year of implementation and four times in the second year of implementation (twice in the fall of 2017 and twice in spring of 2018). Observations were structured on the basis of the SPS manual, and they assessed four key strategies: (1) the deliberate placement of staff at transition hot spots, (2) classroom Meet and Greets, (3) Effective Reminders and Directives, and (4) Hall Scans. Three additional SPS strategies, defusing students who are emotionally charged, handling high-impact incidents, and addressing persistent hall walkers, were deemed too low-frequency to include in the observation protocols. The closed-ended observation protocol allowed for quantifiable ratings

of the frequency and quality of artifacts, activities, and behaviors aligned with essential SPS program components. Observations included transitions between class periods, public spaces during class periods, the cafeteria during lunch, and a hot spot identified by the school staff as highly congested or potentially unsafe. In each setting, observers recorded the use of effective and ineffective staff Reminders and Directives. Finally, observers conducted walkabout to note signage for school expectations and traffic flow, the amount of litter in the hallways, and damage to school property. Interrater agreement on all categories was higher than 80%.

We conducted **interviews** with administrators in all study schools in the fall and spring of both school years. The interview protocol included questions on safety-related topics, such as the challenges that schools faced with respect to school safety; the roles that various school staff played in maintaining safety in schools' public spaces; the type of safety-related data that schools collected; school safety practices for class transitions and hall walking; and school practices for resolving high-impact incidents, such as physical altercations or assisting an emotionally charged student. In SPS schools, we also included questions pertaining specifically to the SPS program.

Implementation Evaluation Analysis

We cleaned and checked all **observation data** to ensure accurate data entry. We assigned a binary score to all variables—0 if the event did not occur and 1 if the event occurred—and then we calculated descriptive statistics for all variables. We transcribed all **interviews** and conducted a narrative analysis to identify themes related to our primary research questions. We also aggregated data related to interview questions into two documents, one for SPS schools and another for control schools. In addition, we conducted multiple read-throughs to identify

emerging themes related to the study's questions. Analysts noted themes within and across SPS and control schools, noted any surprising or inconsistent information, and identified quotes to illustrate themes. Results were summarized in narrative form.

Cost Study

The cost study aimed to accurately estimate the overall cost per student of the SPS activities. Using a subsample of three SPS and two control schools, we computed the average cost per student using the Ingredients Approach (Levin, McEwan, Belfield, Bowden, & Shand, 2018).

Cost Study Methods

We collected comprehensive resource descriptions and specifications via interviews with administrative staff involved in safety activities at each of five sampled study schools. Use of a subsample minimized burden on school staff while still capturing essential information about resources needed for implementation. The semistructured interviews addressed personnel, facilities, materials, equipment, and other inputs. We addressed follow-up questions as necessary by e-mail or phone. For each school, the study team used this information, as well as reference information gathered from NYC DOE data on salaries for position types for which these data were available and district salary schedules for the other staff types. Census data were used to calculate benefits. This information populated a Resource Cost Model (RCM), which organized resource specifications and the associated prices to produce accurate cost calculations.

Cost Study Analysis

Information describing the way schools used each ingredient for safety-related activities in SPS and control schools was entered into the site-specific RCMs. Personnel information

included position type (i.e., associated salary and benefit compensation) and the amount of time devoted to the program. Personnel costs multiplied the cost by the quantity. Non-personnel ingredient specifications included details such as the type, lifespan, and quantity of each resource, and included a 3% rate of depreciation (except for facilities). Cost calculations for facilities included relevant inflation rates from the Bureau of Labor Statistics.[1] Each site's RCM calculated per-student cost for each category of analysis (i.e., overall cost, personnel and non-personnel, strategic component, staff position type). We then used those figures with enrollment data to produce weighted average per-student costs for each analysis component for the SPS and control schools.

Impact Evaluation

To measure the effect of SPS on student and school outcomes, we conducted a randomized trial. Twenty-four schools, all interested in the program and willing to be randomized, were pair-matched according to school characteristics using propensity scores, and then schools in each pair were randomly assigned to SPS or control conditions. Questions for this study were as follows:

1. Do SPS schools provide a safer environment for their students than schools not implementing the program? This will be evident by students' and teachers' reports on safety and the number and severity of incidents of problem behavior in public spaces, including violence and delinquency.

---

[1] Specifically, to inflate the facilities costs, we used the Bureau of Labor Statistics' Employment Cost Index of wages and salaries of workers in the natural resources, construction, and maintenance available, which are available at https://beta.bls.gov/dataViewer/view/timeseries/CIS2020000400000I.

2. To what extent does the Safe Public Spaces program affect other student outcomes—specifically, attendance and achievement—that are linked to problem behavior and sense of safety at school?

Impact Evaluation Methods

The impact evaluation was conducted using data from the Research Alliance for New York City Schools (Research Alliance). Among the core functions of the Research Alliance is maintaining a unique archive of longitudinal data on NYC schools to support ongoing research. The Research Alliance created stripped data sets in accordance with their data security agreement with the NYC DOE, and AIR analyzed those data on their servers. Disciplinary incident data included both total incidents in the school and just those occurring in public spaces; for each of these, we obtained information about whether the incident resulted in a suspension and whether it was a Level 4 or 5 event. Student-level data included mathematics and reading achievement test scores and attendance. In addition, we examined four student-reported variables from NYC's School survey: safety, student-teacher trust, classroom behavior, and bullying.

Finally, we conducted exploratory analyses to determine whether implementation quality affected treatment effects. We tested two measures of implementation: observation data from the implementation study and quality of school leadership data from the NYC School Survey. We tested the latter variable because discussions with Engaging Schools suggested that school leadership was an important factor in SPS implementation.

Impact Evaluation Analysis

We assessed in several ways whether SPS schools provided a safer environment for their students than did control schools. First, we looked at the impact of SPS on the proportion of all

incidents in Year 1 and Year 2 of the study and the proportion of incidents in public spaces. Public spaces included the auditorium, bathroom, cafeteria, lobby, library, hallways, gymnasium, entrance and exit, and close proximity to the school. Second, we looked at the impact of Safe Public Spaces on the severity of all incidents in Year 1 and Year 2 and the severity of incidents in public spaces. The severity of incidents was measured using NYC DOE's five-point system of rating the intensity of incidents, ranging from insubordinate behaviors (1) to seriously dangerous or violent behaviors (5). Third, on the basis of the annual NYC School Survey, we looked at the impact of the intervention on student reports of safety and adult support.

For each of these analyses, we used ordinary least squares regression analysis with student-level data aggregated to the school level. The models included dummy variables for each of the matched pairs and the baseline (2015–16) score of the outcome variable. For student-level exploratory analyses, we used two-level hierarchical modeling with students (L1) nested in schools (L2). All models included dummy variables for each of the matched pairs and the following student covariates: gender, race/ethnicity, grade level, special education status, and a poverty flag. Incident rates included nonunique incidents, meaning that students were counted as many times as they had incidents. The proportion of incidents and suspensions were computed as the average number of incidents and suspensions in the school in any given year. Effect sizes were computed using Hedge's *g*, which is the adjusted difference in means between the treatment and control conditions divided by the pooled weighted standard deviation of the outcome with a correction factor for small sample sizes.

Community Crime Study

The community crime study examined the frequency of juvenile arrests before, during, and after SPS implementation and examined evidence for displacement of juvenile crime or diffusion of benefits to the geographic area immediately around the schools (Braga, Papachristos, & Hureau, 2014; Wang, Liu, & Eck, 2014). However, most of this research has focused on serious crimes such as street violence and drug markets (Johnson, Guerette, & Bowers, 2014; Weisburd, Telep, Hinkle, & Eck, 2010). Does the SPS program affect the community around the school?

Community Crime Study Methods

The community crime study used data on the arrest of juveniles from the NYPD (New York Police Department) data. Data included all arrests for the juvenile crime during the 2-year period of SPS implementation, plus a 1-year follow-up period (through June 2019). The 12 pairs of matched schools were identified and geo-mapped using Quantum GIS software. We mapped block groups in the 2010 U.S. Census in which the schools reside and neighboring block groups into micro-areas. This resulted in 12 experimental school blocks and 11 control blocks in which the schools resided (two of the control schools were in the same census block group). In addition, neighboring blocks were geo-mapped into 70 experimental and 77 control adjacent block groups. Finally, juvenile arrests were mapped into experimental and control areas.

Community Crime Study Analysis

We used auto-regressive integrative moving average (ARIMA) to assess the effects of the initiative on crime. We smoothed the data using Holt-Winters additive model to account for variation in the data. Using the ARIMA time-series method in the Stata 15 statistical software

package, we analyzed arrest data to compare the change in juvenile arrests in the experimental and control sites.

Findings

In this section, we describe the results of our study for each component in turn.

Implementation Evaluation

The implementation evaluation found that SPS was well-implemented in schools, but that comparison schools also had high levels of safety activities similar to key components of the SPS program. School staff reported positive impressions of the program.

To address our research question about fidelity, we examined observation data across both Year 1 and Year 2 to compare SPS and control schools. SPS schools had higher levels of Meet and Greets (observed on 64% of occasions in SPS schools versus 46% in control schools), strategically placed staff in the cafeteria (96% in SPS schools, 86% in control schools), effective reminders in both the hot spot (76% in SPS schools, 61% in control schools) and the cafeteria (87% in SPS schools, 79% in control schools), posting of expectations (71% in SPS schools, 48% in control schools), and traffic flow signs (68% in SPS schools, 45% in control schools) in the cafeteria. Students in SPS schools were also less likely to walk through the halls without a hall pass (78% in SPS schools, 88% in control schools). Last, staff in SPS schools were less likely to use ineffective reminders in both the selected hot spot (18% in SPS schools, 22% in control schools) and cafeteria (39% in SPS schools, 44% in control schools).

We also observed generally flat to decreasing quality of SPS implementation in Year 2, compared with the quality of implementation in Year 1. For example, Meet and Greets (Year 1 = 65%, Year 2 = 62%), staff placed in the cafeteria (Year 1 = 100%, Year 2 = 93%), and effective

reminders in the cafeteria (Year 1 = 91%, Year 2 = 81%) all dropped somewhat. We observed positive growth in only 2 of 12 Year 2 variables: the use of effective reminders in hot spots (Year 1 = 69%, Year 2 = 79%), and the posting of rules and expectations in the cafeteria (Year 1 = 58%, Year 2 = 75%).

Regarding perceptions of successes and challenges, interview data revealed that staff in SPS schools reported greater levels of safety-related data collection and analysis than did control schools, and described the ways in which new data collection and analysis practices led to institutional changes at their schools. SPS schools also attributed improved relationships between teachers and students to involvement in the SPS program. We found no differences between the ways that SPS and control schools defused emotionally charged students, handled high-impact incidents, or managed persistent hall walking, which were also key strategies of the SPS model.

Finally, with respect to how SPS and control schools differed with respect to safety approaches and practices that are essential to the SPS model, we found generally high levels of safety practices in control schools, but not quite as high as in SPS schools. The high level of activities in control schools

Cost Study

Cost estimates are generally reported within a range, called the lower and upper bounds. Because many of the activities in the SPS schools were similar to those implemented in the control schools, the lower bound cost estimate of the program was captured as the difference between the cost associated with the activities of the SPS schools and those activities of control schools: $Cost_{SPS} = Cost_{Treatment} - Cost_{Control}$. The lower-bound cost estimate of

SPS is $973 per student. The overall upper-bound cost estimate, calculated as the cost of all resources required to implement the SPS, was $2,859 per student. When SPS can take the place of some safety-related activities, middle-ground costs will fall between the upper and lower bound estimates and will likely vary according to the context of each school.

Impact Evaluation

The impact evaluation did not detect any statistically significant effects of SPS on student outcomes.

- The proportion of disciplinary incidents overall was not significantly different in SPS schools compared with that of control schools in either Year 1 or 2. The coefficient in Year 1 is negative because the incident rates in SPS schools were consistently higher than those in controls, including the year prior to implementation.

- Schools implementing SPS did not have significantly lower suspension rates than control schools.

- Schools implementing SPS did not have significantly higher scores on student-reported safety, student-teacher trust, classroom behavior, or bullying than schools not implementing SPS in either implementation year.

Table 1 shows the data supporting these findings; descriptive data on study variables may be found in the appendix.
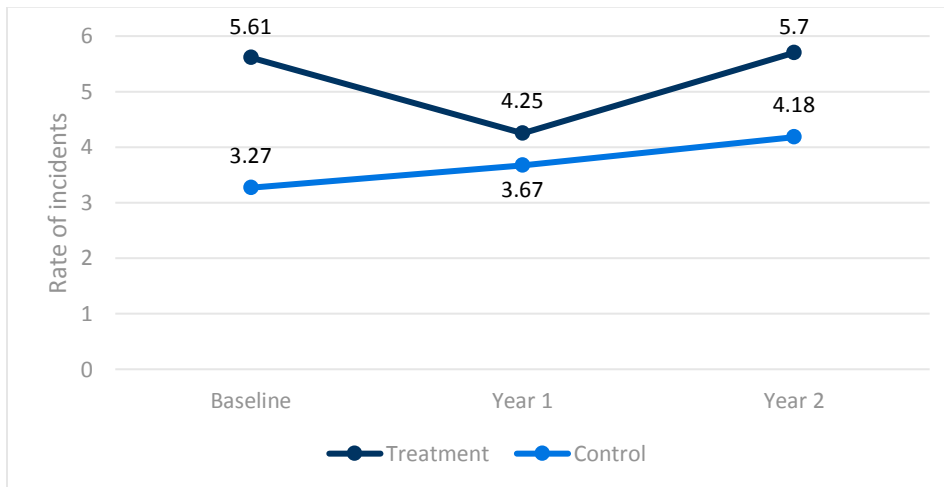
Table 1. Confirmatory Analyses: Impacts on Schools, Years 1 and 2

School outcome variable	Coefficient	Standard error	Proba-bility	Confidence interval		Effect size
Year 1						
Percentage of incidents overall	2.35	6.61	0.73	-12.38	17.07	0.09

School outcome variable	Coefficient	Standard error	Proba-bility	Confidence interval		Effect size
Percentage of incidents in public spaces	1.86	3.40	0.60	-5.71	9.42	0.16
Incidents resulting in suspensions overall	-1.66	2.11	0.45	-6.37	3.05	0.28
Incidents in public spaces resulting in suspensions	-0.46	1.29	0.73	-3.34	2.42	0.15
Level 4 and 5 incidents	-1.70	1.76	0.36	-5.62	2.22	0.21
Level 4 and 5 incidents in public spaces	-0.81	1.21	0.52	-3.50	1.88	0.26
Number of days suspended overall	-0.49	1.74	0.78	-4.43	3.45	0.11
Number of days suspended from public space incidents	2.64	2.24	0.27	-2.41	7.70	0.44
Student-reported safety	0.01	0.03	0.75	-0.05	0.07	0.05
Student-reported student–teacher trust	0.02	0.03	0.38	-0.03	0.08	0.20
Student-reported classroom behavior	0.02	0.04	0.69	-0.10	0.07	0.21
Student-reported bullying	0.03	0.04	0.50	-0.06	0.12	0.14
Year 2						
Percentage of incidents overall	6.55	5.14	0.23	-4.90	17.99	0.28
Percentage of incidents in public spaces	3.71	3.46	0.31	-4.00	11.43	0.31
Incidents resulting in suspensions overall	1.79	1.93	0.37	-2.50	6.08	0.35
Incidents in public spaces resulting in suspensions	1.44	1.16	0.25	-1.15	4.03	0.48
Level 4 and 5 incidents	-0.46	1.24	0.72	-3.23	2.31	0.08
Level 4 and 5 incidents in public spaces	0.59	1.19	0.63	-2.05	3.23	0.24
Number of days suspended overall	-1.04	1.76	0.57	-4.96	2.89	0.23
Number of days suspended from public space incidents	-2.42	1.53	0.15	-5.89	1.04	0.46
Student-reported safety	0.07	0.08	0.38	-0.10	0.25	0.39
Student-reported student–teacher trust	0.08	0.06	0.18	-0.04	0.20	0.49
Student-reported classroom behavior	0.08	0.08	0.33	-0.10	0.26	0.43
Student-reported bullying	0.07	0.08	0.40	-0.11	0.24	0.31

The pattern of the discipline incident data (Figure 1) shows that the SPS and control groups were noticeably different across all years of the study, with incidents more frequent in SPS schools. Incidents declined in Year 1, but then bounced back up in Year 2 to roughly the level where they started. This finding is consistent with implementation study findings showing that the quality of SPS implementation declined somewhat in Year 2. It is also consistent with a high degree of noise in the incident data: suspensions in the baseline year were entirely uncorrelated with suspensions at Year 2.

Figure 1. Incidents in Public Spaces Resulting in Suspensions

We also found that students in schools implementing SPS did not have significantly higher attendance or achievement scores than students in schools not implementing SPS in either implementation year (see Table 2).

Table 2. Exploratory Analyses: Impacts on Students, Years 1 and 2

Student outcome variable	Coefficient	Standard error	Probability	Confidence interval		Effect size
Year 1						
Attendance	-0.24	0.32	0.46	-0.86	0.38	0.02
Mathematics achievement	0.06	0.04	0.14	-0.02	0.15	0.06
English language arts achievement	0.02	0.02	0.44	-0.03	0.06	0.02

Student outcome variable	Coefficient	Standard error	Probability	Confidence interval		Effect size
Year 2						
Attendance	-0.06	0.35	0.87	-0.74	0.62	0.00
Mathematics achievement	0.01	0.05	0.79	-0.08	0.11	0.06
English language arts achievement	0.02	0.03	0.58	-0.05	0.09	0.03

Our tests of whether implementation related to treatment effects used two measures: observation data from the implementation study and quality of school leadership data from the NYC School Survey. For the observation data, we used average Year 1 observation scores to test Year 1 treatment effects and average Year 2 observation scores to test Year 2 treatment effects.

Effects of observed implementation quality on several outcomes were significant, but all of these suggested that better implementing SPS schools had worse outcomes, such as total days suspended in Year 1 and Year 2, days suspended because of incidents in public spaces in Year 2, and student-reported safety and bullying prevention in Year 1. Higher rated quality of school leadership was associated with fewer suspensions in Year 2 (as well as fewer suspensions resulting from incidents in public spaces), but in Year 1 better leadership was associated with more days suspended overall and more days suspended from public space incidents. We caution that these findings may be subject to the same nonequivalence effects due to noisy data that appeared to affect the impact models.

Community Crime Study

Counts of juvenile arrests in SPS and control sites (Figure 2) showed that there were generally more arrests in control sites than there were in SPS sites. Trends in both SPS and control sites were downward prior to the start of the intervention. SPS school sites showed 24% increase in juvenile arrests during the period of the intervention; control sites showed a 10%

decrease during this period. After the intervention, arrests increased slightly in control sites and declined slightly in SPS sites. These differences were not statistically significant.

Figure 2. Counts of Juvenile Arrests in SPS and Control Sites, January 2015 to May 2019
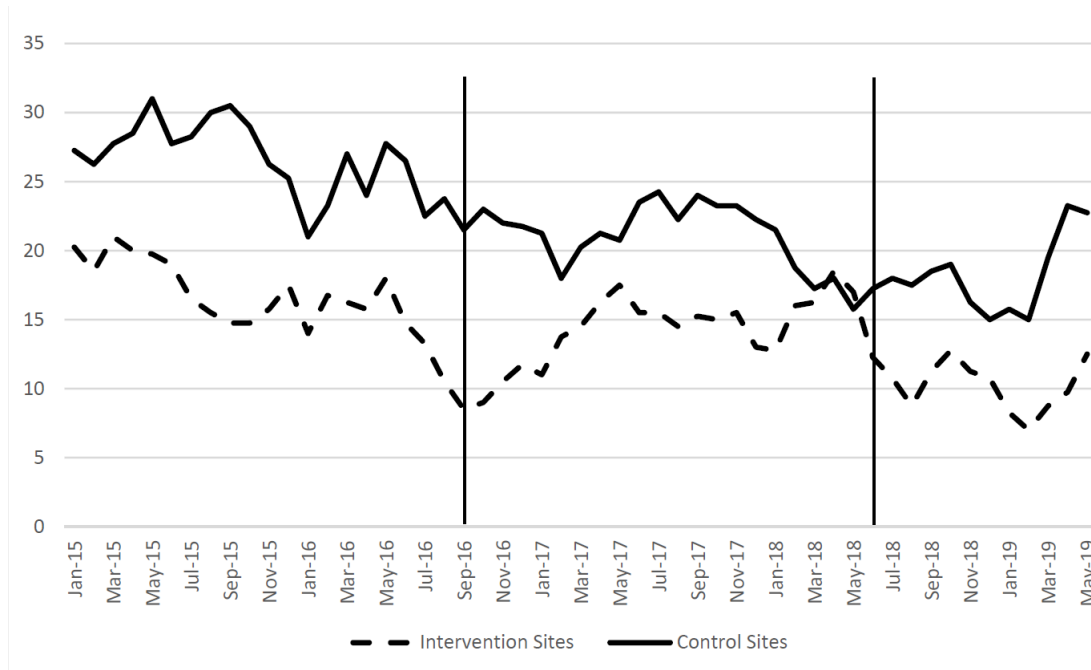
Implications

In New York City, the mayor, chancellor, and NYC DOE leadership team are looking for efficient mechanisms that will ensure the safety and dignity of all students and hasten the decline of crime in schools. This experimental evaluation focused on some of the "toughest" schools and aimed to provide empirical research evidence on the effects of SPS on middle schools characterized by high levels of violence that could lead to refinement and scale-up within NYC.

This carefully implemented, rigorous study showed that the SPS intervention was well-implemented in study schools, but failed to find significant effects of SPS on student outcomes. We can identify three potential causes for this failure. First, the contrast in safety activities

between SPS and control schools was not sharp. Many of the control schools engaged in the same activities SPS promotes, only to a lesser degree. The lack of treatment contrast may have contributed to the lack of effects.

Second, it may be the case that our primary outcome data, discipline incidents in schools, are simply very noisy and a far larger study would be necessary to see any potential effects. The fact that baseline disciplinary incident data were entirely uncorrelated with Year 2 disciplinary incident data suggests that there is notable flux across years. Statistical controls that rely on covariation, such as in our models, are not effective with such noisy data.

A third potential contributor to the failure to see effects is the slight drop-off in implementation quality in SPS schools from Year 1 to Year 2. Plots of raw data showed improvements in most outcomes of interest in SPS schools at the end of Year 1 (although these improvements were not statistically significant) but then in Year 2 these outcomes generally returned to levels similar to those before SPS began. This may reflect a lack of commitment on the part of school staff, or perhaps positive initial effects may have led school staff to "take their foot off the gas pedal" and relax implementation, with consequent adverse effects.

Given the failure to detect intervention effects, it is difficult to identify implications for policy. All schools must address the safety of their school environments; our data on control schools show that schools do accomplish this to a meaningful degree. The benefit of a program like SPS may be felt more powerfully in districts where effective practices are less common.

References

Braga, A. A., Papachristos, A. V., & Hureau, D. M. (2014). The effects of hot spots policing on crime: An updated systematic review and meta-analysis. *Justice Quarterly, 31*(4), 633–663. doi:10.1080/07418825.2012.673632

Cash, A. H., Debnam, K. J., Waasdorp, T. E., Wahl, M., & Bradshaw, C. P. (2019). Adult and student interactions in nonclassroom settings. *Journal of Educational Psychology, 111*(1), 104–117. doi:10.1037/edu0000275

Cash, A. H., Bradshaw, C. P., & Leaf, P. J. (2015). Observations of student behavior in nonclassroom settings: A multilevel examination of location, density, and school context. *Journal of Early Adolescence, 35*(5–6), 597–627. doi:10.1177/0272431614562835

Hong, J. S., & Eamon, M. K. (2012). Students' perceptions of unsafe schools: An ecological systems analysis. *Journal of Child and Family Studies, 21*(3), 428–438.

Johnson, S. D., Guerette, R. T., & Bowers, K. (2014). Crime displacement: What we know, what we don't know, and what it means for crime reduction. *Journal of Experimental Criminology, 10*(4), 549–571. doi:10.1007/s11292-014-9209-4

Levin, H. M., McEwan, P. J., Belfield, C. R., Bowden, A. B., & Shand, R. D. (2018). *Economic evaluation in education: Cost-effectiveness and benefit-cost analysis.* Thousand Oaks, CA: SAGE Publications.

Lipsey, M. W., Howell, J. C., Kelly, M. R., Chapman, G., & Carver, D. (2010). *Improving the effectiveness of juvenile justice programs: A new perspective on evidence-based practice.* Washington, DC: Georgetown University, Center for Juvenile Justice Reform.

Musu, L., Zhang, A., Wang, K., Zhang, J., & Oudekerk, B. A. (2019). *Indicators of school crime and safety: 2018* (NCES 2019-047/NCJ 252571). Washington, DC: Bureau of Justice Statistics and the National Center for Educational Statistics. Retrieved from https://www.bjs.gov/index.cfm?ty=pbdetail&iid=6526

Perkins, H. W., Perkins, J. M., & Craig, D. W. (2014). No safe haven: Locations of harassment and bullying victimization in middle schools. *Journal of School Health, 84*(12), 810–818. doi:10.1111/josh.12208

Vaillancourt, T., Brittain, H., Bennett, L., Arnocky, S., McDougall, P., . . . Cunningham, L. (2010). Places to avoid: Population-based study of student reports of unsafe and high bullying areas at school. *Canadian Journal of School Psychology, 25*(1), 40–54. doi:10.1177/0829573509358686

Wang, N., Liu, L., & Eck, J. E. (2014). Analyzing crime displacement with a simulation approach. *Environment and Planning B: Urban Analytics and City Science, 41*(2), 359–374. https://doi.org/10.1068/b37120

Weisburd, D., Telep, C. W., Hinkle, J. C., & Eck, J. E. (2010). Is problem-oriented policing effective in reducing crime and disorder? Findings from a Campbell systematic review. *Criminology and Public Policy, 9*(1), 139–172. doi:10.1111/j.1745-9133.2010.00617.x

Appendix. Descriptive Statistics for Outcome Variables

	Baseline 2015–16				Year 1 2016–17				Year 2 2017–18			
	Treatment		Control		Treatment		Control		Treatment		Control	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
School-level variables												
Percentage of incidents overall	40.99	22.44	29.50	16.52	39.76	29.87	28.71	17.10	42.27	23.81	31.99	19.61
Percentage of incidents in public spaces	18.51	11.95	11.89	6.65	20.37	13.87	13.92	7.02	21.59	13.13	15.72	8.92
Incidents resulting in suspensions overall	10.17	6.60	6.42	3.52	8.73	5.13	7.74	5.73	10.53	5.66	8.28	3.71
Incidents in public spaces resulting in suspensions	5.61	4.00	3.27	1.46	4.25	2.46	3.67	3.28	5.70	3.13	4.18	2.31
Level 4 and 5 incidents	17.15	9.93	13.23	9.26	13.67	6.89	12.89	8.29	14.04	5.73	13.35	4.33
Level 4 and 5 incidents in public spaces	5.02	3.81	2.84	1.54	3.30	1.88	3.22	3.10	4.26	2.19	3.50	2.25
Number of days suspended overall	8.70	5.59	13.05	7.97	8.45	4.69	9.97	3.51	7.44	5.03	8.76	2.82
Number of days suspended from public space incidents	10.75	7.84	16.68	11.29	8.97	6.44	7.86	4.36	7.81	5.36	9.41	4.32
Student-reported safety	2.95	.19	2.97	0.19	3.00	0.22	3.00	0.13	3.03	0.22	2.96	0.17
Student-reported student–teacher trust	3.09	.13	3.10	0.11	3.13	0.14	3.11	0.07	3.15	0.16	3.08	0.14
Student-reported classroom behavior	2.63	.13	2.69	0.18	2.67	0.10	2.69	0.11	2.71	0.19	2.66	0.17
Student-reported bullying	2.94	.19	2.95	0.21	2.62	0.20	2.62	0.15	2.66	0.23	2.60	0.18
Student-level variables												
Attendance	91.73	9.95	91.35	10.73	90.47	9.59	90.17	10.30	91.02	10.80	90.72	11.45
Mathematics achievement	0.03	0.93	-.13	0.93	0.08	0.95	-0.15	0.95	0.10	0.96	-0.10	0.92
English language arts achievement	0.07	0.92	-.05	0.88	0.07	0.92	-0.07	0.90	0.06	0.93	-0.07	0.93

Established in 1946, the American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research on important social issues and delivers technical assistance, both domestically and internationally, in the areas of education, health, and workforce productivity.

MAKING RESEARCH **RELEVANT**

AMERICAN INSTITUTES FOR RESEARCH

1000 Thomas Jefferson Street NW

Washington, DC 20007-3835 | 202.403.5000

www.air.org

9685_09/19