NCJRS National Criminal Justice Reference Service

Office of Justice Programs
BJA BJS NIJ OJJDP OVC SMART

The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title:	**Rainier Beach Campus Safety Continuum**
Author(s):	**Charlotte Gill, Ph.D., Madeline McPherson, MS, Xiaotian Zheng, MSS, Claudia Gross Shader, Ph.D.**
Document Number:	**308396**
Date Received:	**January 2024**
Award Number:	**2016-CK-BX-0005**

This resource has not been published by the U.S. Department of Justice. This resource is being made publicly available through the Office of Justice Programs' National Criminal Justice Reference Service.

Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Center for Evidence-Based Crime Policy
Department of Criminology, Law & Society

Rainier Beach
Campus Safety Continuum

Final Report

Charlotte Gill, PhD

**with Madeline McPherson, MS, Xiaotian Zheng, MSS,
& Claudia Gross Shader, PhD**

Report prepared for the National Institute of Justice

March 2023

Grant Number: 2016-CK-BX-0005

This resource was prepared by the author(s) using Federal funds provided by the U.S.
Department of Justice. Opinions or points of view expressed are those of the author(s) and do not
necessarily reflect the official position or policies of the U.S. Department of Justice.

The Center for Evidence-Based Crime Policy (CEBCP) in the Department of Criminology, Law and Society at George Mason University seeks to make scientific research a key component in decisions about crime and justice policies. The CEBCP carries out this mission by advancing rigorous studies in criminal justice and criminology through research-practice collaborations, and proactively serving as an informational and translational link to practitioners and the policy community. Learn more about our work at https://cebcp.org and about the Department of Criminology, Law and Society at https://cls.gmu.edu.

Charlotte Gill is an Associate Professor in the Department of Criminology, Law and Society and Deputy Director of the Center for Evidence-Based Crime Policy at George Mason University. Madeline McPherson is a doctoral candidate in the Department of Criminology, Law and Society and a graduate research assistant in the Center for Evidence-Based Crime Policy at George Mason University. Xiaotian (Timo) Zheng is a doctoral student in the Department of Criminology, Law and Society and a graduate research assistant in the Center for Evidence-Based Crime Policy at George Mason University. Claudia Gross Shader is the Research and Evaluation Director at the City of Seattle Office of City Auditor.

This evaluation was funded by the National Institute of Justice, Grant # 2016-CK-BX-0005. The opinions, findings, and conclusions or recommendations expressed in this report are those of the authors and do not necessarily reflect those of the U.S. Department of Justice, National Institute of Justice, or the City of Seattle.

Acknowledgments

The authors would like to thank the following for their hard work, contributions, and support of this research, as well as all current and former members of the Rainier Beach community, Seattle Public Schools, and the research team:

Barb Biondo
Stewart Bowerman
Mark Bridge
Brian Callanan
Clifford Cawthon
Brian Chu
Rebecca Chu
Gregory Davis
Devon DeLena
Lakia Faison
Jenny Frankl
Cathy Girouard
Tricia Hagerty
Danielle Jackson
Marty Jackson
Rachel Jensen
Ricole Jones
L. Cait Kanewske
Naida Kuruvilla
David Lewis
Lori Lynass
Christine Mattfeld
Tolu Mejolagbe
Hana Mohamed
Denise Nazaire, PhD
IB Osuntoki
Sangjun Park, PhD
Heather Prince, PhD
Mahogany Purpose
Patricia Sander
Marisol Santos-Perez
Tanisha Scott
Thomasina Schmitt
Kamaria Sims
Linda Spain
Jabali Stewart
Heidi Tuttle
Zoe Vitter
Cathie Wilmore
David B. Wilson

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Contents

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

List of Tables

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

List of Figures

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

1 Purpose of the Project

The purpose of this project was to develop and evaluate the *Rainier Beach Campus Safety Continuum* (RBCSC), a community-led, place-based, evidence-informed approach to addressing school and community safety and reducing racial disparity in school discipline and police contact through non-punitive approaches in the Rainier Beach neighborhood of Seattle, Washington. The project built upon two existing community crime prevention initiatives developed by the partners of this grant, which included grassroots community organizations such as the Rainier Beach Action Coalition, Seattle Neighborhood Group, and the Boys and Girls Club of King County Rainier Vista Boys and Girls Club; the City of Seattle; and the Center for Evidence-Based Crime Policy at George Mason University (CEBCP): *Rainier Beach: A Beautiful Safe Place for Youth* (ABSPY), a Bureau of Justice Assistance-funded community-led, data-driven, place-based approach to youth crime prevention (Gill & Gross Shader, 2020; Gill et al., 2016), and *Rainier Beach: Beautiful!*, an Office of Juvenile Justice and Delinquency Prevention-funded application of Positive Behavioral Interventions and Supports (PBIS) that extends from local schools into setting shared norms and expectations in community facilities and businesses.

The RBCSC was envisaged as an integrated school-community prevention approach that combined evidence-informed educational practices and community-based crime prevention strategies to provide a continuum of support, in which both young people and adults were encouraged to participate in a positive culture both inside and outside of school. The goal of the approach was to improve school climate and perceptions of community safety, and to reduce crime and punitive discipline, both in school and in the community (i.e. police response to crime involving young people). We aimed to create positive behavioral change by focusing not on individual young people, but rather on the places where crime and victimization involving young people is most concentrated, also recognizing the relevance of the adult-run systems and institutions that shape youthful behavior at those places.

The development and design of the RBCSC weaves together several conceptual threads from criminology and education research, presenting an innovative combination of a number of evidence-informed strategies within a strong place-based approach. Criminological research has well-established that juvenile crime is tightly coupled to place (Weisburd, 2015; Weisburd et al., 2009), but beyond inclusion of the geographic location of school buildings, there is scant research integrating schools into the community landscape when it comes to addressing young people's behavior (Gottfredson et al., 2001; Roman,

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

2002, 2005). However, young people spend the majority of their waking hours at school, which plays a role in how they move and interact within their communities even outside of school hours. One aim of the RBCSC was to address this gap in the literature and to influence school-based officials and community members to think of the needs of local young people comprehensively, involving both schools and community institutions.

1.1 Review of relevant literature

1.1.1 Juvenile crime at place

Crime and place are inextricably linked (see Weisburd, 2015, for a summary of research). Weisburd et al. (2009) found in a study of crime hot spots in Seattle that just 86 street segments of over 26,000 in the city accounted for one-third of all citywide juvenile arrests over a 14-year period. Why is juvenile crime so tightly coupled to place? Environmental theories of crime suggest that people's daily "routine activities," which are determined by the types of places they frequent, shape opportunities for crime and affect people's decisions to commit crime or their risk of victimization (e.g. Cohen & Felson, 1979; Felson, 1987; Felson & Boba, 2010). While these theories apply regardless of age, they suggest that young people—due to the specific types of activities they are permitted to and do engage in—influence the types of places in which they are likely to spend time, and thus commit crime (Felson, 2006).

For example, juveniles are required to be in school at certain times, and juvenile offending tends to cluster around schools during the times when students are either arriving at or leaving school (Gottfredson et al., 2001; Roman, 2002, 2005). Young people's routes of travel to and from school tend to be more predictable and specifically defined than the commutes of adults and can produce opportunities for crime when offenders and potential victims cross paths (Brantingham & Brantingham, 1995). Juveniles, to a much greater extent than adults, are also subject to restrictions on when and where they congregate—for example, they cannot go into bars, occasionally need a parent or guardian to enter a space or establishment (e.g., gyms, pools, etc.), and in some jurisdictions are even subject to curfew at certain times and places.

These factors also determine patterns of juvenile offending at particular places and times. Weisburd et al. (2009) found that the Seattle street segments with the highest concentration of juvenile arrests were the types of places where juveniles tend to hang out without supervision or structured activities—

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

shopping malls, outside schools, and in public spaces—while places with a low concentration of juvenile crime were non-public spaces and adult-oriented locations. These findings also fit with research by Osgood et al. (1996) indicating that unstructured, unsupervised socializing is linked to higher rates of juvenile delinquency and violence.

While these findings are consistent in relation to young people enjoying their free time, an analysis of behavior at school tends to look different both in the research and in practice. At school, juveniles are restricted in their movement, activities, and expectations. School officials manage behavior and where a young person acts out or violates a rule, he or she is met with a number of possible responses to quell or modify the undesirable behavior. In fact, in-school behavior is considered so much differently than community behavior, school officials are unlikely to refer to poor in-school behavior as "delinquency" or "crime" unless it is serious, egregious, or otherwise violates standard norms of behavior. For the purposes of well-designed research, however, school is an important place in which to assess the behavior and actions of young people, thus building upon the crime at place literature. With the rise of school resource officers in recent decades, school and community resources are increasingly paired together to provide safe and secure school campuses and communities. Yet, scholarly work continues to silo school-based and community-based interventions as if they target different populations of young people.

1.1.2 The role of school climate in juvenile crime at place

Notably, the climate within a school or school district has been found to influence a number of factors for students and their families as well as school staff. Defined as the norms, values, and expectations that support safety, relationships, teaching, and learning within schools, school climate—and students' perceptions of it—influences a range of outcomes including students' emotional and mental health, substance use, absenteeism, discipline, academic success, aggression, violence, and positive development (see Thapa et al., 2013, for an extensive literature review of school climate research).

Along with student characteristics, school climate is an important predictor of various forms of school disorder (e.g. Gottfredson et al., 2005; Welsh, 2000, 2001). Findings from a meta-analysis conducted by Reaves et al. (2018) found that addressing or improving school climate is one way to meaningfully reduce in-school problem behavior by students that lasts over time. School climate also impacts academic performance and is associated with risk factors for violence later in life (e.g. Hawkins et al., 1998; Herrenkohl et al., 2000). Despite the significant influence of schools in the lives of most juveniles, there is little research specifically connecting school climate with juvenile offending in the areas surrounding the school, or on optimal strategies for linking school- and community-based interventions. At the same time, the community and place-based context in which the school operates may impact both the behavior and development of students in the school and the school climate itself in a cyclical process (e.g. Duncan & Raudenbush, 1999; Kubrin & Weitzer, 2003). School and community contexts are therefore closely intertwined. Since young people spend so much of their time in school, it is likely that the culture, climate, and behavioral norms that prevail on campus may spill over into how they behave in the community, and vice versa (e.g. Anderson, 1982).

As such, school climate interacts with elements of neighborhood context, such as community disadvantage and perceptions of safety (Anderson, 1982; Duncan & Raudenbush, 1999; Gottfredson & Gottfredson, 1985; Kitsantas et al., 2004; Kubrin & Weitzer, 2003; Limbos & Casteel, 2008; Welsh et al., 1999). Several studies have examined how the broader neighborhood context, including crime rates and indicators of social disorganization (e.g. Shaw & McKay, 1942), and school climate interact to influence young people's perceptions and behavior both inside and outside of school. Kitsantas et al. (2004) found that the strongest predictors of student perceptions of safety in their schools were school climate, community safety, and students' perceived safety of the school relative to the neighborhood. Gottfredson and Gottfredson (1985) found that neighborhood disadvantage predicts violence and crime in schools (see also Limbos & Casteel, 2008; Welsh et al., 1999). Crime and unsafe conditions in the neighborhood and at school can negatively impact school climate and, in turn, increase the risk of academic failure (e.g. Bowen & Bowen, 1999; Henrich et al., 2004; Margolin & Gordis, 2000; McCoy et al., 2013). Ruiz et al. (2018) studied the relationship between socioeconomic status and academic achievement in elementary school-aged students. They found that low socioeconomic status was correlated with lower academic achievement, and violence in the community mediated this relationship. On the other hand, the authors found that a positive school climate was positively associated with academic achievement, which is consistent with prior research that a quality school climate can be especially beneficial for high-risk students in socially disorganized neighborhoods (Battistich et al., 1995). These studies demonstrate the reciprocal, unbounded nature of young peoples' experiences at school and in the community, showing that each influences the other.

These studies suggest that it is difficult to separate the school and community context when consid-

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

ering how to respond to negative in-school behaviors that may ultimately translate to negative behaviors in the community (e.g., juvenile crime) in places close to the school. Fischer and Argyle (2018) found in one rural jurisdiction that school districts that adopted a four-day school week witnessed a roughly 20% increase in juvenile crime, suggesting that schools play a role in limiting juvenile crime in the greater community by keeping young people occupied. Another study found that improving disorganization within a school, defined as improvements to the physical and social environment, can lead to a reduction in violence participation among students (Lindstrom Johnson et al., 2017). Importantly, however, the findings show that students' perception of an improving school environment is key to measurable violence reduction.

Gottfredson (1986) and Sheldon and Epstein (2002) have found that involving the family and community in school activities and planning for school safety is beneficial for school climate and may reduce disciplinary referrals. Furthermore, both school- and community-based research suggests that punitive responses to behavioral issues (such as suspensions and expulsions from school, or arrest in the community) are counterproductive for young people and disproportionately affect youth of color (Brunson & Miller, 2006; Fabelo et al., 2011; Kirk, 2008; Losen, 2014; Mowen & Brent, 2016; Payne & Welch, 2010; Petrosino et al., 2010; Skiba et al., 2002). However, even when schools employ interventions to attempt to limit "exclusion" among students (e.g., time outside the classroom, suspension, expulsion, etc.), any reduction is generally short-term with exclusion rates dropping for the first six months after initial intervention (Valdebenito et al., 2018). Generally, there is much more room for innovation in thinking about how school and community interventions can be integrated to improve climate, behavior, and safety in both settings, which is the impetus behind the RBCSC initiative.

1.1.3 Developing an integrated prevention approach

Theory and research on how youth respond to different types of interventions can be instructive in thinking about how to develop an integrated school-community prevention program. In particular, school discipline and juvenile crime are analogous in that punitive responses to both can be counter-productive (Brunson & Miller, 2006; Fabelo et al., 2011; Kirk, 2008). Suspensions and expulsions from school are significantly associated with both involvement in the juvenile justice system and negative education-related outcomes indirectly associated with crime, such as poor academic attainment and dropout (e.g. Fabelo et

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

al., 2011; Losen, 2014). Mowen and Brent (2016) describe school discipline as a ”negative turning point” for juveniles that can set them on a trajectory toward future delinquency and arrest; they find that multiple suspensions have a cumulative effect on this risk. Outside the school, while hot spots policing is overall effective in reducing crime at place (e.g. Braga et al., 2014), police enforcement at juvenile crime places may contribute to increased justice system processing for youth, which has been found to increase the likelihood of recidivism (e.g. Petrosino et al., 2010). Furthermore, in both the school and community contexts, research suggests that traditional enforcement through school discipline or arrest disproportionally affects youth of color (Brunson & Miller, 2006; Fabelo et al., 2011; Kirk, 2008; Losen, 2014; Payne & Welch, 2010; Skiba et al., 2002).

Thus, in developing an integrated prevention program it is important to look for approaches that seek to change behavior proactively by fostering a healthy, supportive climate in which all community members are encouraged to comply with established social norms and where those who struggle to do so receive supportive assistance. In contrast, approaches that focus on punishing disruptive behavior by individual students have been shown to have only a short-term impact on changing behavior or school climate for the better (Valdebenito et al., 2018). Two school-based approaches that follow this philosophy are Positive Behavioral Interventions and Supports (PBIS) and school-based restorative justice (RJ). Both seek to improve school climate and safety through a whole-school approach emphasizing positive behavioral change, communication of expectations, and equity among all school stakeholders (including adults), along with individualized, non-punitive support for higher-risk students (e.g. Bradshaw, 2013; Fronius et al., 2016).

PBIS is an evidence-based framework for improving students' social and academic outcomes and engagement with school, responding to problematic behavior at different levels, and integrating other evidence-based approaches to violence prevention (Bradshaw, 2013; Horner & Sugai, 2015; Horner et al., 2010; Sugai & Horner, 2006). According to Horner and Sugai (2015), PBIS has been implemented in 21,000 schools in the United States over 20 years. PBIS aims to improve overall school culture while also adding ”tiers” of more intensive support for individual or groups of students with higher levels of need. The three-tier system that characterizes most implementations of PBIS corresponds to the three levels of prevention (primary, secondary, and tertiary). Tier 1 (primary prevention) interventions implemented through the PBIS framework are proactive and focus on preventing problem behaviors from occurring at all by setting common standards and expectations for behavior within different areas of the school

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

(e.g., the classroom, lunch room, play areas). It is estimated that approximately 80 percent of students in a PBIS school would only experience Tier 1 programming (see Figure A1). Tier 2 (secondary prevention) adds more intensive support for the estimated 10-15 percent of students who exhibit problem behaviors and need additional structure and support in order to follow Tier 1 expectations. Finally, Tier 3 (tertiary prevention) focuses on individual support plans for specific students who have been suspended or are otherwise unable to remain in the mainstream classroom, in an effort to ultimately include all students in positive school-based activities. These students are estimated to comprise about 1-5 percent of the student body (Horner & Sugai, 2015; Swain-Bradway et al., 2016). Research on PBIS in schools show that it is effective in reducing disciplinary referrals, out-of-school suspensions, and expulsions and improving academic outcomes and school climate, including perceptions of safety (e.g. Bradshaw et al., 2010; Horner et al., 2009, 2015).

RJ is defined as "a process whereby all the parties with a stake in a particular offense come together to resolve collectively how to deal with the aftermath of the offense and its implications for the future" (Marshall, 1997, cited in Strang, 2002, p. 44). Through the RJ process, which can take a number of forms including face-to-face and family group conferences, victim-offender mediation, and peacemaking circles (Latimer et al., 2001; Presser & Van Voorhis, 2002), community representatives come together to condemn the wrongful act but find ways to repair the harm rather than punish the offender. This problem-solving process supports the offender's reintegration into the community and promotes the dignity of all parties by giving everyone involved a voice, in line with theories of reintegrative shaming (Braithwaite, 1989) and procedural justice (Tyler, 1990). Studies have found that RJ practices are generally effective in reducing recidivism at various stages of the criminal justice system (see Sherman et al., 2015). In the school context, Fronius et al. (2016) suggest that RJ practices are increasingly appealing to school districts as awareness grows about the relationship between punitive school discipline and negative outcomes, and the racial and ethnic disparities in the application of school discipline. There is some research suggesting that school-based RJ reduces disciplinary referrals in general and addresses racial disparities to an extent, as well as increasing students' perceptions that teachers treat them with respect (Gregory et al., 2016). Other studies have found that, when RJ is embedded in the school culture, it can improve school climate, student connectedness, and academic achievement (see Fronius et al., 2016, for a review). However, the evidence base for school-based RJ remains limited and there are currently no rigorous empirical evaluations (Fronius et al., 2016).

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

The RBCSC combines PBIS and RJ in two innovative ways. First, the alignment of PBIS and RJ in the school setting has been encouraged by education researchers but is still a developing practice (Swain-Bradway et al., 2016). However, the key innovation of the RBCSC is the extension of these approaches into the surrounding neighborhood to shift the culture and quality of interactions among adults and youth not only in the school environment but also in local businesses and institutions such as libraries, community centers, and policing. We proposed that PBIS-RJ could leverage community collective efficacy and informal social control to create a continuum of positive support within and outside of school, focusing primarily on changing the way these adult-run institutions operate and regulate behavior in the interests of universal safety. Figure A1 lays out our vision for integrating PBIS and RJ in both the schools and communities, following a proposal by Swain-Bradway et al. (2016) for school-based integration. The implementation of the project in practice is described in detail in the next section.

2 Methodology

Our original evaluation plan examined the effects of integrated school-wide (SW) and community-wide (CW) PBIS and RJ on school climate, school outcomes such as academic achievement and disciplinary referrals, neighborhood crime, and community perceptions of safety. We also sought to examine whether SW- and CW-PBIS-RJ reduce racial and ethnic disparities in these outcomes.

We chose the Rainier Beach neighborhood for our study site because it encapsulates the complex relationship between school climate, community context, and juvenile crime. Rainier Beach is a diverse community of around 5,000 people, with at least 167 primary languages spoken in the zip code area in which it is located (Gill & Gross Shader, 2020). The neighborhood has historically been challenged by violence and community concerns about public safety, with high levels of juvenile- and youth-involved violent crime. Its crime hot spots are entrenched, appearing in Weisburd and colleagues' work on juvenile and overall hot spots going as far back as 1989 (Weisburd et al., 2004, 2009, 2012). As a result, the neighborhood has missed out on economic and development opportunities. However, it also has a strong history of community organizing, as evidenced in part by the two federally-funded initiatives that formed the basis for this project. Thus, this work is grounded in an existing culture of putting data in the hands of the community, building capacity, and implementing evidence-informed strategies.

We envisaged that the project would proceed in three phases, starting at the beginning of the grant period in January 2017: a 12-month planning phase, in which the project partners would create an implementation plan for the combined PBIS-RJ approach imagined in Figure A1; an 18-month training and implementation phase; and an 18-month stabilization and sustainability phase. In reality, the planning phase in particular took much longer as we sought to develop a completely innovative approach with little guidance aside from adapting school-based PBIS protocols and restorative practices for various community settings. We also ran into other challenges; for example, delays in the various partner organizations being able to accept the grant funding and disagreement about which community organizations should be at the table, which led to the project partners participating in our own restorative circle at the end of 2017 to rebuild relationships and trust. Although the school district began training on PBIS in many schools (not just in Rainier Beach) earlier, the full implementation planning process was not completed until late 2018. As a result, in our analysis we consider January 2019 the first month of full school- and community-wide implementation. Seattle was also one of the first places in the United States to be affected by the COVID-19 pandemic, beginning in late February 2020, which further affected the program, although some implementation was able to continue in a modified form.

Our original evaluation plan included analysis of school administrative data, school-based surveys and observations, a three-wave in-person community survey in Rainier Beach and a comparison neighborhood, and process-related data. Due to the delays at the start of the project, and the serious and ongoing impact of the COVID-19 pandemic on schools and communities alike, our ability to fully implement the RBCSC and collect evaluation data has been limited. We were unable to implement the second and third waves of the community survey as planned, but we conducted a mail survey of the same addresses sampled for the baseline survey, as we describe below. Seattle Public Schools (SPS) provided administrative and survey data on disciplinary referrals, academic achievement, and student perceptions of school climate. However, it is important to note that these data are also significantly impacted by the pandemic and related school closures (for example, disciplinary action has been effectively suspended since the beginning of the pandemic). We received calls for service and crime incident data from Seattle Police Department (SPD) consistently throughout the project under a pre-existing data agreement. All primary and secondary data collection protocols were reviewed and approved by the George Mason University Institutional Review Board, and we also obtained research approval and a data agreement from SPS.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

2.1 Implementation of the program

The multidisciplinary RBCSC team kept substantial records of project activities throughout the grant period. Here we use these records, gathered from biannual reporting by each grant partner gathered by the City of Seattle Office of City Auditor (OCA), which oversaw the distribution of funding to each of the local partners, and Tiered Fidelity Inventory (TFI) reports completed by SPS to track implementation of the SW-PBIS work. The biannual reports contained a variety of process and implementation data, including notes from implementation team meetings and community town halls.

The grant period began in January 2017, although in practice we did not receive final budget approval from NIJ until April and the City of Seattle was unable to accept the funds until June, as the city did not have a process in place to accept a federal research grant. Initial project planning involved the ABSPY Core Team, which continued to meet throughout the life of the project (ABSPY activities predate this project and continued throughout the grant period). The planning process involved challenges from the outset. There was overlap in ABSPY and RBCSC personnel, and some of the community partners on the ABSPY Core Team felt they were not appropriately consulted during the development of the proposal. These tensions persisted throughout the life of the grant, and between the overall Core Team and its designated subcommittee, the "NIJ Workgroup."

The organization Huayruro, hired as a technical assistance consultant for the CW-RJ aspect of the project, stepped in to facilitate peace circles for the project partners themselves in the first year of the project, and these circles continued throughout the grant period. Huayruro meeting notes through 2020 document the ongoing tension among the project partners and the key areas in which healing and relationship-building were needed. These included ensuring all team members had a voice in the work and were treated equitably. The documentation suggests the ABSPY Core Team was on a stop-and-go trajectory, as members found it difficult to discuss project activities when there were more considerable personnel and relationship issues among the relevant partners. Importantly for others embarking on similar work in which grassroots community organizations are partnered with government entities like the police, school district, and other city services, some of these tensions were driven by community members' perceptions and experiences of systemic and institutional racism at the hands of government. These tensions needed to be worked through and agreements reached about where the balance of power lay at the table before progress could be made on implementation.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Despite these ongoing challenges, the team was able to move the project forward. The latter part of 2017 was largely dedicated to building the structure of the project and establishing contracts between OCA and the other community agencies. These contracts included Sound Supports, which is the technical assistance provider overseeing PBIS activities, particularly in the Seattle Public Schools. In the summer of 2017, Seattle Channel, the City of Seattle's cable channel, filmed and produced an 8-minute video in 2017 that highlighted ABSPY's work in the community and features interviews with a number of community members serving in various roles. In 2018, the NIJ Workgroup began meeting weekly to dig into the focal pieces of program activities—PBIS and RJ—both as a philosophy and an evidence-based set of practices. In some ways, the two areas seemed to have opposite trajectories for implementation: PBIS was developed primarily for school-based settings and needed to be adapted to the community, whereas RJ was developed primarily for community settings and needed to be adapted to the schools. The natural progression of NIJ grant activities in 2018 put the NIJ Workgroup in a place to begin developing implementation plans for both SW- and CW-PBIS and RJ, respectively.

The combination of SW-PBIS and SW-RJ, and the extension of these two processes to the community, are innovative. While RBCSC is the first program that includes the community extension, other schools have explored the combination of the two practices in the educational setting. In the spring of 2018 the NIJ Workgroup embarked on two site visits to learn more about how other schools and organizations have accomplished some of the same goals. Members of the team visited Federal Way, WA School District (30 minutes south of Rainier Beach) on March 13 and Chicago Public Schools from May 28-30. Both trips focused on learning how the two districts had implemented PBIS and RJ in their schools. Also during this year, restorative justice case management staff were either identified or hired at three of the treatment schools to help facilitate some of the SW-RJ activities. These team members were hired as social workers, counselors, or youth service assistants.

Rainier Beach Action Coalition (RBAC), the lead partner for CW-PBIS, held a community town hall in Rainier Beach in July 2018 to explain what PBIS is and the vision for integrating it into community settings. In the same vein, the Boys and Girls Club of King County (BGCKC), the CW-RJ lead agency, held a Peacemaking Circle Retreat in September 2018 for some of the community partners to detail how Peace Circles work and facilitate healing in the community. By December of 2018, RBAC had developed a CW-PBIS implementation plan, which was approved by the ABSPY Core Team. The plan was aligned with SW-PBIS in that it offered Tier I, II, and III supports.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

In 2019, the NIJ Workgroup continued to meet weekly until June and then switched to a biweekly structure for the latter half of the year. Contracts were updated as necessary between OCA and the various community organizations. In May, members of the team attended the Coalition of Schools Educating Boys of Color (COSEBOC) conference in Detroit to learn about how to implement project activities in a culturally-appropriate way for the highest-risk groups of students in SPS. Seattle Channel developed four more short videos to document project activities:

- Great Expectations for All: Implementing PBIS in Rainier Beach Schools
- United Campus: Building a Safe, Respectful and Responsible Community
- Universal Positive Environments
- Peace Circles in Rainier Beach

The schools implemented PBIS-RJ, mainly at Tier I, during 2019, but the momentum was somewhat slower compared to the trainings, staff hiring, and initial implementation that was occurring during 2018. No additional RJ staff were hired at the treatment schools this year. However, three of the schools (ES4, MS5, and HS5)[1] began holding restorative circles this year, and ES3 held a leadership retreat for its 5th-graders that focused on restorative practices.

In the community in 2019, RBAC and the Seattle Neighborhood Group (SNG) developed more campaigns to communicate PBIS in the community and begin establishing the "shared norms" that form the basis of Tier 1 activities. Their Social Norms Campaign used visuals like yard signs, t-shirts, brochures, and banners to share the "Be^3" messaging—Be Respectful, Be Responsible, Be Safe—that was developed by the community as part of the *Rainier Beach: Beautiful!* initiative. The team also began developing a CW-PBIS Tiered Fidelity Inventory, modeled after the PBIS implementation tracking form used by staff. CW-RJ activities included another Peacemaking Circle workshop held in October 2019, which had 14 participants from the community partner organizations. Four "Circle Keepers" were hired to build capacity to hold community Peace Circles. The Circle Keeper position was especially important as it provided meaningful employment for young people from the Rainier Beach community.

In 2020, much changed for the NIJ Workgroup and project. Barely two months into the new year, COVID-19 reached the United States and began spreading in the community. The American "Patient

[1] See the following section for our school naming convention, used in this report to protect the confidentiality of students within the project schools.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Zero" was identified in King County, where our project was located, in mid- to late-February. By the second week of March, closures and cancellations began to happen throughout Seattle and the nation. Monday, March 16 was the first day many non-essential employees began to work from home. Two NIJ Workgroup meetings were canceled during this hectic time, which led to a roughly six-week break between meetings. OCA's NIJ Coordinator also left the team during this time to pursue a new opportunity working in public health, and a new coordinator had to be hired. During 2020 the OCA team also produced an information packet on Peacemaking Circles for undergraduates at Seattle University and conducted an internal training for OCA staff on Behavior-Specific Praise. Members of the NIJ Workgroup also attended a virtual PBIS Leadership Forum.

All grant activities were affected by the COVID-19 pandemic. OCA and CEBCP employees began teleworking, as did any non-essential staff at the Rainier Beach community organizations. At first, SPS planned to be closed for two weeks, but that was quickly extended to June due to a stay-at-home order issued by Governor Jay Inslee. Given that King County was a COVID-19 hotspot early on, the SPS community was directly affected by COVID-19. Healthy teachers (and the district) were prioritizing covering the classes of ill teachers, which created a distraction from project activities, at best, and led to their deprioritization, at worst. Nonetheless, virtual TFIs were conducted for ES4, ES5, and MS5 during 2020, and ES4 and ES5 developed a PBIS handbook and resources to guide implementation. They were the only schools to complete this deliverable.

Adding to the challenges of 2020, George Floyd was killed by police in Minneapolis on Memorial Day, and racial justice protests broke out across the globe throughout June and July shortly after. For months afterward, individuals, corporations, and public agencies pledged to "correct" systemic racism and a nationwide discussion of "anti-racism" (Kendi, 2019) took center stage. Though never explicitly stated as a result of the 2020 racial justice movement, around this time SPS decided to revise the Seattle TFI. Recognizing that the existing TFI tool was not imbued with anti-racism, the SPS team spent months revising measures and procedures. Their goal was to improve equity in school achievement, discipline, etc. across Seattle Public Schools. During this revision process, no TFIs were completed. The ABSPY Core Team also began reflecting on and reconsidering its relationship with the Seattle Police Department (SPD), who were represented on the team until this point but ended up without a representative, in part because the department's long-term representative was called away to work with the new Chief, who came in after SPD's former Chief Carmen Best resigned.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

In the community, SNG and RBAC experienced some significant turnover in high-level positions. SNG lost their Executive Director and Program Manager. RBAC's PBIS Program Coordinator had gone on maternity leave in November 2019 and decided not to return to work in 2020. As a result of the turnover, SNG developed a "Lookback" report of the broader ABSPY initiative as an orientation guide for future new staff. RBAC hired a new Program Specialist and Program Coordinator by June. Despite the challenges, the community-wide PBIS team developed a PBIS matrix based on the Be^3 model that showed how each of the three community values (Be Respectful, Be Responsible, Be Safe) could be implemented in different settings. RBAC also conducted CPTED assessments and walkthroughs with community organizations and businesses by request.

CW-RJ activities continued, but feasibility continued to be an issue since Peace Circles are traditionally held in person. When the pandemic began, BGCKC continued to hold in-person activities longer than the other organizations, as some of their programs double as childcare for working parents. They eventually transitioned Peace Circles to a virtual format on March 31 and kept richly detailed notes on how how this transition happened and how youth participants were faring in general. By year-end, approximately 4-5 Circles were taking place each week. BGCKC also began a podcast toward the end of 2020 called the "Youth Insider Podcast." Importantly, the community Peace Circles helped residents deal with the trauma of the pandemic, the racial justice reckoning, and a number of high-profile shootings and homicides that occurred in the Rainier Beach community in 2020 and 2021. The community connections that arose from ABSPY and RBCSC enabled community organizations and residents to mobilize from the earliest days of the pandemic to support vulnerable and isolated community members and ensure needed resources like food, medication, and vaccinations were provided to Rainier Beach.

The grant was originally due to end in December 2020. However, having essentially lost a year at the beginning of the project due to the delays in receiving and accepting funding and the relationship challenges on the team, and almost another year due to COVID-19—during which we were barely able to spend down grant funding—we requested a one-year extension to the grant. Unfortunately, our request came at the same time as the Office of Justice Programs (OJP) was implementing the new JustGrants system. Our grant record was lost in the transition to the new system and our request could not be processed. After two months there was still no resolution to the issue and the grant technically expired at the end of 2020. This was a significant setback for our team. Several staff members hired on the grant had to be laid off. This further exacerbated the tensions on the team, as grant partners who worked for government and educational institutions did not need to rely on the money to pay their salaries but many community organization partners did. Those who did lose their jobs were overwhelmingly young people of color from the Rainier Beach neighborhood. Ultimately George Mason University stepped in to guarantee some of the community funding through March 2021. The JustGrants issue was resolved by March.

We immediately developed and submitted a second extension request through the end of 2022, knowing that we were already two years behind at this point. This request was not approved until late 2021. Grant activities did take place in 2021 but at a much slower pace, as team members were uncertain about what the future of the program looked like. Due to the grant extension, all of the grant contracts and MOAs had to be amended to adjust the periods of performance. We also lost another OCA project coordinator during this year.

Within SPS, the new racial equity-based TFI tool was still being finalized. The project team members from SPS held various orientations on the new tool in late 2021. By this point, ES5 was pulling ahead of the other schools in terms of implementation, and had begun developing Tier II materials and adapting their Distance Learning Matrix to in-person school, as students began returning in the 2021-22 school year.

In the community, SNG and RBAC continued to conduct community assessments with organizations and businesses who requested them. To continue to spread the word about CW-PBIS, a coloring book was developed that shares PBIS messaging. Over 800 coloring books were handed out in August 2021 at the Back2School Bash. Two murals were commissioned to be painted in the community that incorporated the "Be^3" messages: the Be'er Sheva mural and the Pho Van mural. RBAC held a Public Safety Deep Dive event in October that was intended to share information about the project with community members, and toward the end of the year, began storyboarding four short, TikTok-style videos to share on social media and educate people about PBIS.

In terms of CW-RJ, BGCKC continued to document the regular Circle Keeping activities and offer rich information about timing, location, and participants. BGCKC, OCA, and CEBCP collaborated to create a curriculum for Circle Keeper training, which was completed by the end of 2021 in the form of a "Circle Keeping Handbook." The Youth Insider podcast continued and BGCKC partnered directly with Rainier Beach High School to provide alternatives to suspension in the form of restorative practices, the goal of which was to help mediate conflict while providing social/emotional support to students. Finally, BGCKC

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

began to develop plans to hold Parent Circles, which continued into 2022.

In 2022, the NIJ Workgroup continued to meet on a biweekly basis. OCA brought on a new NIJ Coordinator, who had previously worked as a youth staff member with BGCKC and had a strong understanding of young people in the Rainier Beach community. In April, a number of Workgroup members attended and presented at the Northwest PBIS Conference in Seattle. OCA connected with Seattle Channel to produce a culminating video describing the implementation of the RBCSC. Filming for the video took place between May and early August, and the final product was available by the fall.

On the CW-PBIS front, RBAC held more Public Safety Deep Dives in late May and September 2022. They continued to conduct CPTED assessments of community organizations and businesses by request. RBAC had a booth at the Block Party held in June by BGCKC and other community organizations. Similarly, and as in past years, RBAC attended the Back2School Bash in August where they handed out PBIS coloring books among other PBIS-related items. For CW-RJ, BGCKC continued to hold Circles every week and has steadily been increasing capacity to cover additional locations. They have also continued to develop a concept for Parent Circles and worked to increase the number of Youth Circle Keepers. Both CW-PBIS and CW-RJ were sustained with funding from other local sources after implementation funding from the NIJ grant ended in December 2022.

In the schools, the new TFI was distributed to the NIJ Workgroup in the fall of 2022 and is getting closer to being rolled out, but had not yet been used by the end of the project. In general, the schools have struggled to bounce back to their previous implementation operations as they continue to deal with the fallout of the pandemic and subsequent teacher and staff shortages. SPS's goal for 2022 was to finalize PBIS resources "with renewed focus on students' voices and engaging with families," but only ES5 (and to some extent ES4) truly reached this stage, based on the documentation from OCA's biannual reports.

2.2 Project design

As discussed above, we chose the Rainier Beach neighborhood and its schools as the implementation site for this program in order to build on existing initiatives. Due to the limited number of schools in the neighborhood and the proximity of the schools to one another, which we felt would affect our ability to assess effects on crime and other outcomes, we did not conduct a randomized controlled trial. How-

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

ever, we selected two comparison areas elsewhere in Seattle that had similar types of schools and/or demographic profiles and crime rates. The nature of schools in the area primarily drove the selection; for example, we selected one neighborhood in a different part of the city that featured a Kindergarten through 8th grade (K-8) elementary/middle school and a high school in close proximity to each other, mirroring Rainier Beach's K-8 and high schools. The other area we selected for comparison was larger, but fell within the same police precinct as Rainier Beach and featured a number of elementary and middle schools, a high school, and an alternative high school, since Rainier Beach also has an alternative school. Nonetheless, it is important to note that these comparison areas are not randomly assigned or statistically matched with Rainier Beach. We took steps to account for this in our analyses as much as possible, as we describe in the following section.

SPS provided school climate survey and administrative data for each of the schools in our treatment and comparison areas. We describe the contents of these datasets in more detail in the following section. Because the school data are protected by FERPA, we have anonymized the school names in this report to reduce the risk that students might be indirectly identified via any of our analyses.[2] We coded the school names according to whether they are elementary (ES), middle or K-8 (MS), high (HS), or alternative high (AS) schools. The treatment neighborhood (Rainier Beach) contains three elementary schools (ES3, ES4, ES5), two middle or K-8 schools (MS1, MS5), one high school (HS5), and an alternative high school (AS5). Across the two comparison sites there are two elementary schools (ES7, ES13), three middle/K-8 schools (MS6, MS11, MS12), two high schools (HS6, HS7), and an alternative high school (AS8). We originally included a third elementary school (ES10) in our comparison area, but during the project period the school closed for renovations and temporarily relocated to a different neighborhood. Although we collected data for this site, we elected to exclude it from all of our analyses because the RBCSC was intended to be closely aligned to place/community and we felt that the change of location compromised this.

Similarly, due to the focus of our program on aligning school and community supports and interventions, we chose to assess community and crime outcomes within a defined geographic area around each school. We identified calls for police service and crime incident reports recorded by SPD that fell within a 1000ft buffer around each school in the treatment and comparison areas.[3] In both the treatment and comparison areas, some of the schools were located in such close proximity to each other that we could

[2] Per SPS request, we also aggregate data as much as possible so there are no variables that identify fewer than ten students.
[3] See Gill et al. (2015) for more details about the geocoding process we have developed for Seattle police data.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

not create separate 1000ft buffers around each of them. We elected to create combined buffers, so multiple schools are included in sites 5, 6, and 7.[4] We used the same 1000ft buffers to sample addresses for the community survey. Prior to conducting the first wave of the survey, we obtained from ArcGIS a list of all addresses that fell within the 1000ft buffers. Local researchers conducted an in-person census to visually verify the addresses and document apartment numbers and multi-family dwellings that were not captured in the program. We used this final list of addresses to draw our sample for the baseline community survey, as described in the next section.

3 Data and Analytic Strategy

3.1 School climate survey

SPS conducts a research-based school climate survey, formally called the Student Survey of School Climate, in the fall and spring of each year with students in grades 3-12. The results are used within the school district to assess and plan for student success and guide school- and district-level improvements. They are also made available in the aggregate to the public via an online data portal. The surveys ask students a variety of questions (items) about their school experiences and perceptions, which (for the years we analyze, as explained below) are organized on the SPS data portal into eight constructs:

1. Healthy community: Students' enjoyment of school and perception that students are treated fairly and with respect by adults and peers.
2. Belonging: The extent to which students feel they belong and are cared about at school.
3. Classroom environment: Students' perceptions of interactions with their peers in a learning context.
4. School safety: Students' feelings of safety in school and the surrounding neighborhood, as well as how well they think the school addresses bullying.
5. Motivation and inclusion: How well students think their teachers motivate them to learn and grow.
6. Pedagogical effectiveness: Students' perceptions of their teachers' performance in the classroom.
7. Learning mindset: Whether students believe they work hard, challenge themselves, and set future goals.

[4] An additional benefit of this approach is that it further maintains the anonymity of some of the schools' data, especially in the treatment neighborhood where the schools may be more easily identifiable.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

8. Social-emotional learning: Students' perceived ability to self-regulate and relate to others.

SPS provided us with raw data files of individual student responses to each survey item for all of our project schools from the 2014-15 school year through the 2021-22 school year, with the exception of 2020-21. Although surveys were conducted that year, they were designed to assess students' experiences of remote learning during the pandemic and were not comparable to other years. For the 2014-15 academic year we received the Spring 2015 data only; for subsequent years from 2015-16 through 2019-20 we received both the Spring and Fall datasets (with the exception of Spring 2020, which was cancelled), as well as additional data for surveys conducted in Winter (2016-2020). For the 2021-22 school year, we received data for Fall and Spring.

We ultimately decided only to analyze the Spring datasets for 2016 through 2019. The Fall and Winter surveys were not conducted consistently across each school, and in some cases most or all of our project schools did not participate in these waves (in general, participation rates varied substantially across schools and years). In addition, the individual survey items and the overarching constructs they measured in the Spring 2015 survey were somewhat different from those used in subsequent years, so we dropped this first year of data. In 2021-22, the survey items and constructs were substantially overhauled. With a few exceptions, different items were used to measure the constructs that did remain consistent with prior years. The new survey instrument also had different response options for a number of items, and was updated to better represent students' self-reported gender and racial/ethnic identities. Thus, we felt these surveys were also not comparable to prior years. While the 2019 survey was conducted only a few months after the overall implementation start date, note that some schools were conducting training and initial Tier 1 PBIS implementation throughout the analysis period and the choice of 2019 as a start date reflects full implementation, including the community piece. Nonetheless, this is a limitation of the analysis. The cut-off point of 2019 also means that our school climate analysis is not affected by the COVID-19 pandemic.

Thus, in this report we analyze three years of data pre-implementation (2016, 2017, and 2018) and one year post-implementation (2019), all of which use the same items to assess the constructs described above. Each construct contains between 4 and 6 items, all of which are measured on 5-point scales. Questions in the first six constructs listed above are measured on agreement scales, while the Learning Mindset and Social-Emotional Learning items are measured according to the extent to which students

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

identify with each item/statement on a scale ranging from ”not like me at all” to ”very much like me.” The datasets we received did not include any summary measure for each construct, so we chose to create scales based on the mean value of the students' responses to each item in the construct. Table A1 shows descriptive statistics for each scale by wave, as well as the Cronbach's alpha (α) for each one. As we would expect given that these constructs were created intentionally by the survey developers, all of the alpha scores exceed 0.7, which is the rule of thumb for a ”good” scale.

There was one inconsistent item in the surveys we analyzed: in 2018 there was a change to the question asking students to self-report their academic performance. In the first two years, students were asked ”what were your grades like last year?” with 4 response options: Mostly As, Mostly Bs, Mostly Cs, Mostly Ds/Es. In 2018 and 2019 they were asked ”what kind of grades do you usually get?” The response scale included five options: Very high, High, Good, Some good/some not, Not very good. In order to use these questions as a single control variable in all years, we created a new combined variable, ”Self-reported grades,” with four response options. We combined ”High” and ”Very high” from the newer question and equated that to ”Mostly As” in the older version of the question. The remaining categories aligned with each other, e.g. ”Good” = ”Mostly Bs,” etc.

To analyze the data, we developed multilevel mixed-effects linear regression models for each construct, following the approach of Kochel and Weisburd (2017). These models include fixed effects for survey wave, treatment assignment, and control variables. The coefficient of interest is the interaction term *wave* $\times$ *treatment assignment*; specifically, the interaction between 2019 (the only post-implementation wave we were able to analyze) and treatment assignment. This interaction term compares 2019 to 2016, the first pre-intervention year in the dataset. However, we opted to also include interaction effects for each wave rather than simply compare pre- and post-implementation data to better account for the imbalance in the number of years included in each time period. The models also include a random effect for school to account for the nesting of students within schools, as well as robust standard errors to further account for clustering. Analyses were performed in Stata 17.

An important limitation of the data is the lack of an identifying number that tracks students throughout their school experience, unlike the school administrative data we describe later. This means we are not able to account for the nesting of multiple student responses each year or track the movement of students between, for example, elementary and middle school during the timeframe of the dataset. To partially account for this we opted to analyze elementary, middle, and high school respondents separately rather than combining the data and controlling for school type.[5] Finally, all the models include controls for gender, race, the extent to which students reported English is spoken in their home, and their self-reported grades. Gender was the only variable that did not significantly differ between treatment and comparison groups at baseline (2016), but we opted to include it in the models as it may have affected students' responses to specific survey items or constructs (see Tables A2-A4 for baseline characteristics and differences).

3.2 School administrative data

In addition to the school climate survey data, SPS provided us with a variety of administrative datasets that they collect for student tracking and monitoring purposes. These datasets covered the academic years 2014-15 to 2021-22, including the 2020-21 school year when the district was closed due to the pandemic, and contained information about 28,638 students, including those who were enrolled at one of our project schools during the timeframe of the data. The datasets contained at least one row per student per year; sometimes more if the student transferred schools within a school year. Unlike the climate survey data, the administrative datasets contained a unique proxy identification number that SPS created based on the students' names prior to transferring the data to us, so we were able to track students over time and across schools, at least within the SPS system. The proxy identifier also allowed us to match student data across each individual administrative dataset. The specific datasets provided were as follows:

- **Enrollment history.** This dataset contained one row per enrollment and comprised students' demographic information and enrollment history (N=83,272). This included school and grade; gender and race as recorded by the school district; the student's living situation (e.g. living with both parents, mother/father only, other family etc); the primary language spoken by the student and in the student's home; projected graduation year; and whether the student was eligible for and/or receiving a variety of services, including English-language learner (ELL), advanced learning, and special education. All students were included in this dataset, even if they did not have observations in the other datasets listed below.

[5] Some of the schools in our project sites are also K-8 schools, meaning that they included survey responses for both elementary and middle school students. Thus, although students are likely double-counted across years in our analysis, separating the models by school type ensured we did not double-count students within schools.

- **Academic achievement.** We received two datasets pertaining to academic achievement measures: GPA data and state assessment results. The GPA dataset contained one row per student per year for middle and high schoolers only. The data points included cumulative GPA and credits earned each year. The state assessment dataset included test results for English Language Arts (ELA) and Math. These tests are conducted at all school levels, with multiple tests in elementary and middle school and one test in high school. They are conducted in spring and fall, but we only received spring data. The datasets contained one row per student per test type (ELA/Math) per year, as well as whether or not the test was attempted, the score if so, and whether or not the score met state standards. Both datasets were affected by the COVID-19 pandemic. In the Spring 2020 semester when the pandemic began, the school district grading policy was As or incompletes. In the 2020-21 school year, the grading policy was As, Bs, Cs, or incompletes. This affects the GPA data. State assessments were not conducted in Spring 2020 or 2021.
- **Attendance.** This dataset contained one record per student per enrollment per year, and reflected the number of days students were eligible to be in school that enrollment period, the number of excused and unexcused absent days, and the number of tardy days. Absence data are undercounted in 2019-20 and 2020-21 due to the shift to remote learning during the pandemic, and tardies were undercounted in 2021-22 until around mid-October.
- **Discipline.** This dataset only tracked discipline data for students who had any disciplinary actions (N=1,021, one row per student per year). This dataset was only available for 2015-16 onwards. The dataset contained total disciplinary actions, short-term suspensions, long-term suspensions, and expulsions. Discipline data was not collected while students were learning remotely during the pandemic.

To prepare the administrative data for analysis, we combined all of the files into a single dataset with one row per student per academic year. Prior to analysis we decided to drop records for the following reasons:

- One of the alternative high schools in the project also operates a GED program and a program for incarcerated students. We received data for students enrolled in these programs as well as the school's main campus and opted to exclude them from the analysis due to the very different experiences of the students and the lack of physical presence in the project area.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

- We dropped records for the 2014-15 academic year because of the lack of discipline data for that year. This also better aligned the timeframes of the administrative data and the school climate survey data.
- We excluded students who did not have both pre- and post-intervention enrollment data.
- We excluded students who transferred between treatment and comparison schools during the analysis period to reduce potential contamination.

These adjustments created a final dataset with 34,416 observations and 6,917 unique students (treatment group N = 1,972, comparison group N = 4,945). We subsequently made adjustments for a handful of students who still had more than one record in a given year; these students were typically enrolled in alternative high schools and had brief (e.g. one-week) breaks in enrollments, or moved from a traditional to an alternative high school. In these cases we kept the enrollment record with the highest number of enrolled days after verifying that the other outcome variables remained the same. A few students also had more than one state test record in a given year because they had an initial incomplete attempt followed by a completed attempt, or they took the test at two different levels (e.g. 6th grade and high school) in the same year. In these cases we dropped the incomplete or lower grade test. Finally, due to the small number of students within middle schools (likely due to the short period of time students spend in middle school, i.e. grades 6-8, relative to elementary and high school), the analytic models we describe below would not run correctly. We opted to investigate separately the effects of the program for students in elementary and high schools only.

We used propensity score-weighted multilevel mixed-effects regression models with robust standard errors to analyze school administrative outcomes. In most cases these were linear regression models for continuous variables, but we used multilevel mixed-effects negative binomial regression for the discipline model, which is based on the count of disciplinary actions. Similar to our analysis of the school climate survey data, the interaction between treatment assignment and time (pre- vs. during implementation) is the coefficient of interest in these models. Because we were able to track students via a unique identifier in the school administrative datasets, we were able to use propensity score analysis techniques (e.g. Rosenbaum & Rubin, 1983, 1984) to create a subset of students in treatment and comparison schools who were more comparable to each other in the absence of randomizing or matching schools in our evaluation design. Propensity score analysis balances the covariates that predict receiving the treatment

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

across samples of treated and similarly-situated non-treated individuals or groups. We elected to then include the propensity score as a weight in the final outcome analysis models, which means we do not have to control for the wide variety of factors that might affect the outcome at this stage (e.g. Gill & Wilson, 2017).

Propensity score analysis is a two-stage process in which we first construct a logistic regression model on treatment assignment to explore the selection mechanism; i.e., which covariates (here, student characteristics) predict treatment assignment. The clustered nature of the school data (i.e. students nested in schools) complicates the propensity score estimation process here. Technically, the estimation of the propensity score should account for the clustering (e.g. Arpino & Cannas, 2016; Arpino & Mealli, 2011). However, in our case the clustering variable (school) perfectly predicted treatment assignment because we had already sorted schools into treatment and comparison conditions as part of the research design. As a result, we decided to use ”naive” propensity score estimation and account for the clustering in the final analysis using the multilevel mixed-effects models with robust standard errors.

The covariates we included in the propensity score estimation model were gender;[6] race; whether or not the student lived with both parents, whether the student's primary language was English and whether English was spoken at home; and whether the student received and/or was eligible for English-language learner services and advanced learning services (Tables A13-A14). We estimated the propensity score using only each student's first pre-intervention observation in the dataset as there was some, but very little, time variance in these covariates that further complicated the process, and we estimated propensity scores for the elementary school and high school samples separately. Propensity score estimation was conducted using the `psmatch2` package in Stata 17 (Leuven & Sianesi, 2016), with matching of 3 nearest neighbors to account for the imbalance of sample sizes in the treatment and comparison groups, and using a caliper of .25 to identify matches. The models performed reasonably well in rebalancing the samples. The `psmatch2` package reports Rubin's *B* and Rubin's *R*, which Rubin (2001) recommends should be less than 25 and between 0.5 and 2 respectively. For elementary schools these values were 16.8 and 0.48 respectively, and for high schools 8.8 and 0.1. Thus Rubin's *R* was slightly out of range for both samples. Figures A6 and A7 visually represent the effectiveness of the matching process at re-

[6] We excluded students whose gender was recorded as non-binary as the number was very small; we recognize that the number of students who do not identify specifically as male or female is likely larger, given that students did not self-identify their gender in this dataset.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

ducing bias across the two school samples.[7] Finally, we calculated the inverse probability weight of the propensity score (1/ps) to use as the weight in the final models.

The outcome measures we selected for analysis were as follows (see Table A15).

- Academic achievement: cumulative GPA (high school only); ELA state test score; Math state test score.
- Attendance: Total absences, excused absences, unexcused absences, absences as a proportion of eligible days, tardies as a proportion of eligible days
- Discipline: Count of disciplinary actions. We did not look at suspensions or expulsions separately as the overall number of disciplinary actions were very small and there were only 10 expulsions across the entire dataset, so suspensions and total actions were strongly collinear. We analyzed the total count of disciplinary actions only. No records were provided for students who had no disciplinary actions, so we coded all of these records as zero for this variable; however, we cannot confirm whether these are true zeros or missing data.[8]

A limitation of the outcome models is that we could not appropriately control for the academic year. We aggregated all data for 2014-15 through 2017-18 into a "pre-implementation" variable and all data for 2018-19 onwards as "during implementation." Note that we used 2018-19 as the first during-implementation year even though full implementation of the program began in January 2019. Although the community implementation did not begin until January, some PBIS/RJ training and Tier 1 implementation was occurring prior to that date. This start date also gave us a balanced number of years in the pre- and during-implementation period for most analyses. Nonetheless, we were unable to include the specific academic year as a control variable or a random effect. Using it as a control variable created too much instability in the models, and it was not appropriate to use it as a random effect because students and schools are not nested within years or vice versa. This also means that we are unable to effectively control for the impact of the pandemic in these models.

[7] Note that we did not treat race as a factor variable when we ran the propensity score estimation model to create these graphs, in order to simplify the visualization. However, the model that we used to estimate the propensity score used in the final outcome analysis did include race as a factor variable. The two models performed similarly in terms of the matching.

[8] We coded records for 2014-15 as missing rather than zero because no discipline data was provided for that academic year.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

3.3 Crime and calls for police service

SPD has a longstanding data agreement with the CEBCP, going back to 2010. We receive monthly downloads of calls for service data from their computer-aided dispatch (CAD) system and incident reports from the records management system (RMS). CAD data includes both 911 calls from the public and police-initiated activity (including logs of breaks, special patrols etc.). RMS data includes police reports of validated offenses and incidents. SPD uses the National Incident-Based Reporting System (NIBRS) to classify reports as offenses (crime events) or incidents (non-crime events reported by police, such as a death by natural causes or vehicle impound). Offenses are further classified into Group A offenses, which are more serious, and Group B, which are minor crimes, and categorized as crimes against persons, property, or society. The RMS data also includes an "entity dataset," which provides details and roles (suspect, victim, etc.) of people involved in each offense or incident. SPD also provides us the full address of each call and incident to allow for more accurate geocoding to street segments, as the publicly-available X-Y coordinates are offset to preserve confidentiality.

Our analysis focused only on calls for service and police reports that are crime-related in nature (offenses). We coded calls for service using the three NIBRS categories, based on the description of the call. We added categories for other crime-related calls, traffic, alarms, police administrative logs, and non-crime calls. With the exception of the other crime-related calls category, these additional categories were not used in the analysis. We retained all offense reports categorized as person, property, society, and other (we created the 'other' category from offenses with the NIBRS code "90Z - All Other Crimes"). We used dates of birth in the entity dataset to calculate whether suspects and victims were under 18 at the time of the offense and coded each incident according to whether it involved a juvenile suspect or victim. Because of the small number overall, we defined "juvenile offenses" as any offense involving a juvenile suspect *and/or* victim, as the overarching goal of our initiative was to improve safety and neighborhood climate for all young people.

We analyzed all crime outcomes using difference-in-differences random effects negative binomial regression models.[9] These models estimate the effects of the initiative on monthly crime outcomes in the treatment and comparison sites according to the months in which the initiative was active. Given our longstanding relationship with SPD, we were able to include a long time-series of data for all three

[9] This approach follows Kondo et al. (2015) but uses negative binomial regression to account for overdispersion.

crime outcomes. Our pre-intervention analysis period ran from July 2014-December 2018, and our post-intervention period ran from January 2019-June 2022 (96 months total). In each model, the unit of analysis is the monthly count of the outcome (i) in the combined treatment and comparison sites (t). The models include indicator variables for whether the overall program was active or inactive (P_{it}), treatment vs comparison (A_{it}), controls for seasonality, overall monthly trend, and spatial autocorrelation,[10] and additional controls for the overall effect of the pandemic and related school closures.[11] The parameter of interest is a difference-in-differences interaction term, ($P_{it} \times A_{it}$). Finally, the model includes a random effect to account for clustering of outcomes within each school buffer area. We report the exponentiated coefficients—incidence rate ratios (IRR). Here, the IRR represents the ratio of crime counts in the treatment areas to crime counts in the comparison areas associated with the program. We performed the analyses in Stata 15. Table A23 shows the pre-intervention period mean monthly counts of each outcome in the individual treatment and comparison buffers.

3.4 Community survey

Our final data collection and analysis strategy involved a survey of community members living within the buffer areas around each project school. We aimed to assess whether the intervention affected community perceptions of social cohesion, collective efficacy, safety, disorder, and policing. We developed our own survey instrument, which drew upon measures from the Project on Human Development in Chicago Neighborhoods (e.g. Morenoff et al., 2001; Sampson & Morenoff, 1997; Sampson et al., 1997) and earlier community surveys we designed and conducted for Rainier Beach: A Beautiful Safe Place for Youth (e.g. Gill & Prince, 2020; see also Gill & Gross Shader, 2020).

As we discussed above, we originally intended to conduct a three-wave in-person survey of addresses in the treatment and comparison areas. Due to the COVID-19 pandemic we were unable to conduct the data collection as planned and ultimately pivoted to a two-wave survey in which the second wave was conducted by mail. The implementation of the first wave of the survey in the summer and fall of 2019

[10] There was a high level of autocorrelation in the buffer areas, perhaps due to their small size. In some models we controlled for autocorrelation over 4 months. We calculated separate monthly lag variables to control for autocorrelation in each buffer area, based on the natural log of the monthly crime count from the previous month, 2 months, etc., adding 1 to each value to account for months where no crimes were recorded.

[11] 'Pandemic active' is a simple pre/post indicator variable. Seattle was one of the first places in the US to experience the effects of the pandemic, beginning in late February 2020. We coded all months from March 2020 to June 2022 as 'pandemic active.' Seattle Public Schools closed on March 13, 2020 and reopened in full in September 2021, although certain students were allowed back in person at various points prior to that date. We coded months from March 2020-August 2021 inclusive as 'schools closed.'

was conducted as planned. We closely followed the methods of Weisburd et al. (2021) in designing the door-to-door survey process. We set a goal of conducting 200 surveys in treatment locations and 200 in comparison locations, following the power calculations in our grant proposal. After local researchers conducted the residential census described above and finalized the address list for each location, we sampled 2.5 times the number of addresses per survey we hoped to complete (i.e. roughly 500 addresses per group or 1,000 total). In practice, due to the uneven number of schools in each group and the geographic overlap of some schools, we weighted our sampling approach by school, releasing approximately 72 addresses per treatment school and 63 per comparison school.[12] This led to an overall goal of 30 surveys per treatment school location and 25 surveys per comparison school location. Where schools overlapped geographically we added up these goals to get an overall target number for the buffer area; for example, we aimed for 120 completed surveys in site 5, a treatment area containing 4 schools (30*4) and 50 completed surveys in site 6, a comparison area containing 2 schools (25*2). The lead author drew the sample by assigning each address within each group a random number using the Excel formula and sorting from smallest to largest. The first set of addresses that fell into the target sample for each site (number of completed surveys*2.5) were released.

Teams of at least two researchers, including a team leader, visited each address at various times, primarily on weekdays but also on some weekends, to request interviews. If a resident answered the door, the interviewer would verify whether they were at least 18 years old or ask if an adult was present and engage in a brief recruitment script. Participants who were interested then went over a consent form with the researchers, and the interview began once informed consent was obtained. While we did not have full coverage of the variety of languages spoken in Rainier Beach and the comparison neighborhoods, members of the research team were proficient or native speakers of several languages, including Spanish, Somali, and Mandarin Chinese, and were able to interpret or help respondents work through a written English version of the survey as needed. Table A28 shows that the research team met or exceeded the target number of completed surveys in almost all sites in Wave 1.

We delayed the next wave of the survey multiple times due to the COVID-19 pandemic and missed the opportunity to conduct two more waves as we frequently re-evaluated the safety of participants and the research team. Eventually we opted to conduct one more wave of the survey by mail, as we were

[12] These numbers include the neighborhood around ES10, the school that was ultimately dropped from the analysis because of the temporary change of location.

concerned about the continuing effects of the pandemic and the extensive effort that would be needed to rehire and retrain a research team after so much time had passed. This survey was administered in the fall of 2022. After deleting addresses that were listed as no longer valid after the first wave of the survey, we mailed a paper copy of the survey to every address at which a Wave 1 survey was completed (N=456, including the ES10 neighborhood), and sampled a further 544 addresses that were contacted in Wave 1 but did not complete a survey, for a total of 1,000 mailed surveys. Respondents could choose to fill out the survey on paper and return it in a prepaid envelope, or scan a QR code/type in a URL on the survey cover sheet to complete the survey online. To address the linguistic diversity of the communities, each mailer also included an information sheet translated into Amharic, Traditional Chinese, Oromo, Somali, Spanish, and Vietnamese—the languages most commonly used by families with children in Seattle Public Schools as well as in Rainier Beach—inviting speakers of those languages to scan the QR code and take the survey in their own language. We worked with a language access company based in the Pacific Northwest, which had expertise in the specific languages spoken in Seattle, to translate the survey instrument and related documentation.

Unfortunately, as Table A27 shows, engagement with the mail survey was far less successful than the door-to-door survey. We received only 49 valid responses, of which 7 were excluded from the analysis because they were from residents near the ES10 neighborhood. Eighteen of the responses, not including those from ES10, were from individuals at addresses where a survey was completed in Wave 1, and 24 were from addresses that had not completed a survey before. Around 40 survey mailers were also returned to us by the US Postal Service as undeliverable.

Due to the lack of engagement with the Wave 2 follow-up survey, our analysis is limited. As with the school climate survey, we combined items into scales based on overarching concepts of social cohesion, collective efficacy, community engagement, police effectiveness, police legitimacy, feelings of safety, and perceptions of disorder. The scales were based on the mean of each participant's response to the individual items within each construct. Table A29 shows the descriptive statistics for each outcome scale, including Cronbach's α and the number of items. We also analyzed the effects of the program on three single-item outcome measures that we felt were especially relevant to the program goals: perceived change in neighborhood safety and police protection over the past five years, police visibility (likelihood of seeing police in the neighborhood), and satisfaction with police.

We use linear regression models to assess the scaled outcomes, and ordered logistic regression to assess the single-item outcomes, all of which were measured using three- or four-point Likert scales. As in our other analyses, the coefficient of interest is the interaction between treatment and time (Wave 2 vs Wave 1). Ideally we would have used multilevel mixed-effects models, as with the school outcomes, for the nesting of respondents within addresses and buffer areas, or at the very least controlled for the buffer area. However, due to the extremely low response rate in Wave 2 and the small number of households who responded in both waves, any attempt to include random effects or control for buffer areas introduced too much instability into the models. Clustered standard errors also did not work, so we use basic one-level fixed-effects models for all analyses with robust standard errors to account for general misspecification. All models control for respondent characteristics that significantly differed between the treatment and control group respondents in the Wave 1 survey (Table A28): race, child currently in school, born in United States, highest level of education completed, home status (rent vs. own), household income, and crime victimization in past year.

4 Project Findings

Our project findings include statistically significant reductions in police calls for service and offenses, and non-significant reductions in offenses specifically involving young people. However, our analyses of school administrative and climate data and our community survey were hampered by pandemic-related data collection and recording issues and do not present a clear picture of the impacts of those aspects of the program. We describe the findings from each dataset in more detail below.

4.1 School climate

Tables A5-A12 show the results of the multilevel mixed-effects linear regression models by school type (elementary, middle, high) for each of the SPS school climate survey scales. Overall, the results are mixed and do not paint a conclusive picture of the impact of the program on students' perceptions of school climate. For the healthy community construct, the interaction term for post-implementation (2019) by treatment assignment is negative for elementary and high schools, meaning that students in these schools were less likely to agree that their school community was healthy compared to 2016 (Table A5). For high school respondents, this interaction was statistically significant (see also Figure A2). However, for middle school respondents there was a slight improvement in this construct, albeit non-significant. Note that

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

these trends are largely reflected in the two pre-intervention interaction terms as well, so the difference in 2019 may not be a direct result of the program.

We see similar trends for the other scales. For belonging, we also see a slight improvement in middle schools but declines for elementary and high schools. However, these trends appear to have been in play throughout the analysis period and the interaction terms with treatment assignment are not statistically significant in any year (Table A6). For classroom environment, the interaction coefficients are negative for all years and non-significant (Table A7). There were no significant program effects on the school safety scale, but there were small improvements for elementary and middle school students (Table A8).

For motivation and inclusion, pedagogical effectiveness, and learning mindset, the results mirror those for the healthy community scale. Elementary students scored lower on these scales, although this was not statistically significant. Middle school students scored higher, but this was also not significant. Scores were significantly lower for high school students in treatment schools post-implementation (Tables A9-A11; Figures A3-A5). Finally, for social-emotional learning scores were slightly lower, but not significantly so, for all school types (Table A12).

Note that a limitation of all of these models is the inability to calculate the Wald chi-square, as reflected in the tables. This issue occurs in mixed models when the number of coefficients exceeds the number of clusters (i.e. the school clusters modeled in the random effect). While the overall significance of the model (i.e. the test of whether all coefficients are jointly zero) is not relevant to our conclusions, this is an indication that the model may not be ideal for the data and results should be treated with caution.

4.2 School administrative outcomes

Tables A16-A22 show the results of the propensity score-weighted multilevel mixed-effects models on school administrative outcomes. As with the climate data, there are mixed findings across this set of outcomes. Among elementary school students, state English-language arts (ELA) and math assessment scores were significantly lower in the treatment schools during the program (Table A16). Similarly, ELA scores were significantly lower for high school students. Math scores were also lower, but this was not statistically significant. There was no effect on GPA for high school students (Table A17).[13] The predicted margin plots for each of these models show that scores actually improved in both groups in the during-

[13] The models for high school state assessments only include random effects for schools, not students, since students are only assessed once during high school.

intervention period, but the improvement was not as great in the treatment schools. For high school students, scores in the treatment group remained relatively stable but improved slightly in the comparison group (Figures A8-A10).

Tables A18 and A19 show the effects of the program on a variety of attendance-related outcomes for elementary school students, including total, excused, and unexcused absences (Table A18) and absences and tardiness adjusted for the number of eligible attendance days (Table A19). There are no statistically significant differences for these outcomes, although all but the proportion of eligible days recorded as tardy were lower in the treatment group during the intervention. The same outcomes for high school students are shown in Tables A20 and A21. Here, all outcomes are lower in the treatment group during the intervention, and these findings are statistically significant for the number of excused absences and the proportion of eligible days marked as unexcused absent. Figure A11 shows that the number of excused absences decreased in the treatment group but increased in the comparison group. Figure A12 indicates that while the proportion of eligible days marked as unexcused absent increased in both groups, the increase was slightly less steep for the treatment group.

Finally, Table A22 shows the results of the multilevel mixed-effects negative binomial models for the count of disciplinary actions per student in elementary and high schools. Coefficients in Table A22 are exponentiated to show the incidence rate ratio (IRR), so the interaction term represents the ratio of disciplinary action counts in the treatment group to the counts in the comparison areas during the implementation period associated with treatment. There were substantially fewer disciplinary actions in both school types in the treatment areas during the implementation period: the IRR for the interaction term in the elementary schools model represents an 89% reduction in disciplinary actions and is statistically significant. For high schools, the interaction term is not statistically significant but still represents a 23% reduction. Note that the number of disciplinary actions is very small overall, and we coded all students who did not have disciplinary actions recorded as having zero actions, but some may be missing data. Note also that no disciplinary actions were recorded for part of the 2019-20 academic year or in 2020-21.

As we discussed in the Methodology section, none of these models controls for the effects of individual academic years. Because the COVID-19 pandemic affected a great deal of the during-implementation period, it is likely that the results we see here are driven by the effects of the pandemic rather than the RBCSC program. Although students are matched in this analysis, Rainier Beach as a neighborhood was significantly affected by the pandemic and students in the treatment schools within this neighborhood

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

may have experienced more extensive hardships than similarly-situated students in schools located in less marginalized communities. On the other hand, undercounting and limited reporting during this period may be driving lower rates for some outcomes (disciplinary actions in particular), although we would expect this to be consistent across treatment and comparison schools because pandemic-related policy changes were made at the central SPS level. Nonetheless, we urge caution in interpreting these results.

4.3 Racial disparities in school outcomes

One of SPS's key goals in implementing school-based PBIS and restorative practices was to attempt to reduce racial disparities in school outcomes. This aligned with the district's strategic goals during the project period, one of which was "educational excellence and equity for every student." After running the overall analyses for school climate and administrative outcomes, we also investigated whether race mediated the effects of the treatment on each outcome. Following Gross Shader (2020), who drew upon an analytical approach used by public health researchers (McDougal et al., 2017), we reran each model including a three-way interaction term between treatment, time, and race and then calculated stratum-specific coefficients comparing the program effects for White students (the reference category) against those for students of other races. We found a number of significant differences between White students and students of other races from these analyses, even when the overall outcome was not statistically significant. Due to the large number of findings overall, we focus here on the mediating effects of race in the school climate and administrative models that were statistically significant overall. Note that to simplify the school climate survey models we used a simple pre/post time variable instead of interactions for each year in the dataset here.

For the school climate outcomes, recall that in high schools the program was associated with significantly lower ratings for the healthy community, motivation and inclusion, pedagogical effectiveness, and learning mindset scales. For healthy community, the *treatment* $\times$ *time* $\times$ *race* interaction was statistically significant for Asian and multiracial students. While the overall interaction of *treatment* $\times$ *time* in the main model was negative, the stratum-specific coefficient for White students was small but positive (b = .012). However, for Asian students the stratum-specific coefficient was -.195 (significantly different from White students, χ^2 = 57.684, p<.0001) and for multiracial students -.397 (χ^2 = 6.122, p = .013).

For motivation and inclusion in high schools, Black, Hispanic, and multiracial students all significantly differed from White students. Again, the *treatment* $\times$ *time* coefficient for White students was small but positive (and statistically significant, $b = .076$), but it was negative for students of other races. For Black students, the stratum-specific coefficient was -.120, for Hispanic students -.235, and for multiracial students -.392. All of these significantly differed from the stratum-specific coefficient for White students (Black: $\chi^2 = 4.972$, $p = .026$; Hispanic: $\chi^2 = 38.393$, $p<.0001$; multiracial: $\chi^2 = 8.074$, $p = .004$).

We see a similar pattern for pedagogical effectiveness in high schools, with White students experiencing a small (though not statistically significant) positive effect while Hispanic and multiracial students gave significantly lower ratings in this domain. For White students, $b = .096$, for Hispanic students -.189 (comparison to White students: $\chi^2 = 4.794$, $p = .029$), and for multiracial students -.353 (comparison to White students: $\chi^2 = 4.745$, $p = .029$).

Finally, there were different effects of race for learning mindset in high schools. The program was associated with significantly lower ratings for White students ($b = -.131$), but positive effects for Black ($b = .087$) and Asian ($b = .035$) students. The stratum-specific coefficient for Black students was statistically significant. The effects for both Black and Asian students were significantly different from White students (Black: $\chi^2 = 9.906$, $p = .002$; Asian: $\chi^2 = 5.377$, $p = .020$).

In our analysis of school administrative data for elementary school students, we saw significant negative effects associated with the program period on ELA and math state assessment scores, and a significant positive effect on school discipline. Interestingly, there were some positive mediating effects of race in the state assessment models, in that some groups of non-White students experienced significant improvements in test scores compared to their White counterparts. In the ELA test, scores for White students were lower in the treatment schools during the program ($b = -31.235$), in line with the overall model (although the stratum-specific coefficient for White students was not statistically significant). However, for Native American and Pacific Islander students in the treatment group, test scores were significantly *higher* during the program period (comparison to White students: $\chi^2 = 49.885$, $p<.0001$). In the math test, scores were significantly lower during treatment for White students ($b = -28.191$), but significantly higher for Hispanic ($b = 24.297$, difference from White students: $\chi^2 = 6.475$, $p = .011$) and Native American/Pacific Islander ($b = 19.415$, difference from White students: $\chi^2 = 6.475$, $p<.0001$) students. We were unable to calculate stratum-specific coefficients for disciplinary actions because the number of actions were so small that the model with the three-way interaction effect was unstable.

In high schools we saw overall significant negative effects associated with the program for ELA scores and positive effects for excused absences and unexcused absences as a proportion of eligible days. Similar to the results for elementary school students, White high school students experienced significantly lower ELA scores during the treatment (b = -84.753), but Black (b = 9.188, not significant), Asian (b = 23.742, significant), and multiracial (b = 55.802, significant) students all earned significantly higher scores compared to their White counterparts (Black: $\chi^2 = 29.118$, p<.0001; Asian: $\chi^2 = 285.021$, p<.0001; multiracial: $\chi^2 = 67.933$, p<.0001). However, Hispanic students' scores were also significantly lower compared to White students (b = -44.271, significant; χ^2 compared to White students = 19.282, p<.0001).

For excused absences, the mediating effects of race were mixed. Although there was a significant reduction in excused absences overall, each *treatment* $\times$ *time* $\times$ *race* interaction for non-White students was positive (i.e. absences increased) and, with the exception of Hispanic students, statistically significant. The stratum-specific program effects for Asian, multiracial, and Native American/Pacific Islander students significantly differed from that of White students, and the results were mixed. The stratum-specific program effect coefficient for Asian students also indicated a reduction in excused absences, though at a smaller magnitude than the effect for White students (b = -.799 vs -2.451 for White students; difference from White students: $\chi^2 = 24.372$, p<.0001). However, the coefficients for the stratum-specific program effects for multiracial and Native American/Pacific Islander students were both positive (b = 2.281 and 4.288 respectively. Differences from White students: $\chi^2 = 8.406$, p = .004 (multiracial), χ^2 = 74.406, p<.0001 (Native American/Pacific Islander)), indicating that these two groups of students did not benefit from the possible program effects on excused absences in the same way as White students did.

Finally, we also found a number of differences by race in the proportion of eligible attendance days classed as unexcused absences. Overall, the program is associated with a significant reduction in this ratio. In the overall race-adjusted model, almost all of the three-way interaction terms with the exception of Black students indicated statistically significant reductions relative to White students (for Black students, the interaction term also indicates a reduction but it was not significant). In contrast with the other models, White students experienced a significant increase in the proportion of eligible days classed as unexcused absent (b = .102, while Asian (b = -.032), Hispanic (b = -.062), and Native American/Pacific Islander (-.098) students all experienced a reduction, although the stratum-specific coefficient for Hispanic students was not statistically significant. The stratum-specific effects for Asian, Hispanic, and Native American/Pacific Islander students were all significantly different from the stratum specific effect for White students (Asian: $\chi^2 = 106.547$, p<.0001; Hispanic: $\chi^2 = 6.448$, p = .011; Native American/Pacific

Islander: χ^2 = 17.217, p<.0001).

4.4 Crime and calls for police service

Figure A14 shows the overall trend in calls for service in the combined treatment school buffer areas (orange line) and comparison school buffer areas (blue line) across the entire analysis period. Calls in the treatment areas were already trending slightly lower than the comparison areas prior to the intervention start date, perhaps because of the other interventions taking place in the treatment neighborhood. However, the comparison areas saw a much more steep increase in calls than the treatment areas during the spring and summer of 2020, when the pandemic and police protests were taking place. Figure A15 shows the pre/post January 2019 change in calls in the treatment and comparison sites, adjusted for the unequal number of months in the pre- and post-intervention periods. Although there were 15% more calls in the treatment sites post-intervention, there were 84% more in the comparison sites. Table A24 shows that the intervention is associated with an 11% lower rate of calls for service, controlling for seasonality, trend, autocorrelation, and the pandemic and school closures. This result is statistically significant (IRR $= .887, p = .003$).

Figure A17 shows that offenses have consistently been higher in the comparison areas throughout the analysis period. However, offenses have continued to trend downward in the treatment areas, while the pattern in the comparison areas is less clear. Overall, there were 33% fewer offenses in the treatment areas in the post-intervention period compared to 12% fewer in the comparison areas (Figure A18). Similar to the model for calls for service, the intervention is associated with a statistically significant 12% lower rate of offenses (IRR $= .875, p = .006$; Table A25).

Finally, Figure A20 shows the overall trend in offenses involving juvenile suspects and/or victims. The monthly counts of offenses in both the treatment and comparison sites are very low, and the patterns are relatively similar. Figure A21 shows that juvenile offenses were considerably lower in both the treatment and comparison sites in the post-intervention period, with a percentage decrease of 52% in the treatment sites and 39% in the comparison sites. The statistical model is similar to those for calls for service and total offenses: the rate of juvenile offenses was 12% lower in the treatment sites during the intervention period, controlling for the other variables. However, for juvenile offenses this finding was not statistically significant (IRR $= .871, p = .376$; Table A26).

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

4.5 Community survey

The results of the analysis of community survey outcomes are displayed in Tables A30-A35. Not surprisingly given the lack of responses to the mail survey in Wave 2, there are very few significant effects and the coefficients in some of the models are unusually large or small due to the very small numbers in each category, so results should be interpreted with caution. The coefficient for the *treatment* $\times$ *time* interaction is small and positive, but non-significant, for the social cohesion and collective efficacy scales (Table A30), and the community engagement scale (Table A32), indicating slight improvements in these areas.

Other results are less encouraging. The odds ratio for the *treatment* $\times$ *time* interaction in the model for change in perceptions of neighborhood safety over the past five years (a three-point scale indicating that safety has gotten worse, stayed the same, or improved) is very small (.022) and statistically significant, indicating that perceptions of safety worsened in the treatment areas during the treatment period. The finding for perceived change in police protection over the same time period is also very small (.031), though not statistically significant (both models shown in Table A31. The unusually small sizes of these odds ratios reflect the discrepancy in sample sizes between Waves 1 and 2 and should be interpreted with caution. Perceptions of police effectiveness and legitimacy were also lower in the treatment sites in Wave 2, although again these were not statistically significant (Table A33). There was no change in perceived visibility of police in the community, but similar to the perceptions of neighborhood change, the odds ratio for satisfaction with police was very small (in the less favorable direction) and non-significant (Table A34).[14] Finally, Table A35 shows that residents in the treatment areas felt less safe and perceived higher levels of disorder in Wave 2, although again these findings are not statistically significant.

5 Implications for Policy and Practice

Young people in historically marginalized communities like Rainier Beach are at risk of academic failure, dropout, juvenile justice system involvement, and victimization due to a range of risk factors that stem from both the schools they attend and the communities in which they live. School-based factors and neighborhood context interact to create unsafe school environments that exacerbate structural challenges like social disorganization, poverty, and crime. These problems are especially profound for stu-

[14] In Wave 1, we only asked the policing questions, with the exception of satisfaction with police, to respondents who said they had contact with the police in the past year. We did not limit responses in Wave 2, but the overall N in these models is much lower than the other survey models.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

dents of color, who are substantially more likely to be disciplined in school and arrested and formally processed through the juvenile justice system, even when controlling for actual involvement in disruptive behavior and crime. In our current social context, governments, social institutions, and communities are increasingly called upon to recognize and mitigate racial inequity and ensure that responses to problematic behaviors address underlying risk factors without inadvertently targeting vulnerable populations. The RBCSC sought to address these issues by building on existing evidence-informed practices and combining promising approaches from a range of community settings to improve the adult-run systems that impact outcomes for young people. PBIS improves school climate and other academic and social outcomes that can be risk or protective factors for youth. The addition of RJ offered an opportunity to engage in supportive conflict resolution and alternatives to punitive discipline. While school-based RJ has not been as rigorously tested as PBIS, research suggests that it may reduce disciplinary referrals, improve school climate, and address racial disparities in discipline. The combination of PBIS and RJ in schools is an emerging practice, and the extension of these practices to community settings is an innovation of this study.

As we documented in the Methodology section, the implementation and evaluation of the RBCSC was beset by challenges from the start. We experienced serious setbacks in the first year of the project due to the lack of buy-in from established community organizations. While we were able to come to some agreements about ensuring broad representation in the initiative (through our own restorative justice processes), these relational tensions persisted throughout the grant period and hindered access and progress at times. We also experienced funding challenges, including a 6-month delay in receiving grant funds after the project start date and a period during which the grant expired because our extension request, although approved, could not be processed. The latter delay in particular further increased tensions and reduced morale on the team, as well as leading to layoffs among the most economically vulnerable members of the initiative. The COVID-19 pandemic, which could not have been predicted at the outset, hit the project especially hard due to its location (the same county where the United States' first COVID-19 case was identified), the focus on school and community activities, all of which were shut down during the pandemic, and its focus on the most marginalized and vulnerable populations, who were the hardest hit by COVID. Even with multiple extensions, it was understandably difficult for some of the participating organizations to bounce back and focus on implementation, especially on the school side (although, as we discuss below, the initial capacity-building efforts of the project benefited the com-

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

munity in unexpected ways). As we have described throughout this report, the pandemic also irretrievably impacted both our ability to collect reliable data and the reliability of administrative data on crime and school outcomes that was already routinely collected.

A further important challenge was that the program was extremely ambitious, both in terms of scope and innovation and in its attempts to bring together a variety of community organizations and government institutions that traditionally work in silos. The planning process took much longer than anticipated and implementation was somewhat piecemeal. For example, SPS was already interested in developing a PBIS approach prior to the project start date, and was already moving ahead with training and some degree of implementation in both treatment *and* non-treatment schools before the community piece, which had to be developed from scratch, was ready to go. As a result, students in our comparison schools may also have been exposed to PBIS, although they did not receive the RJ component or experience any CW-PBIS or CW-RJ development in their neighborhoods. On the other hand, the schools had varying levels of capacity to implement SW-PBIS or RJ, despite support from SPS and the grant partners. Some schools in the program (e.g. ES5 and MS5) began implementing PBIS at Tier II and even III by the end of the project period, while others (such as MS1 and AS5) never reached full implementation at Tier I. From an evaluation standpoint, this piecemeal implementation, extended planning period, and lack of communication/tension with some partners who had reservations about the process by which this project was brought to the table led to delays in beginning data collection, which were then compounded by the pandemic. In hindsight we should have started collecting process data and conducting ongoing outcome analyses earlier and accounted for the variability in start dates, rather than waiting until we believed all of the partners were ready to go. However, we were able to supplement some of the data collection efforts; for example, with school climate survey data collected routinely by SPS (although, as we have described, these data are affected by the pandemic and other changes in data collection protocols, since they were not originally collected for the purpose of this project).

Nonetheless, our project description also shows that a great deal of progress was made, especially in the community setting. The CW-PBIS implementation team, led by RBAC, worked closely with the PBIS training and technical assistance advisors from SPS to translate the principles of SW-PBIS into a community setting. They developed data collection instruments and processes similar to those developed for schools, which mirrored the SW-PBIS tiers of support and integrated the CW-RJ and other pre-existing ABSPY activities. For example, the adoption of the "Be^3" principles formed the basis of Tier I CW-PBIS, and shared norms and expectations around these principles were communicated through public art and messaging campaigns. At Tier II, ABSPY's Safe Passage team, which watches over and assists school-aged children in getting to and from school safely, is an example of an intervention that assists a smaller group of community members who need more support. Finally, at Tier III, community peacemaking circles led by the CW-RJ team provided individualized support and services for young people involved in conflict or affected by community crime.

The CW-PBIS team also provided walkthroughs for local residents and businesses to help them prevent crime (e.g. through Crime Prevention Through Environmental Design, or CPTED, assessments) and support their clientele in following the "Be^3" values. The work of the CW-PBIS team culminated in the development of a CW-PBIS Handbook, which documents the development and implementation of CW-PBIS and contains a wealth of examples and resources to help other communities implement a similar program. Similarly, the CW-RJ team hired a group of young people in the community and trained them in the peace circle-keeping process while also providing them with a hugely important opportunity for economic support and leadership development. They have conducted circles with hundreds of young people and other community members, both virtually and in-person, to provide emotional support, conflict resolution, and a space to process the trauma of the pandemic and instances of violent crime in the community.

The variety of challenges documented above have limited our ability to draw firm conclusions about the effectiveness of this initiative or provide answers to all of our research questions. Furthermore, the focus on Rainier Beach allowed us to capitalize on an existing foundation of community organizing but prevented us from implementing a more rigorous evaluation design. As a result, our findings are mixed. The school climate survey results were generally not favorable—where there were significant effects, they were generally in the wrong direction, especially for high school students. However, because data collection stopped during the pandemic and then the survey was substantially revised for post-pandemic implementation, we could only look at data up to the 2019 school year. We aimed to control for this by looking at the effects of the program by year relative to 2016, when some schools began thinking about and training on PBIS, but it is unlikely that our analysis could have picked up the staggered implementation that was occurring across the time period. We find similarly mixed results for the school administrative data, although the program was associated with some statistically significant improvements in excused absences and unexcused absences as a proportion of eligible attendance days. For

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

elementary school students, there was also a significant and substantial reduction in disciplinary actions. However, as we noted in the Findings section, these results are likely skewed due to the pandemic (for example, no disciplinary actions were recorded while children were learning remotely). Our analysis also does not control for individual academic years, which would have allowed us to model this slightly more accurately. Further, we were not able to calculate results for middle schools on any of the school administrative data points. Importantly, an exploratory analysis of program effects stratified by race showed that students of color experienced the program differently from White students on each of the significant outcomes. These experiences were sometimes, though by no means always, positive. For example, in elementary schools Hispanic and Native American/Pacific Islander students earned significantly higher test scores than White students, even though test scores decreased overall. We found similar results for Black, Asian, and multiracial students in high schools. However, multiracial and Native American/Pacific Islander high school students did not benefit from the reduction in excused absences in the same way as White students.

Our community survey was severely hampered by a pivot to a mail survey format for the follow-up wave, which had a very low rate of engagement compared to our baseline survey, which involved door-to-door in-person data collection. As a result, we found very few statistically significant effects, and even significant effects should be interpreted with caution. Respondents in the follow-up survey perceived that neighborhood safety had gotten worse in the past 5 years, they felt less safe, and perceived more disorder. In addition to the low response rate, it is important to remember that the two survey waves were conducted pre- and post-pandemic, which may also be driving these findings. Nonetheless, our most promising findings come from the analysis of crime incidents and calls for police service. The results show that the initiative is associated with modest but statistically significant reductions in police calls for service and recorded offenses in the areas immediately surrounding the treatment schools. We see similar decreases in offenses involving juvenile suspects or victims. While this finding was not statistically significant, it is possible that the small number of juvenile-involved offenses in these small buffer zones limited our ability to detect a significant difference.

Considering these results in light of the program implementation and challenges, it appears that our strongest positive findings align with the aspects of the program that were most consistently implemented and were able to persist, albeit in modified form, during the COVID-19 pandemic. It is important to stress that the SW-PBIS and SW-RJ initiatives are in no way an implementation failure—many of the

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

schools did very meaningful work in this space, and both school staff and SPS central staff worked extremely hard to maintain aspects of the program as much as possible during and after the pandemic, during a period where there was very little time to focus on new initiatives and personnel were being pulled in all directions to cover for illness and understaffing (for example, at one point after schools reopened one of our PBIS technical assistance providers had to be pulled back to the central office to answer phones in order to document staff and student absences). Although it was not fully implemented by the end of the project, the Seattle TFI is a radical reimagining of established PBIS implementation tracking documents that places anti-racism and culturally competent programming at the heart of student support services. Given the variations we see in the effect of the school-based program by race, this is a valuable and much-needed development. Furthermore, PBIS is an established evidence-based practice in schools, so it is most likely that our data could not adequately capture its impact, given the piecemeal implementation (including some schools that started implementation before our project formally began, and the development of PBIS in some non-treatment schools) and our reliance on administrative datasets that, in addition to being affected by the pandemic, were not designed with the specific goal of capturing PBIS or SW-RJ implementation or outcomes. In short, our data collection and analysis may have been too far removed from the reality of implementation of SW-PBIS-RJ to fully capture its effects.

On the other hand, while they took a long time to develop, the CW-PBIS-RJ elements of the program ultimately addressed neighborhood crime and public safety at a much more direct level. The CW-PBIS initiative focused specifically on providing awareness, training, and direct support for improving public safety at specific neighborhood sites, many of which fell directly within the buffer areas we selected for our crime analysis. They drew on established crime prevention practices like CPTED, which provided concrete recommendations for businesses and other community establishments to secure their spaces, while also providing training, support, and follow-up on how to maintain a positive and welcoming environment for young people within these improved areas. The CW-RJ Peace Circles provided immediate opportunities to debrief, process, grieve, and heal after incidents of community violence, with the Circle Keepers and implementation team essentially serving as "first responders" for the community when traumatic events occurred. All of these activities took place against a backdrop of, and often directly incorporated, a well-established and effective youth crime prevention program (ABSPY) that has been operating in the community for a decade. Ultimately, there is a lot of overlap here that makes it difficult to parse out the specific effects of the RBCSC, but the positive results for crime prevention point to the fact

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

that there were clear, direct, and ongoing crime prevention efforts occurring during the implementation period of this project.

Reflecting on these findings and the implementation challenges, we offer the following recommendations for policy and practice to avoid some of the challenges we experienced:

1. Community crime prevention is complex, and the addition of governmental institutions adds more layers of complexity that make program development and implementation difficult. While it is not always possible within the timeline of grant-funded research, long planning periods (upwards of two years) are likely needed to parse out the various activities that will be conducted and who will be responsible for which areas. Sustainability planning also needs to be part of the conversation from the beginning, especially given the level of staff turnover in these types of organizations. Who will be responsible for implementation if the original person leaves?
2. Related to the above, it is crucial to start with honest and open dialogue about who is, is not, and should be at the table. As we have described, our failure to have this conversation prior to applying for the grant damaged relationships and caused long-term issues with trust and morale. This is particularly important when working with members of communities who have historically been marginalized, oppressed, and 'over-researched but under-served.' Engaging in our own restorative circles with external facilitation helped us have difficult conversations and eventually gain consensus and healing around some of the concerns.
3. On the other hand, it is important that a willing and capable leader is identified when implementing such a complex project, even if the community coalition is inclusive and highly democratic. It is important that everyone's voice is heard, and ultimately someone has to be responsible for making the final decision about how the work will be done. Thus, the leader has to have the support and legitimacy of the whole group, and there needs to be a clear decision-making process. For example, do final decisions need to be unanimous? What is the timeline by which a decision needs to be made, that also provides sufficient time for everyone to gather the information they need? Should some parties' decisions be weighted more heavily than others to account for historic marginalization? One of the reasons it was challenging to align implementation was because of the very different bureaucratic structures of the school side and the community side. SPS, as a large governmental bureaucracy, has a clear leadership hierarchy and established decision-making process,

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

as do most if not all school districts. PBIS was originally designed to fit within such a framework. When a variety of community groups come together there is a strong sense of shared values but no clear leader—no group was considered to be more important than another in this initiative. This made the community side of the program more authentic and responsive to the needs of the neighborhood, but also harder to implement and track.

4. We strongly encourage communities to work with local research partners. CEBCP has a longstanding relationship with the City of Seattle and the Rainier Beach community. However, the pandemic made it very clear how difficult it is to conduct research from a distance. While being local would not have prevented issues like school closures and changes in data collection procedures, we likely could still have been more responsive to the changing nature of the program during this time (and, even before the pandemic, to the varying start dates of different elements of implementation). As it was, even with a great deal of resources for travel, we were not on the ground regularly enough to see the intervention unfold, and once the pandemic started we were obviously unable to travel at all. From an evaluation perspective, complex interventions need an embedded research team to keep the evaluation design and data collection on track. Furthermore, this was an innovative intervention that included education and community settings. Future researchers in this area might consider a multi-disciplinary team that also includes education researchers with knowledge of PBIS and RJ.

This study is the first of its kind to provide evidence that PBIS and RJ can successfully be extended from schools to communities within a crime prevention framework and show significant crime prevention benefits. More work is needed, using a more rigorous research design and a wider variety of reliable data points, to more precisely establish the mechanisms by which this process is most successful and the conditions under which other communities could make a similar approach work. Nonetheless, while it was not always easy, the RBCSC shows that a coalition of community organizations and government institutions can, with the right preparation, work together to look within themselves and assess how to change the adult-run systems in a neighborhood for the benefit of young people.

Importantly, whether or not it is ultimately reflected in the data, the implementation of this program has been incredibly transformative for the Rainier Beach community, often in ways we did not anticipate at the outset of the project. In particular, as we have described and our culminating video directly illus-

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

trates, initiatives like the community Peace Circles (CW-RJ) have helped residents respond to and heal from traumatic events and serious crime incidents. The identification of shared values and the community connections that arose from them allowed residents and local organizations to mobilize during the pandemic, enabling them to get food and supplies to isolated residents and advocate for resources like mobile COVID-19 testing and vaccination clinics to be set up in the neighborhood. There have been challenges throughout the process, including (beyond the effects of the pandemic) the lack of obvious leadership or structure/hierarchy in the community compared to a school district bureaucracy. This can present difficulties when multiple organizations and private entities come together with their own interests and concerns. On the other hand, our findings do bear out the idea that there is a role for community ownership in such efforts and that it could be protective for communities and young people, at least from a crime standpoint, even in the face of massive social upheaval. An important question is whether such an initiative could be translated to a different community that does not have such a robust history of community organizing. Nonetheless, the hope and empowerment that it has provided to the residents of Rainier Beach is hard to deny.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Center for Evidence-Based Crime Policy
Department of Criminology, Law & Society

Rainier Beach
Campus Safety Continuum

Final Report

Appendices

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Appendix A: Tables & Figures

Tables

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A1: Descriptive statistics for school climate survey outcomes

		2016 (Pre)		2017 (Pre)		2018 (Pre)		2019 (During)	
	α (Items)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
Healthy community[a]	.771 (5)	5,194	3.28 (.73)	5,217	3.22 (.75)	4,549	3.20 (.74)	4,660	3.12 (.79)
Belonging[a]	.755 (6)	5,194	3.67 (.72)	5,216	3.60 (.70)	4,546	3.62 (.71)	4,655	3.58 (.71)
Classroom environment[a]	.825 (4)	5,192	3.15 (.79)	5,207	3.13 (.81)	4,499	3.13 (.82)	4,640	3.06 (.84)
School safety[a]	.793 (5)	5,189	3.26 (.79)	5,208	3.23 (.82)	4,516	3.25 (.83)	4,641	3.18 (.84)
Motivation & inclusion[a]	.842 (5)	5,175	3.57 (.84)	5,197	3.54 (.83)	4,503	3.53 (.85)	4,627	3.47 (.85)
Pedagogical effectiveness[a]	.864 (6)	5,193	3.69 (.78)	5,188	3.66 (.78)	4,523	3.66 (.77)	4,624	3.61 (.79)
Learning mindset[b]	.837 (6)	5,081	3.89 (.73)	5,128	3.83 (.76)	4,472	3.82 (.76)	4,538	3.75 (.76)
Social-emotional learning[b]	.801 (6)	5,067	3.86 (.74)	5,116	3.78 (.77)	4,457	3.81 (.76)	4,539	3.76 (.75)

[a] Outcomes based on a 5-point agreement scale (1 = strongly disagree, 3 = neither agree nor disagree, 5 = strongly agree).
[b] Outcomes based on a 5-point scale (1 = not like me at all, 3 = somewhat like me, 5 = very much like me).

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A2: Baseline (2016) sample characteristics by group (school climate survey data, elementary schools)

	Comparison (%) N = 441	Treatment (%) N = 534	χ^2 *(p)*
Gender			
Female	52.1	49.0	1.390 (.499)
Male	43.2	44.9	
Prefer not to state	4.7	6.0	
Race/Ethnic origin			
Black/African-American	28.0	32.0	89.350*** (<.0001)
Asian-American	10.6	26.4	
White	25.0	5.9	
Hispanic/Latinx	7.9	10.8	
Native American	2.7	3.0	
Pacific Islander	2.5	2.6	
Multiracial	23.3	19.3	
How often is English spoken at home?			
Rarely or never	2.3	6.4	29.960*** (<.0001)
Sometimes	20.3	25.1	
Most of the time	22.9	30.6	
Always	54.4	37.9	
Self-reported grades			
Not very good (Ds/Es)	2.4	4.1	9.921* (.019)
Some good/some not (Cs)	6.2	11.1	
Good (Bs)	41.9	43.7	
High/Very high (As)	49.5	41.1	

Significant differences between treatment and comparison groups at baseline: * $p<.05$, ** $p<.01$, *** $p<.001$

Note: Students could not be uniquely identified in the dataset. There may be multiple observations per student.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A3: Baseline (2016) sample characteristics by group (school climate survey data, middle schools)

	Comparison (%) N = 1,824	Treatment (%) N = 669	χ^2 (p)
Gender			
Female	47.3	47.4	.425 (.809)
Male	49.0	48.4	
Prefer not to state	3.6	4.2	
Race/Ethnic origin			
Black/African-American	21.0	27.6	222.353*** (<.0001)
Asian-American	16.6	35.0	
White	29.3	4.8	
Hispanic/Latinx	12.4	11.4	
Native American	0.7	0.5	
Pacific Islander	2.2	6.3	
Multiracial	17.9	14.4	
How often is English spoken at home?			
Rarely or never	3.8	5.7	95.416*** (<.0001)
Sometimes	16.4	25.8	
Most of the time	25.5	36.6	
Always	54.3	31.9	
Self-reported grades			
Not very good (Ds/Es)	2.7	3.3	25.669*** (<.0001)
Some good/some not (Cs)	13.0	5.7	
Good (Bs)	39.6	41.2	
High/Very high (As)	44.8	49.8	

Significant differences between treatment and comparison groups at baseline: * p<.05, ** p<.01, *** p<.001

Note: Students could not be uniquely identified in the dataset. There may be multiple observations per student.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A4: Baseline (2016) sample characteristics by group (school climate survey data, high schools)

	Comparison (%) N = 1,639	Treatment (%) N = 115	χ^2 (*p*)
Gender			
Female	47.6	45.8	.553 (.758)
Male	48.5	51.4	
Prefer not to state	3.9	2.8	
Race/Ethnic origin			
Black/African-American	21.5	49.5	54.026*** (<.0001)
Asian-American	33.4	16.5	
White	14.7	1.0	
Hispanic/Latinx	12.3	10.3	
Native American	1.5	1.0	
Pacific Islander	4.8	8.2	
Multiracial	11.8	13.4	
How often is English spoken at home?			
Rarely or never	8.8	5.6	2.339 (.505)
Sometimes	21.3	25.0	
Most of the time	29.0	25.9	
Always	40.9	43.5	
Self-reported grades			
Not very good (Ds/Es)	6.7	9.5	1.749 (.626)
Some good/some not (Cs)	19.3	17.1	
Good (Bs)	34.8	37.1	
High/Very high (As)	39.2	36.2	

Significant differences between treatment and comparison groups at baseline: * p<.05, ** p<.01, *** p<.001

Note: Students could not be uniquely identified in the dataset. There may be multiple observations per student.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A5: Multilevel mixed effects linear regression on healthy community scale, by wave

	Elementary schools	Middle schools	High schools
Fixed effects	*b* (Robust SE)	*b* (Robust SE)	*b* (Robust SE)
Wave (ref:2016)			
2017	-.041 (.076)	-.070 (.081)	-.065*** (.005)
2018	.049 (.185)	-.169 (.090)	-.069*** (.012)
2019	-.090 (.248)	-.343*** (.049)	-.067* (.033)
Treatment	.002 (.116)	.113 (.139)	-.019 (.166)
2017 $\times$ Treatment	-.080 (.103)	-.095 (.106)	.052 (.084)
2018 $\times$ Treatment	-.214 (.253)	.084 (.169)	-.121* (.047)
2019 $\times$ Treatment	-.049 (.255)	.094 (.151)	-.195* (.095)
Gender (ref:Female)			
Male	-.090** (.029)	.076* (.036)	.075* (.033)
Prefer not to state	-.361*** (.066)	-.291** (.111)	-.281*** (.065)
Race (ref:Black/African-American)			
Asian-American	.035 (.083)	.024 (.025)	-.031 (.024)
White	.016 (.045)	-.062 (.042)	-.026 (.026)
Hispanic/Latinx	.147* (.064)	.092*** (.020)	.041 (.028)
Native American	.176 (.090)	-.078 (.070)	.109* (.047)
Pacific Islander	.152 (.113)	.111 (.059)	.060 (.044)
Multiracial	.036 (.046)	-.042* (.017)	-.102*** (.022)
English spoken at home (ref:Always)			
Rarely or never	.133 (.070)	.044 (.037)	.113*** (.034)
Sometimes	.100* (.040)	.058 (.035)	.122** (.038)
Most of the time	.045 (.041)	.012 (.015)	.029 (.035)
Self-reported grades(ref:High/Very High)			
Not very good (Ds/Es)	-.351*** (.061)	-.513*** (.047)	-.501*** (.065)
Some good/some not (Cs)	-.112* (.053)	-.251*** (.034)	-.219*** (.009)
Good (Bs)	-.052* (.022)	-.092*** (.012)	-.127*** (.011)
Constant	3.532*** (.081)	3.217*** (.106)	3.458*** (.168)
Random effects	σ (Robust SE)	σ (Robust SE)	σ (Robust SE)
School	.058 (.017)	.113 (.033)	.201 (.062)
Residual	.778 (.018)	.703 (.015)	.680 (.009)
Log pseudolikelihood	-3654.435	-7676.037	-6778.070
Wald χ^2	-	-	-
N	3127	7193	6549

Multilevel mixed effects linear regression
$*p < .05; **p < .01; ***p < .001$

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A6: Multilevel mixed effects linear regression on belonging scale, by wave

	Elementary schools	Middle schools	High schools
Fixed effects	*b* (Robust SE)	*b* (Robust SE)	*b* (Robust SE)
Wave (ref:2016)			
2017	-.106 (.070)	-.036* (.016)	-.065*** (.013)
2018	.042 (.165)	-.044 (.031)	-.052** (.017)
2019	-.046 (.202)	-.146*** (.016)	-.016 (.021)
Treatment	.064 (.063)	-.024 (.075)	-.049 (.133)
2017 $\times$ Treatment	-.058 (.081)	-.036 (.067)	.028 (.079)
2018 $\times$ Treatment	-.195 (.192)	.019 (.046)	-.145 (.076)
2019 $\times$ Treatment	-.053 (.204)	.045 (.067)	-.140 (.090)
Gender (ref:Female)			
Male	-.084** (.032)	-.006 (.027)	.007 (.010)
Prefer not to state	-.353*** (.068)	-.420*** (.046)	-.342*** (.071)
Race (ref:Black/African-American)			
Asian-American	.015 (.054)	.003 (.010)	.080** (.028)
White	.063 (.041)	.042 (.053)	.162*** (.030)
Hispanic/Latinx	.166*** (.035)	.122*** (.023)	.116** (.036)
Native American	.163** (.053)	-.102 (.058)	.121*** (.021)
Pacific Islander	.076 (.087)	.090* (.045)	.156** (.056)
Multiracial	.073 (.041)	.001 (.032)	-.000 (.029)
English spoken at home (ref:Always)			
Rarely or never	-.071* (.035)	-.058 (.041)	-.047 (.045)
Sometimes	.048 (.041)	-.051 (.035)	-.002 (.038)
Most of the time	.028 (.039)	-.034 (.022)	-.049* (.024)
Self-reported grades(ref:High/Very High)			
Not very good (Ds/Es)	-.317*** (.082)	-.523*** (.084)	-.457*** (.067)
Some good/some not (Cs)	-.152*** (.044)	-.256*** (.040)	-.223*** (.021)
Good (Bs)	-.058* (.027)	-.134*** (.013)	-.145*** (.010)
Constant	3.927*** (.043)	3.761*** (.053)	3.775*** (.096)
Random effects	σ (Robust SE)	σ (Robust SE)	σ (Robust SE)
School	.103 (.021)	.069 (.020)	.130 (.040)
Residual	.693 (.016)	.674 (.016)	.645 (.011)
Log pseudolikelihood	-3298.691	-7371.435	-6433.238
Wald χ^2	-	-	-
N	3126	7188	6551

Multilevel mixed effects linear regression
$*p < .05; **p < .01; ***p < .001$

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A7: Multilevel mixed effects linear regression on classroom environment scale, by wave

	Elementary schools	Middle schools	High schools
Fixed effects	*b* (Robust SE)	*b* (Robust SE)	*b* (Robust SE)
Wave (ref:2016)			
2017	.020 (.053)	-.061 (.060)	-.010 (.034)
2018	.190 (.204)	-.079 (.082)	-.053*** (.011)
2019	.049 (.250)	-.199*** (.040)	-.034 (.043)
Treatment	-.055 (.115)	.055 (.120)	-.051 (.097)
2017 $\times$ Treatment	-.025 (.084)	-.062 (.097)	.003 (.130)
2018 $\times$ Treatment	-.263 (.270)	.005 (.159)	-.187 (.107)
2019 $\times$ Treatment	-.071 (.258)	-.009 (.088)	-.169 (.153)
Gender (ref:Female)			
Male	.038 (.027)	.095*** (.026)	.039 (.036)
Prefer not to state	-.168* (.083)	-.199* (.087)	-.272*** (.047)
Race (ref:Black/African-American)			
Asian-American	-.067 (.090)	.041 (.027)	-.012 (.032)
White	-.059 (.046)	-.030 (.132)	-.102*** (.008)
Hispanic/Latinx	.042 (.064)	.030 (.016)	-.016 (.045)
Native American	.161* (.080)	-.145 (.106)	-.011 (.114)
Pacific Islander	.056 (.106)	.058 (.046)	.049 (.035)
Multiracial	-.023 (.038)	-.030 (.029)	-.148*** (.024)
English spoken at home (ref:Always)			
Rarely or never	.128 (.119)	-.016 (.077)	.035 (.025)
Sometimes	.172** (.064)	.044 (.046)	.055 (.042)
Most of the time	.087 (.059)	.038*** (.010)	.013 (.036)
Self-reported grades(ref:High/Very High)			
Not very good (Ds/Es)	-.346*** (.086)	-.402*** (.045)	-.395*** (.047)
Some good/some not (Cs)	-.103* (.044)	-.212*** (.059)	-.105*** (.013)
Good (Bs)	-.061 (.036)	-.077 (.043)	-.094*** (.025)
Constant	3.220*** (.108)	3.006*** (.118)	3.422*** (.093)
Random effects	σ (Robust SE)	σ (Robust SE)	σ (Robust SE)
School	.039 (.018)	.066 (.027)	.083 (.024)
Residual	.878 (.008)	.792 (.010)	.700 (.019)
Log pseudolikelihood	-4026.921	-8486.718	-6947.017
Wald χ^2	-	-	-
N	3123	7150	6538

Multilevel mixed effects linear regression
$*p < .05; **p < .01; ***p < .001$

Table A8: Multilevel mixed effects linear regression on school safety scale, by wave

	Elementary schools	Middle schools	High schools
Fixed effects	*b* (Robust SE)	*b* (Robust SE)	*b* (Robust SE)
Wave (ref:2016)			
2017	-.123* (.056)	-.032 (.115)	-.027 (.020)
2018	-.017 (.140)	-.059 (.101)	-.058 (.032)
2019	-.068 (.227)	-.196*** (.015)	-.037 (.028)
Treatment	-.087 (.068)	.073 (.134)	-.158 (.151)
2017 $\times$ Treatment	.004 (.091)	-.058 (.136)	.114 (.068)
2018 $\times$ Treatment	-.074 (.219)	.092 (.149)	.033 (.056)
2019 $\times$ Treatment	.038 (.242)	.075 (.169)	-.046 (.053)
Gender (ref:Female)			
Male	-.015 (.046)	.160*** (.029)	.137*** (.025)
Prefer not to state	-.349*** (.058)	-.288*** (.067)	-.344*** (.079)
Race (ref:Black/African-American)			
Asian-American	.027 (.085)	-.084** (.027)	-.108*** (.032)
White	.056 (.055)	-.174*** (.039)	-.099* (.050)
Hispanic/Latinx	.197*** (.039)	-.016 (.036)	-.018 (.037)
Native American	.275*** (.062)	.049 (.060)	.043 (.074)
Pacific Islander	.109 (.115)	.012 (.019)	.008 (.051)
Multiracial	.009 (.056)	-.088*** (.024)	-.152*** (.039)
English spoken at home (ref:Always)			
Rarely or never	.130 (.082)	.019 (.057)	.094 (.049)
Sometimes	.104* (.048)	.048 (.043)	.057** (.021)
Most of the time	.063 (.047)	.007 (.022)	.006 (.023)
Self-reported grades(ref:High/Very High)			
Not very good (Ds/Es)	-.334*** (.093)	-.456*** (.039)	-.415*** (.065)
Some good/some not (Cs)	-.014 (.046)	-.172*** (.025)	-.142*** (.015)
Good (Bs)	.010 (.039)	-.040 (.023)	-.088*** (.015)
Constant	3.526*** (.049)	3.184*** (.093)	3.493*** (.153)
Random effects	σ (Robust SE)	σ (Robust SE)	σ (Robust SE)
School	.092 (.019)	.127 (.036)	.201 (.056)
Residual	.882 (.011)	.779 (.018)	.693 (.010)
Log pseudolikelihood	-4045.725	-8384.381	-6886.016
Wald χ^2	-	-	-
N	3124	7164	6538

Multilevel mixed effects linear regression
$*p < .05; **p < .01; ***p < .001$

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A9: Multilevel mixed effects linear regression on motivation & inclusion scale, by wave

	Elementary schools	Middle schools	High schools
Fixed effects	*b* (Robust SE)	*b* (Robust SE)	*b* (Robust SE)
Wave (ref:2016)			
2017	-.075*** (.015)	.012 (.025)	-.058 (.037)
2018	.051 (.078)	-.072* (.030)	-.046* (.022)
2019	-.020 (.177)	-.200*** (.054)	-.064*** (.009)
Treatment	.015 (.058)	.032 (.129)	.014 (.161)
2017 $\times$ Treatment	-.032 (.057)	-.077 (.084)	.048 (.063)
2018 $\times$ Treatment	-.239 (.135)	.023 (.070)	-.079 (.049)
2019 $\times$ Treatment	-.078 (.196)	.044 (.119)	-.151*** (.027)
Gender (ref:Female)			
Male	-.078* (.036)	.011 (.035)	-.018 (.015)
Prefer not to state	-.380*** (.057)	-.469*** (.046)	-.338*** (.053)
Race (ref:Black/African-American)			
Asian-American	-.055 (.075)	-.061* (.029)	-.089* (.045)
White	-.139** (.045)	-.118*** (.025)	-.077 (.041)
Hispanic/Latinx	.132** (.047)	.001 (.044)	.020 (.059)
Native American	.072 (.121)	-.163 (.087)	.000 (.029)
Pacific Islander	.114 (.094)	.076* (.034)	.120** (.045)
Multiracial	.020 (.049)	-.073* (.037)	-.161*** (.028)
English spoken at home (ref:Always)			
Rarely or never	.180* (.090)	.051 (.071)	.117** (.038)
Sometimes	.098*** (.021)	.029 (.033)	.111*** (.029)
Most of the time	.027 (.041)	-.015 (.023)	.018 (.022)
Self-reported grades(ref:High/Very High)			
Not very good (Ds/Es)	-.474*** (.075)	-.715*** (.074)	-.618*** (.074)
Some good/some not (Cs)	-.179*** (.030)	-.309*** (.041)	-.295*** (.018)
Good (Bs)	-.067 (.035)	-.156*** (.022)	-.154*** (.011)
Constant	4.162*** (.058)	3.686*** (.092)	3.722*** (.157)
Random effects	σ (Robust SE)	σ (Robust SE)	σ (Robust SE)
School	.028 (.018)	.112 (.030)	.203 (.060)
Residual	.792 (.009)	.788 (.018)	.729 (.006)
Log pseudolikelihood	-3699.516	-8456.509	-7216.876
Wald χ^2	-	-	-
N	3120	7155	6531

Multilevel mixed effects linear regression
$*p < .05; **p < .01; ***p < .001$

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A10: Multilevel mixed effects linear regression on pedagogical effectiveness scale, by wave

	Elementary schools	Middle schools	High schools
Fixed effects	*b* (Robust SE)	*b* (Robust SE)	*b* (Robust SE)
Wave (ref:2016)			
2017	-.058 (.051)	.012 (.025)	-.061 (.045)
2018	.057 (.074)	-.072* (.030)	-.029 (.049)
2019	.017 (.157)	-.200*** (.054)	-.044 (.034)
Treatment	.108*** (.023)	.032 (.129)	.051 (.149)
2017 $\times$ Treatment	-.049 (.075)	-.077 (.084)	-.016 (.064)
2018 $\times$ Treatment	-.265* (.103)	.023 (.070)	-.133* (.058)
2019 $\times$ Treatment	-.083 (.164)	.044 (.119)	-.151** (.052)
Gender (ref:Female)			
Male	-.067* (.030)	.011 (.035)	-.035* (.014)
Prefer not to state	-.335*** (.052)	-.469*** (.046)	-.313*** (.040)
Race (ref:Black/African-American)			
Asian-American	-.100 (.067)	-.061* (.029)	.023 (.037)
White	-.120*** (.036)	-.118*** (.025)	-.074 (.043)
Hispanic/Latinx	.083** (.029)	.001 (.044)	.033 (.038)
Native American	.123** (.046)	-.163 (.087)	-.087** (.031)
Pacific Islander	.009 (.106)	.076* (.034)	.090* (.040)
Multiracial	.018 (.024)	-.073* (.037)	-.104* (.042)
English spoken at home (ref:Always)			
Rarely or never	.128 (.088)	.051 (.071)	.051 (.038)
Sometimes	.096*** (.017)	.029 (.033)	.059*** (.017)
Most of the time	.046 (.035)	-.015 (.023)	.005 (.009)
Self-reported grades(ref:High/Very High)			
Not very good (Ds/Es)	-.441*** (.094)	-.715*** (.074)	-.515*** (.058)
Some good/some not (Cs)	-.145*** (.041)	-.309*** (.041)	-.233*** (.020)
Good (Bs)	-.037 (.041)	-.156*** (.022)	-.137*** (.008)
Constant	4.112*** (.031)	3.686*** (.092)	3.836*** (.135)
Random effects	σ (Robust SE)	σ (Robust SE)	σ (Robust SE)
School	.061 (.015)	.112 (.030)	.185 (.051)
Residual	.751 (.018)	.788 (.018)	.688 (.008)
Log pseudolikelihood	-3543.331	-8456.509	-6849.827
Wald χ^2	-	-	-
N	3123	7155	6543

Multilevel mixed effects linear regression
$*p < .05; **p < .01; ***p < .001$

Table A11: Multilevel mixed effects linear regression on learning mindset scale, by wave

	Elementary schools	Middle schools	High schools
Fixed effects	*b* (Robust SE)	*b* (Robust SE)	*b* (Robust SE)
Wave (ref:2016)			
2017	-.121** (.042)	-.024 (.016)	-.076*** (.006)
2018	.048 (.036)	-.064* (.031)	-.086*** (.024)
2019	-.048 (.060)	-.178*** (.022)	-.105** (.032)
Treatment	-.061 (.045)	-.139*** (.021)	.069 (.065)
2017 × Treatment	.093 (.073)	-.013 (.018)	-.017 (.015)
2018 × Treatment	-.162* (.065)	.081** (.031)	-.184*** (.037)
2019 × Treatment	-.022 (.081)	.093 (.060)	-.097* (.046)
Gender (ref:Female)			
Male	-.015 (.031)	.027 (.053)	-.006 (.008)
Prefer not to state	-.225*** (.066)	-.314*** (.052)	-.284** (.092)
Race (ref:Black/African-American)			
Asian-American	-.120** (.038)	-.241*** (.024)	-.252*** (.024)
White	-.029 (.029)	-.192*** (.023)	-.175*** (.032)
Hispanic/Latinx	.054 (.050)	-.129** (.040)	-.083 (.069)
Native American	-.174 (.111)	-.329*** (.035)	-.188*** (.035)
Pacific Islander	.061 (.070)	-.109** (.039)	-.090*** (.026)
Multiracial	.023 (.029)	-.103* (.041)	-.169*** (.045)
English spoken at home (ref:Always)			
Rarely or never	.019 (.057)	-.009 (.018)	-.092 (.083)
Sometimes	.089* (.043)	.013 (.030)	-.014 (.023)
Most of the time	.052 (.028)	-.015 (.013)	-.038 (.030)
Self-reported grades(ref:High/Very High)			
Not very good (Ds/Es)	-.931*** (.142)	-1.068*** (.090)	-.899*** (.099)
Some good/some not (Cs)	-.539*** (.016)	-.655*** (.036)	-.575*** (.021)
Good (Bs)	-.272*** (.023)	-.320*** (.023)	-.306*** (.010)
Constant	4.210*** (.064)	4.302*** (.044)	4.311*** (.032)
Random effects	σ (Robust SE)	σ (Robust SE)	σ (Robust SE)
School	.033 (.016)	.000 (.)	.044 (.020)
Residual	.712 (.019)	.678 (.019)	.684 (.016)
Log pseudolikelihood	-3382.486	-7415.517	-6785.616
Wald χ^2	-	-	-
N	3133	7196	6522

Multilevel mixed effects linear regression
$*p < .05; **p < .01; ***p < .001$

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A12: Multilevel mixed effects linear regression on social-emotional learning scale, by wave

	Elementary schools	Middle schools	High schools
Fixed effects	*b* (Robust SE)	*b* (Robust SE)	*b* (Robust SE)
Wave (ref:2016)			
2017	-.136*** (.011)	-.018 (.016)	-.071* (.029)
2018	.041 (.030)	-.017 (.023)	-.038 (.020)
2019	-.023 (.074)	-.089*** (.025)	-.043 (.024)
Treatment	-.014 (.048)	-.014 (.032)	-.054 (.031)
2017 × Treatment	.023 (.036)	-.065 (.035)	-.009 (.037)
2018 × Treatment	-.228* (.104)	-.014 (.028)	-.080 (.079)
2019 × Treatment	-.151 (.090)	-.057 (.035)	-.041 (.026)
Gender (ref:Female)			
Male	-.175*** (.013)	-.063** (.020)	-.049 (.027)
Prefer not to state	-.286** (.089)	-.303*** (.041)	-.395*** (.054)
Race (ref:Black/African-American)			
Asian-American	.023 (.041)	-.012 (.046)	-.047 (.029)
White	.136*** (.041)	.058 (.051)	.100* (.044)
Hispanic/Latinx	.159* (.068)	.005 (.053)	-.016 (.042)
Native American	-.144 (.074)	-.132*** (.039)	-.143 (.075)
Pacific Islander	.036 (.115)	.032 (.060)	.080 (.049)
Multiracial	.048 (.034)	.010 (.024)	-.103* (.044)
English spoken at home (ref:Always)			
Rarely or never	.023 (.079)	.006 (.068)	-.009 (.025)
Sometimes	.112** (.039)	-.008 (.036)	.018 (.018)
Most of the time	.026 (.033)	-.005 (.024)	.030*** (.008)
Self-reported grades(ref:High/Very High)			
Not very good (Ds/Es)	-.727*** (.143)	-.782*** (.124)	-.589*** (.080)
Some good/some not (Cs)	-.314*** (.034)	-.436*** (.018)	-.343*** (.019)
Good (Bs)	-.159*** (.028)	-.198*** (.014)	-.160*** (.010)
Constant	4.068*** (.058)	4.028*** (.057)	4.067*** (.035)
Random effects	σ (Robust SE)	σ (Robust SE)	σ (Robust SE)
School	.030 (.010)	.032 (.029)	.000 (.)
Residual	.758 (.015)	.705 (.022)	.690 (.020)
Log pseudolikelihood	-3573.506	-7673.178	-6845.340
Wald χ^2	-	-	-
N	3129	7177	6533

Multilevel mixed effects linear regression
$*p < .05; **p < .01; ***p < .001$

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A13: Pre-intervention sample characteristics by group (school administrative data, elementary schools)

	Comparison (%) N = 1,209	Treatment (%) N = 1,400	χ^2 (p)
Gender			
Female	49.3	47.4	.907 (.341)
Male	50.7	52.6	
Race/Ethnicity			
Black	29.5	44.4	417.128*** (<.0001)
Asian	12.7	24.3	
Hispanic	11.2	16.6	
White	33.9	4.4	
Multiracial	11.7	9.1	
American Indian	0.3	0.4	
Pacific Islander	0.6	0.9	
Lives with both parents	68.9	60.7	15.257*** (<.0001)
Primary language is English	72.2	49.6	137.760*** (<.0001)
English spoken at home	71.9	51.1	117.742*** (<.0001)
Receiving English-Language Learner (ELL) services	24.6	37.3	48.737*** (<.0001)
Eligible for ELL services	24.3	37.3	50.759*** (<.0001)
Receiving/eligible for Special Education services	9.5	13.6	10.356*** (<.0001)
Receiving/eligible for Advanced Learner services	10.5	0.3	142.056*** (<.0001)

Significant differences between treatment and comparison groups at baseline: * *p*<.05, ** *p*<.01, *** *p*<.001

Note: Descriptive statistics based on first pre-intervention observation for students who had pre- and during-intervention observations.

Table A14: Pre-intervention sample characteristics by group (school administrative data, high schools)

	Comparison (%) N = 1,553	Treatment (%) N = 440	χ^2 *(p)*
Gender			
Female	47.0	46.6	.015 (.901)
Male	53.0	53.4	
Race/Ethnicity			
Black	26.0	41.0	108.367*** (<.0001)
Asian	29.0	29.5	
Hispanic	20.2	14.9	
White	14.8	2.3	
Multiracial	7.8	8.3	
American Indian	1.4	0.2	
Pacific Islander	0.7	3.8	
Lives with both parents	55.8	48.9	6.709** (.010)
Primary language is English	53.5	44.8	10.640*** (.001)
English spoken at home	52.8	44.6	9.449** (.002)
Receiving English-Language Learner (ELL) services	16.0	27.9	33.062*** (<.0001)
Eligible for ELL services	16.7	27.7	27.417*** (<.0001)
Receiving/eligible for Special Education Services	17.3	16.4	.166 (.684)
Receiving/eligible for Advanced Learner Services	8.1	2.0	20.545*** (<.0001)

Significant differences between treatment and comparison groups at baseline: * *p*<.05, ** *p*<.01, *** *p*<.001

Note: Descriptive statistics based on first pre-intervention observation for students who had pre- and during-intervention observations.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A15: Dependent variables for school administrative data analysis, pre- and during-implementation

	Elementary Schools		High Schools	
	N	Mean (SD)	N	Mean (SD)
Cumulative GPA[a]				
Pre	-	-	5,032	2.57 (1.03)
During	-	-	8,724	2.87 (.85)
State test score: English-Language Arts				
Pre	2,350	2,455.89 (97.32)	1,375	2,596.93 (120.25)
During	1,410	2,470.52 (103.96)	908	2,612.07 (120.82)
State test score: Mathematics				
Pre	2,355	2,469.38 (88.57)	909	2,597.93 (146.14)
During	1,402	2,464.94 (96.03)	867	2,572.61 (136.85)
Total absences				
Pre	6,525	8.07 (8.50)	5,004	27.91 (33.52)
During	4,672	9.09 (11.06)	8,575	29.86 (33.75)
Excused absences				
Pre	6,525	5.95 (6.70)	5,004	6.09 (7.98)
During	4,672	5.63 (7.61)	8,575	5.81 (8.67)
Unexcused absences				
Pre	6,525	2.13 (4.48)	5,004	21.83 (31.14)
During	4,672	3.45 (6.66)	8,575	24.05 (31.39)
Absences as proportion of eligible days				
Pre	6,525	.05 (0.06)	5,004	.11 (.19)
During	4,672	.06 (.08)	8,575	.14 (.22)
Tardies as proportion of eligible days				
Pre	6,525	.06 (.10)	5,004	.10 (.13)
During	4,672	.06 (.11)	8,575	.08 (.12)
Number of disciplinary actions[b]				
Pre	4,302	.02 (.23)	4,050	.07 (.41)
During	3,917	.01 (.11)	7,809	.03 (.24)

Numbers reflect all student records across the study period. There are multiple records per student.

[a] Cumulative GPA only recorded for middle and high school students.

[b] Count of disciplinary actions was only provided for students with at least one action. Students without a count were coded as 0, but we do not know if they represent true zeros or missing data.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A16: Propensity score-weighted multilevel mixed effects linear regression on academic achievement (elementary schools)

	State ELA score	State math score
Fixed effects	*b* (Robust SE)	*b* (Robust SE)
Treatment	-28.925* (14.197)	-9.909 (12.087)
During implementation	76.819*** (4.485)	62.860*** (2.259)
Treatment × During	-20.743* (8.545)	-35.029*** (2.927)
Constant	2438.995*** (10.588)	2443.052*** (3.320)
Random effects	σ (Robust SE)	σ (Robust SE)
School	14.491 (3.727)	14.370 (5.088)
Student	94.421 (4.468)	89.397 (4.601)
Residual	37.397 (1.123)	29.538 (1.904)
Log pseudolikelihood	-144081.872	-138623.227
Wald χ^2	800.780***	1316.081***
N	2447	2450

Propensity score-weighted multilevel mixed effects linear regression
$*p < .05; **p < .01; ***p < .001$

Table A17: Propensity score-weighted multilevel mixed effects linear regression on academic achievement (high schools)

	Cumulative GPA	State ELA score	State math score
Fixed effects	*b* (Robust SE)	*b* (Robust SE)	*b* (Robust SE)
Treatment	-.287 (.333)	-67.670 (62.768)	-83.877 (85.061)
During implementation	.004 (.025)	25.329** (9.564)	-13.779 (12.106)
Treatment × During	.005 (.025)	-33.968*** (9.698)	-10.567 (12.112)
Constant	2.779*** (.204)	2629.865*** (43.074)	2621.876*** (66.894)
Random effects	σ (Robust SE)	σ (Robust SE)	σ (Robust SE)
School	.300 (.048)	60.826 (10.106)	87.599 (17.917)
Student	.825 (.022)	-	-
Residual	.165 (.012)	107.431 (4.034)	120.328 (5.437)
Log pseudolikelihood	15613.749	-75145.136	-63208.313
Wald χ^2	189020.394***	47.961***	10302.188***
N	6473	1413	1090

Propensity score-weighted multilevel mixed effects linear regression
$*p < .05; **p < .01; ***p < .001$

Table A18: Propensity score-weighted multilevel mixed effects linear regression on absences (elementary schools)

	Total absences	Excused absences	Unexcused absences
Fixed effects	*b* (Robust SE)	*b* (Robust SE)	*b* (Robust SE)
Treatment	1.465 (.764)	-.391 (1.181)	1.860** (.721)
During implementation	-.339 (.280)	-.863 (.680)	.524 (.513)
Treatment $\times$ During	-.600 (.413)	-.095 (.764)	-.482 (.655)
Constant	8.128*** (.654)	6.654*** (1.119)	1.478** (.481)
Random effects	σ (Robust SE)	σ (Robust SE)	σ (Robust SE)
School	.845 (.201)	1.268 (.344)	.844 (.156)
Student	7.198 (.605)	5.325 (.960)	3.360 (.446)
Residual	4.572 (.493)	3.900 (.591)	2.257 (.362)
Log pseudolikelihood	-189554.691	-179230.256	-144628.966
Wald χ^2	13.468**	9.627*	7.689
N	7286	7286	7286

Propensity score-weighted multilevel mixed effects linear regression
$*p < .05; **p < .01; ***p < .001$

Table A19: Propensity score-weighted multilevel mixed effects linear regression on proportional absences (elementary schools)

	Unexcused/Total absence ratio	Absent/Eligible attendance days ratio	Tardy/Eligible attendance days ratio
Fixed effects	*b* (Robust SE)	*b* (Robust SE)	*b* (Robust SE)
Treatment	.153* (.075)	.008 (.005)	.031** (.012)
During implementation	.115* (.056)	.005 (.004)	-.002 (.004)
Treatment $\times$ During	-.100 (.063)	-.005 (.004)	.005 (.004)
Constant	.160** (.058)	.049*** (.004)	.048*** (.007)
Random effects	σ (Robust SE)	σ (Robust SE)	σ (Robust SE)
School	.087 (.016)	.007 (.001)	.014 (.004)
Student	.169 (.014)	.049 (.007)	.079 (.007)
Residual	.211 (.024)	.036 (.008)	.042 (.007)
Log pseudolikelihood	6344.022	117613.788	108218.978
Wald χ^2	8.686*	11.709**	9.760*
N	6731	7286	7286

Propensity score-weighted multilevel mixed effects linear regression
$*p < .05; **p < .01; ***p < .001$

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A20: Propensity score-weighted multilevel mixed effects linear regression on absences (high schools)

	Total absences	Excused absences	Unexcused absences
Fixed effects	*b* (Robust SE)	*b* (Robust SE)	*b* (Robust SE)
Treatment	-4.031 (16.227)	.026 (1.891)	-4.257 (14.515)
During implementation	8.144*** (1.538)	.445 (.228)	7.692*** (1.671)
Treatment × During	-1.360 (1.639)	-.825*** (.228)	-.568 (1.772)
Constant	30.045* (14.514)	6.177*** (1.440)	23.876 (13.206)
Random effects	σ (Robust SE)	σ (Robust SE)	σ (Robust SE)
School	19.023 (5.057)	1.992 (.492)	17.273 (4.711)
Student	17.684 (2.294)	5.193 (.485)	15.932 (2.690)
Residual	14.570 (4.450)	7.322 (3.049)	13.444 (4.332)
Log pseudolikelihood	-225388.076	-187155.961	-220984.073
Wald χ^2	659.553***	2246.542***	455.057***
N	6334	6334	6334

Propensity score-weighted multilevel mixed effects linear regression
$*p < .05; **p < .01; ***p < .001$

Table A21: Propensity score-weighted multilevel mixed effects linear regression on proportional absences (high schools)

	Unexcused/Total absence ratio	Absent/Eligible attendance days ratio	Tardy/Eligible attendance days ratio
Fixed effects	*b* (Robust SE)	*b* (Robust SE)	*b* (Robust SE)
Treatment	.078 (.102)	-.040 (.125)	.029 (.025)
During implementation	.138*** (.008)	.063** (.020)	-.008 (.010)
Treatment × During	-.026* (.012)	-.033 (.020)	-.007 (.010)
Constant	.614*** (.095)	.160 (.108)	.085*** (.020)
Random effects	σ (Robust SE)	σ (Robust SE)	σ (Robust SE)
School	.117 (.032)	.147 (.036)	.021 (.005)
Student	.208 (.010)	.094 (.021)	.078 (.008)
Residual	.201 (.011)	.096 (.022)	.060 (.006)
Log pseudolikelihood	7017.924	47710.260	72564.424
Wald χ^2	507.445***	141.189***	106612.645***
N	6219	6334	6334

Propensity score-weighted multilevel mixed effects linear regression
$*p < .05; **p < .01; ***p < .001$

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A22: Propensity score-weighted multilevel mixed effects negative binomial regression on count of disciplinary actions

	Elementary schools	High schools
Fixed effects	IRR (Robust SE)	IRR (Robust SE)
Treatment	35.439* (56.061)	1.407 (.425)
During implementation	2.577*** (.601)	.412*** (.086)
Treatment × During	.110*** (.062)	.768 (.170)
Constant	.001*** (.001)	.045*** (.014)
Dispersion parameter		
ln(Alpha)	4.000*** (.429)	2.932*** (.566)
Random effects	σ (Robust SE)	σ (Robust SE)
School	1.132 (.904)	.328 (.083)
Log pseudolikelihood	-1355.014	-6883.740
Wald chi2	24.56***	356.89***
N	6719	6517

Propensity score-weighted multilevel mixed effects negative binomial regression

Exponentiated coefficients (incidence rate ratio, IRR) - alpha not exponentiated

$*p < .05; **p < .01; ***p < .001$

Table A23: Mean monthly pre-intervention counts of calls for service and offenses

	Calls for service		All offenses		Juvenile offenses	
	Mean	SD	Mean	SD	Mean	SD
Treatment Sites						
MS1	28.1	8.6	7.6	3.4	0.8	0.9
ES3	14.0	4.2	5.7	3.0	0.2	0.4
ES4	18.9	6.0	7.1	3.4	0.9	1.3
ES/MS/AS/HS5	210.4	47.6	43.4	9.1	5.3	3.2
Comparison Sites						
MS/HS6	49.7	15.9	16.5	5.0	2.7	2.4
ES/HS7	109.2	28.9	34.4	8.2	2.6	2.3
AS8	58.9	12.3	21.5	4.5	0.7	1.0
MS11	11.1	7.2	5.0	3.0	0.4	0.6
MS12	86.3	17.6	24.3	6.7	1.7	1.7
ES13	22.3	8.3	6.5	2.9	0.3	0.6

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A24: Random effects negative binomial regression on calls for service

	Calls for service
Fixed effects	IRR (SE)
Interventions active	1.031 (.045)
Treatment areas	.926 (.097)
Active × Treatment	.887** (.036)
Month (ref:Jan)	
Feb	.886* (.042)
Mar	1.015 (.047)
Apr	.961 (.044)
May	1.108* (.050)
Jun	1.165*** (.051)
Jul	1.081 (.051)
Aug	.903* (.044)
Sep	.927 (.044)
Oct	1.000 (.045)
Nov	.910* (.042)
Dec	.897* (.042)
COVID-19 pandemic active	1.157** (.057)
COVID-19 school closures active	1.053 (.036)
Trend	1.000 (.001)
Autocorrelation controls	
1 month	1.384*** (.047)
2 months	1.209*** (.041)
3 months	1.178*** (.038)
Constant	.843 (.121)
Dispersion parameters	
ln_r	13.121 (6.401)
ln_s	51.360 (25.732)
Log likelihood	-3838.152
Wald χ^2	1510.493***
N	930

Random effects negative binomial regression
Exponentiated coefficients (incidence rate ratio, IRR)
$*p < .05; **p < .01; ***p < .001$

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A25: Random effects negative binomial regression on total offenses

	All offenses
Fixed effects	IRR (SE)
Interventions active	.905* (.044)
Treatment areas	1.032 (.191)
Active × Treatment	.875** (.043)
Month (ref:Jan)	
Feb	.919 (.052)
Mar	1.126* (.061)
Apr	1.064 (.058)
May	1.189** (.064)
Jun	1.102 (.060)
Jul	1.023 (.058)
Aug	1.056 (.059)
Sep	1.069 (.060)
Oct	1.078 (.060)
Nov	1.076 (.058)
Dec	.980 (.054)
COVID-19 pandemic active	.997 (.058)
COVID-19 school closures active	1.032 (.051)
Trend	1.000 (.001)
Autocorrelation controls	
1 month	1.194*** (.042)
2 months	1.145*** (.040)
3 months	1.151*** (.040)
4 months	1.110** (.038)
Constant	4.814*** (1.148)
Dispersion parameters	
ln_r	20.685 (10.598)
ln_s	12.247 (6.387)
Log likelihood	-2648.635
Wald χ^2	253.192***
N	920

Random effects negative binomial regression
Exponentiated coefficients (incidence rate ratio, IRR)
$*p < .05; **p < .01; ***p < .001$

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A26: Random effects negative binomial regression on offenses involving juveniles

	Juvenile offenses
Fixed effects	IRR (SE)
Interventions active	.826 (.128)
Treatment areas	1.656 (.568)
Active × Treatment	.871 (.136)
Month (ref:Jan)	
Feb	.946 (.169)
Mar	1.133 (.198)
Apr	1.085 (.189)
May	1.460* (.240)
Jun	1.110 (.193)
Jul	.795 (.155)
Aug	.768 (.146)
Sep	.931 (.169)
Oct	1.194 (.201)
Nov	.948 (.167)
Dec	.842 (.154)
COVID-19 pandemic active	.659* (.127)
COVID-19 school closures active	.833 (.154)
Trend	1.002 (.003)
Autocorrelation controls	
1 month	1.325*** (.093)
Constant	2.274** (.611)
Dispersion parameters	
ln_r	6.223 (3.033)
ln_s	2.055 (.920)
Log likelihood	-1214.002
Wald χ^2	114.343***
N	950

Random effects negative binomial regression
Exponentiated coefficients (incidence rate ratio, IRR)
$*p < .05; **p < .01; ***p < .001$

Table A27: Number of community surveys completed by site and wave

	Survey wave		
	1 (Pre)	2 (During)	Total
Treatment sites			
MS1	34	4	38
ES3	32	3	35
ES4	30	4	34
ES/MS/AS/HS5	125	11	136
Comparison sites			
MS/HS6	49	4	53
ES/HS7	56	3	59
AS8	25	6	31
ES11	28	2	30
MS12	25	2	27
ES13	28	3	31
Total	432	42	474

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A28: Survey respondent characteristics by group at baseline (wave 1)

	Comparison (%)	Treatment (%)	χ^2 *(p)*
Gender			
Female	58.1	54.3	1.838 (.399)
Male	40.6	45.2	
Other/Multiple	1.4	0.5	
Age			
18-25	12.7	15.1	1.070 (.983)
26-35	26.3	25.1	
36-45	21.6	22.4	
46-55	16.0	15.5	
56-65	8.9	7.3	
Over 75	5.2	5.9	
Race/ethnicity			
Black/African-American/ African immigrant	12.1	34.0	51.533*** (<.0001)
White	59.1	28.8	
Hispanic/Latinx	3.7	8.4	
Asian	9.3	14.9	
Other/Mixed/Multiple	15.8	14.0	
Length of time lived in neighborhood			
Less than 5 years	50.7	43.9	2.420 (.298)
5-10 years	20.3	20.8	
More than 10 years	29.0	35.3	
Has children			
	48.9	54.1	.859 (.354)
Child currently in school			
	47.7	69.2	7.345** (.007)
Born in US			
	81.9	62.7	20.068*** (<.0001)
Highest level of education			
Elementary/middle/some high school	1.9	7.9	40.319*** (<.0001)
High school diploma/GED	8.1	23.6	
Some college credit	16.1	18.5	
Associate's/Bachelor's degree	42.2	36.1	
Masters/graduate/professional degree	31.8	13.9	
Employment			
Full-time	51.4	50.0	6.898 (.141)
Part-time	18.7	14.8	
Not working	11.2	18.5	
Retired	15.9	15.7	
Other	2.8	0.9	

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Survey respondent characteristics by group at baseline (continued)

	Comparison (%)	Treatment (%)	χ^2 (*p*)
Attending school			
Part-time	4.3	7.4	1.924 (.382)
Full-time	6.2	6.5	
Home status			
Rent	43.1	56.4	7.431** (.006)
Own	56.9	43.6	
Household income			
<$20,000	8.0	20.3	37.778*** (<.0001)
$20,000-$34,999	11.5	18.8	
$35,000-$49,999	12.0	19.8	
$50,000-$74,999	15.0	14.2	
$75,000-$99,999	10.0	8.1	
$100,000+	43.5	18.8	
Contact with police in past year			
	33.0	32.4	.018 (.893)
Victim of crime in past year			
	33.0	24.2	4.050* (.044)

Significant differences between treatment and comparison group at baseline: * p<.05, ** p<.01, *** p<.001

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A29: Descriptive statistics for survey outcomes

		Wave 1		Wave 2	
	α (Items)	N	Mean (SD)	N	Mean (SD)
Social cohesion[a]	.796 (7)	433	2.96 (.46)	41	2.92 (.51)
Collective efficacy[b]	.905 (16)	427	2.84 (.51)	42	2.59 (.62)
Change in neighborhood safety[c]	-	366	2.38 (.64)	34	2.24 (.70)
Change in police protection[c]	-	312	2.26 (.58)	30	1.73 (.58)
Community engagement[d]	.710 (7)	422	.35 (.29)	41	.36 (.26)
Police visibility[e]	-	138	2.84 (.98)	42	2.33 (.90)
Police effectiveness[f]	.859 (3)	129	.69 (.41)	30	.47 (.48)
Police legitimacy[f]	.748 (2)	122	.58 (.45)	38	.28 (.43)
Satisfaction with police[g]	-	350	3.00 (.79)	36	2.47 (.74)
Feelings of safety[a]	.856 (6)	434	3.19 (.49)	42	3.05 (.55)
Perceptions of disorder[h]	.883 (12)	427	1.76 (.63)	41	2.21 (.65)

[a] Outcomes based on a 4-point agreement scale (1 = strongly disagree, 4 = strongly agree).
[b] Outcomes based on a 4-point likelihood scale (1 = very unlikely, 4 = very likely).
[c] Outcomes based on a 3-point scale (1 = gotten worse, 2 = stayed the same, 3 = much better).
[d] Outcomes based on a 2-point scale (0 = no, 1 = yes).
[e] Outcomes based on a 4-point likelihood scale (1 = very unlikely, 4 = very likely). In Wave 1 this question was only asked to people who had contact with police in past year.
[f] Outcomes based on a 2-point scale (0 = no, 1 = yes). In Wave 1 this question was only asked to people who had contact with police in past year.
[g] Outcomes based on a 4-point satisfaction scale (1 = very unsatisfied, 4 = very satisfied).
[h] Outcomes based on a 4-point scale (1 = not a problem, 4 = big problem).

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A30: Social cohesion and collective efficacy

	Social cohesion	Collective efficacy
	b (Robust SE)	b (Robust SE)
Wave 2	-.096 (.315)	-.081 (.446)
Treatment	-.009 (.095)	-.263* (.100)
Wave 2 × Treatment	.150 (.429)	.139 (.612)
Race (ref:Black/African-American/African Immigrant)		
White	.114 (.124)	.219 (.152)
Hispanic/Latinx	-.023 (.165)	.166 (.219)
Asian	-.023 (.147)	.060 (.193)
Other/Mixed/Multiple	-.181 (.122)	.119 (.143)
Child currently in school	.133 (.077)	.156 (.083)
Born in US	-.024 (.103)	.029 (.141)
Highest level of education (ref:Less than high school)		
High school diploma/GED	-.169 (.227)	-.267 (.206)
Some college credit	-.200 (.233)	-.256 (.263)
Associate's/Bachelor's degree	-.067 (.231)	-.252 (.271)
Masters/graduate/professional degree	.069 (.254)	-.234 (.277)
Home status (ref:Rent)		
Own	.096 (.095)	.148 (.151)
Household income (ref:<$20,000)		
$20,000-$34,999	.114 (.186)	.197 (.191)
$35,000-$49,999	.375 (.200)	.155 (.218)
$50,000-$74,999	.222 (.228)	.280 (.245)
$75,000-$99,999	.333 (.212)	.272 (.257)
$100,000+	.259 (.193)	.190 (.242)
Victim of crime in past year	-.034 (.080)	-.007 (.090)
Constant	2.786*** (.328)	2.723*** (.231)
F	2.90***	1.93*
R^2	.295	.237
RMSE	.436	.503
N	139	137

Linear regression with robust standard errors.
* p<.05, ** p<.01, *** p<.001

Table A31: Neighborhood changes in past 5 years

	Change in neighborhood safety	Change in police protection
	OR (Robust SE)	OR (Robust SE)
Wave 2	1.877 (2.293)	.276 (.362)
Treatment	.621 (.278)	.772 (.437)
Wave 2 $\times$ Treatment	.022* (.037)	.031 (.059)
Race (ref:Black/African-American/African Immigrant)		
White	.286 (.218)	.283 (.241)
Hispanic/Latinx	3.849 (4.088)	.338 (.388)
Asian	.199* (.126)	.153* (.132)
Other/Mixed/Multiple	.517 (.386)	.190* (.144)
Child currently in school	1.987 (.820)	.344* (.172)
Born in US	1.041 (.618)	.458 (.310)
Home status (ref:Rent)		
Own	.572 (.393)	.954 (.761)
Household income (ref:<$20,000)		
$20,000-$34,999	3.506 (2.461)	.602 (.538)
$35,000-$49,999	1.944 (1.456)	.252 (.255)
$50,000-$74,999	2.615 (2.234)	.339 (.373)
$75,000-$99,999	8.178* (7.583)	.129 (.141)
$100,000+	4.325 (3.286)	.130* (.123)
Victim of crime in past year	.527 (.247)	.827 (.377)
cut1	*.054** (.049)*	*.001*** (.001)*
cut2	*.822 (.652)*	*.056* (.070)*
Log pseudolikelihood	-98.609	-74.488
Pseudo R^2	.154	.232
Wald χ^2	44.477***	38.031**
N	125	106

Control for education omitted due to collinearity
Ordered logistic regression with robust standard errors
Exponentiated coefficients (odds ratios)
* p<.05, ** p<.01, *** p<.001

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A32: Community engagement

	Community engagement
	b (Robust SE)
Wave 2	-.272*** (.061)
Treatment	.001 (.054)
Wave 2 × Treatment	.305 (.207)
Race (ref:Black/African-American/African Immigrant)	
White	.034 (.084)
Hispanic/Latinx	-.077 (.103)
Asian	.093 (.085)
Other/Mixed/Multiple	.147* (.069)
Child currently in school	.102* (.046)
Born in US	-.004 (.055)
Highest level of education (ref:Less than high school)	
High school diploma/GED	.061 (.125)
Some college credit	.049 (.127)
Associate's/Bachelor's degree	.055 (.137)
Masters/graduate/professional degree	.208 (.139)
Home status (ref:Rent)	
Own	.206*** (.058)
Household income (ref:<$20,000)	
$20,000-$34,999	-.005 (.082)
$35,000-$49,999	-.023 (.100)
$50,000-$74,999	-.096 (.096)
$75,000-$99,999	-.209* (.096)
$100,000+	-.106 (.098)
Victim of crime in past year	.067 (.042)
Constant	.124 (.144)
F	4.20***
R^2	.292
RMSE	.242
N	139

Linear regression with robust standard errors.
* *p*<.05, ** *p*<.01, *** *p*<.001

Table A33: Perceptions of police (continuous outcomes)

	Police effectiveness	Police legitimacy
	b (Robust SE)	b (Robust SE)
Wave 2	.061 (.406)	-.225 (.254)
Treatment	-.094 (.194)	.246 (.144)
Wave 2 × Treatment	-.780 (.399)	-.549 (.317)
Race (ref:Black/African-American/African Immigrant)		
White	-.158 (.276)	.323 (.186)
Hispanic/Latinx	.084 (.234)	.446 (.485)
Asian	.038 (.231)	.092 (.235)
Other/Mixed/Multiple	.264 (.212)	.427* (.181)
Child currently in school	-.133 (.188)	-.029 (.138)
Born in US	.025 (.212)	-.311 (.180)
Highest level of education (ref:High school/GED)[a]		
Some college credit	-.218 (.202)	.147 (.289)
Associate's/Bachelor's degree	-.474* (.208)	.093 (.294)
Masters/graduate/professional degree	-.536 (.295)	.161 (.305)
Home status (ref:Rent)		
Own	.131 (.227)	-.250 (.217)
Household income (ref:<$20,000)		
$20,000-$34,999	-.166 (.165)	.065 (.312)
$35,000-$49,999	.460 (.234)	-.085 (.452)
$50,000-$74,999	-.211 (.160)	-.490* (.230)
$75,000-$99,999	.346 (.323)	-.394 (.267)
$100,000+	-.055 (.234)	-.139 (.249)
Victim of crime in past year	-.017 (.212)	-.115 (.145)
Constant	1.221*** (.278)	.798* (.364)
F	-	2.44*
R^2	.473	.599
RMSE	.423	.362
N	50	51

[a] No observations for primary/middle/some high school.
Linear regression with robust standard errors.
* $p<.05$, ** $p<.01$, *** $p<.001$

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table A34: Perceptions of police (categorical outcomes)

	Visibility of police	Satisfaction with police
	OR (Robust SE)	OR (Robust SE)
Wave 2	.425 (.446)	.658 (.621)
Treatment	.783 (.798)	3.296* (1.994)
Wave 2 $\times$ Treatment	1.033 (2.057)	.023 (.044)
Race (ref:Black/African-American/African Immigrant)		
White	1.133 (.810)	.213** (.126)
Hispanic/Latinx	<.0001*** (<.0001)	.224 (.236)
Asian	1.634 (3.398)	.669 (.580)
Other/Mixed/Multiple	5.791 (5.272)	.414 (.242)
Child currently in school	1.717 (1.452)	.734 (.389)
Born in US	.414 (.398)	.496 (.297)
Highest level of education (ref:Less than high school)[a]		
High school diploma/GED	-	.945 (1.163)
Some college credit	-	.213 (.241)
Associate's/Bachelor's degree	-	.463 (.609)
Masters/graduate/professional degree	-	.207 (.287)
Home status (ref:Rent)		
Own	2.243 (4.427)	5.363* (4.477)
Household income (ref:<$20,000)		
$20,000-$34,999	47.624* (74.819)	.995 (.897)
$35,000-$49,999	2.468 (4.132)	.469 (.422)
$50,000-$74,999	1.581 (3.093)	.920 (1.110)
$75,000-$99,999	3.472 (5.006)	.205 (.192)
$100,000+	2.599 (3.907)	.257 (.298)
Victim of crime in past year	.864 (.631)	.606 (.347)
cut1	*.269 (.602)*	*.005*** (.007)*
cut2	*1.685 (3.869)*	*.020** (.028)*
cut3	*13.050 (29.989)*	*.872 (1.170)*
Log pseudolikelihood	-61.630	-96.403
Pseudo R^2	.143	.200
Wald χ^2	.	36.783*
N	56	109

[a] Education variable omitted from police visibility model due to collinearity.
Ordered logistic regression with robust standard errors
Exponentiated coefficients (odds ratios)
* p<.05, ** p<.01, *** p<.001

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

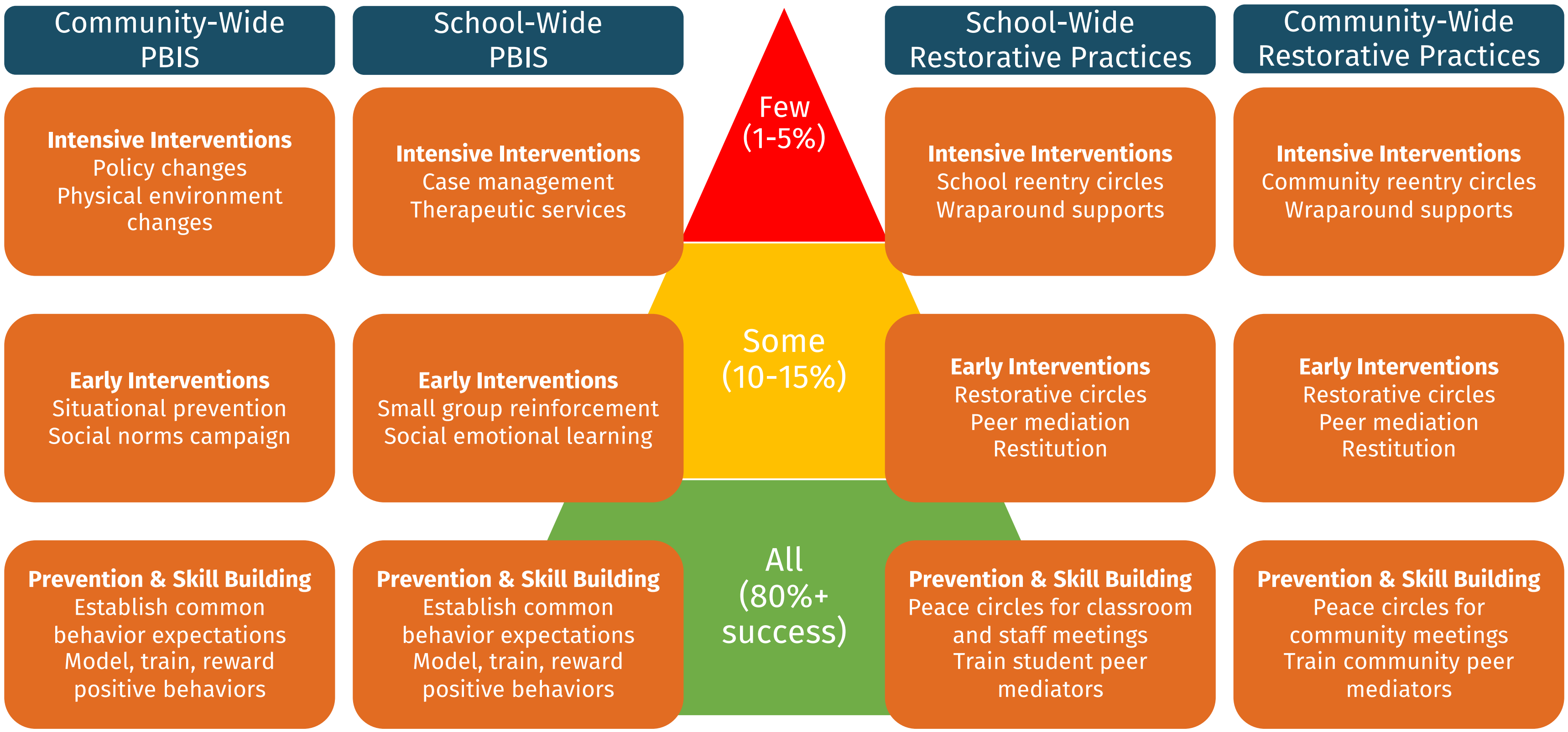
Table A35: Perceptions of safety and disorder

	Feelings of safety	Perceptions of disorder
	b (Robust SE)	b (Robust SE)
Wave 2	-.280 (.293)	.460 (.275)
Treatment	-.139 (.105)	.255 (.134)
Wave 2 × Treatment	-.293 (.371)	.320 (.416)
Race (ref:Black/African-American/African Immigrant)		
White	.193 (.148)	.137 (.172)
Hispanic/Latinx	.072 (.194)	-.213 (.253)
Asian	.042 (.155)	.024 (.230)
Other/Mixed/Multiple	-.133 (.132)	.028 (.164)
Child currently in school	.008 (.091)	.125 (.119)
Born in US	.020 (.114)	.027 (.154)
Highest level of education (ref:Less than high school)		
High school diploma/GED	-.090 (.220)	.319 (.300)
Some college credit	-.056 (.226)	.177 (.287)
Associate's/Bachelor's degree	-.130 (.244)	.038 (.297)
Masters/graduate/professional degree	.028 (.258)	.162 (.313)
Home status (ref:Rent)		
Own	.017 (.140)	.185 (.183)
Household income (ref:<$20,000)		
$20,000-$34,999	.018 (.164)	-.031 (.268)
$35,000-$49,999	.232 (.182)	-.212 (.310)
$50,000-$74,999	.185 (.190)	-.236 (.281)
$75,000-$99,999	.239 (.221)	-.244 (.309)
$100,000+	.197 (.193)	-.283 (.274)
Victim of crime in past year	-.034 (.089)	.192 (.106)
Constant	3.123*** (.259)	1.427*** (.346)
F	1.97*	2.17**
R^2	.250	.190
RMSE	.474	.639
N	139	138

Linear regression with robust standard errors.
* p<.05, ** p<.01, *** p<.001

Figures

Figure A1: Proposed extension of SW-PBIS and RJ practices to community settings



Adapted from https://www.pbis.org/pbis/what-is-pbis and Swain-Bradway et al. (2016)

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Figure A2: Predicted high school healthy community scale by treatment assignment and survey wave

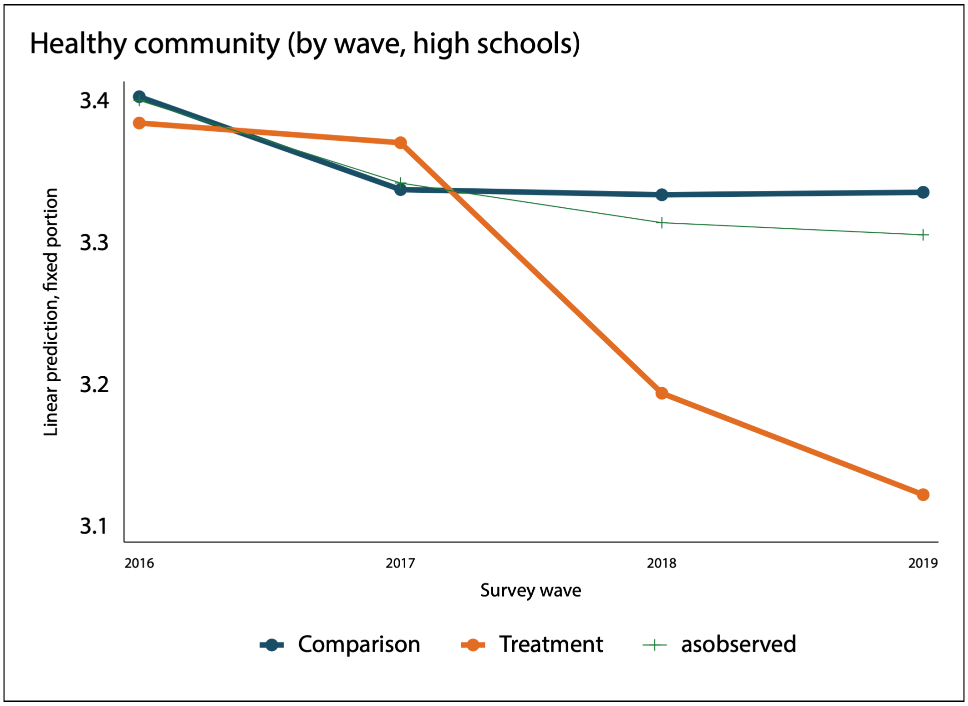


Figure A3: Predicted high school motivation & inclusion scale by treatment assignment and survey wave

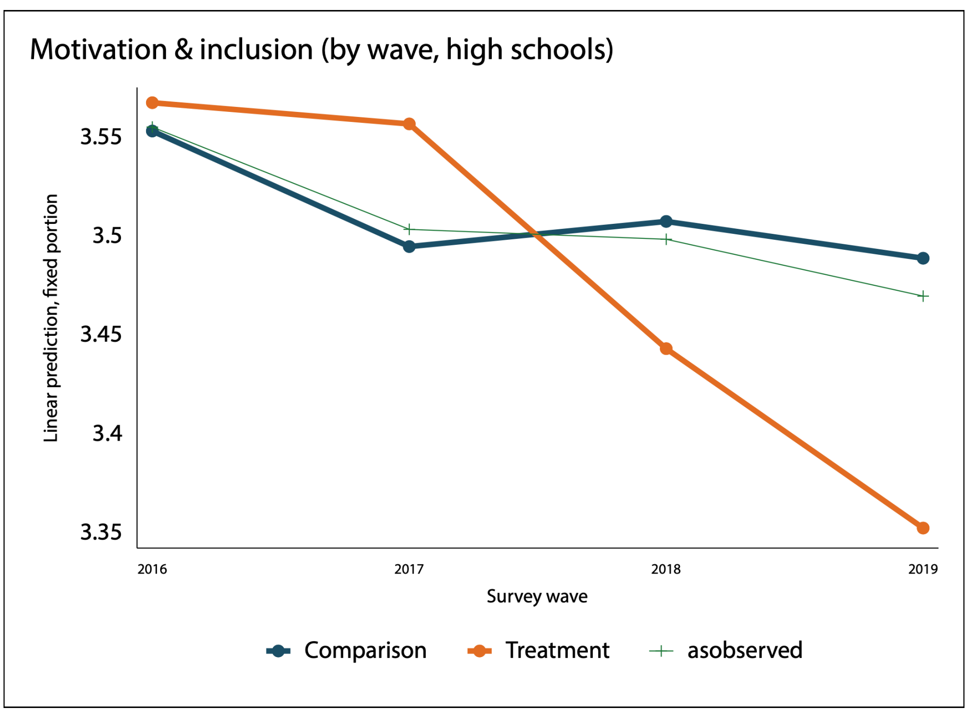


Figure A4: Predicted high school pedagogical effectiveness scale by treatment assignment and survey wave

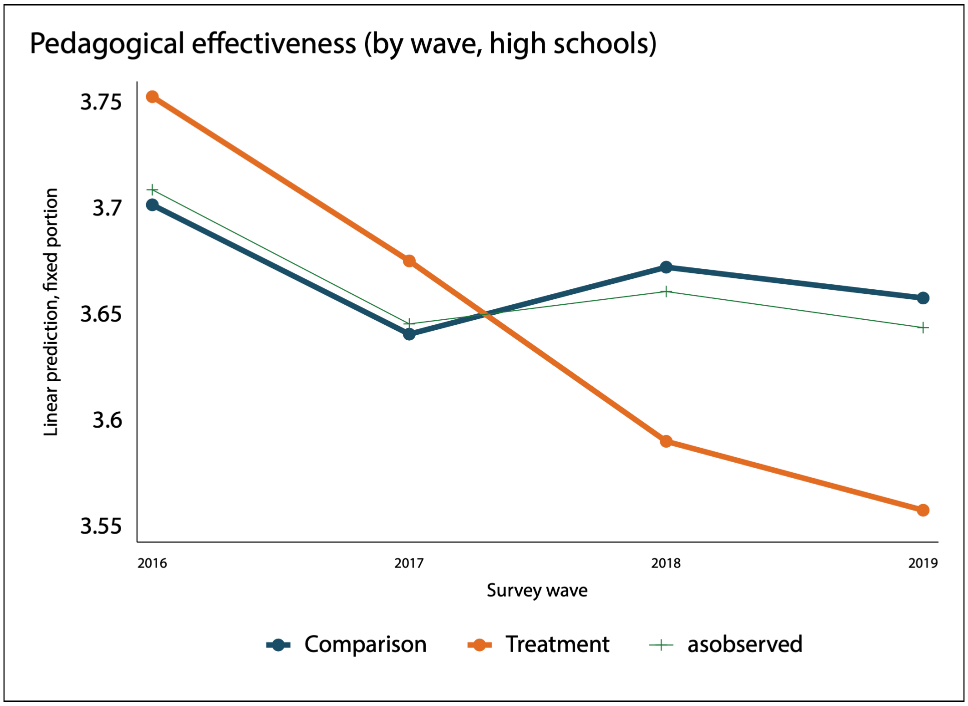


Figure A5: Predicted high school learning mindset scale by treatment assignment and survey wave

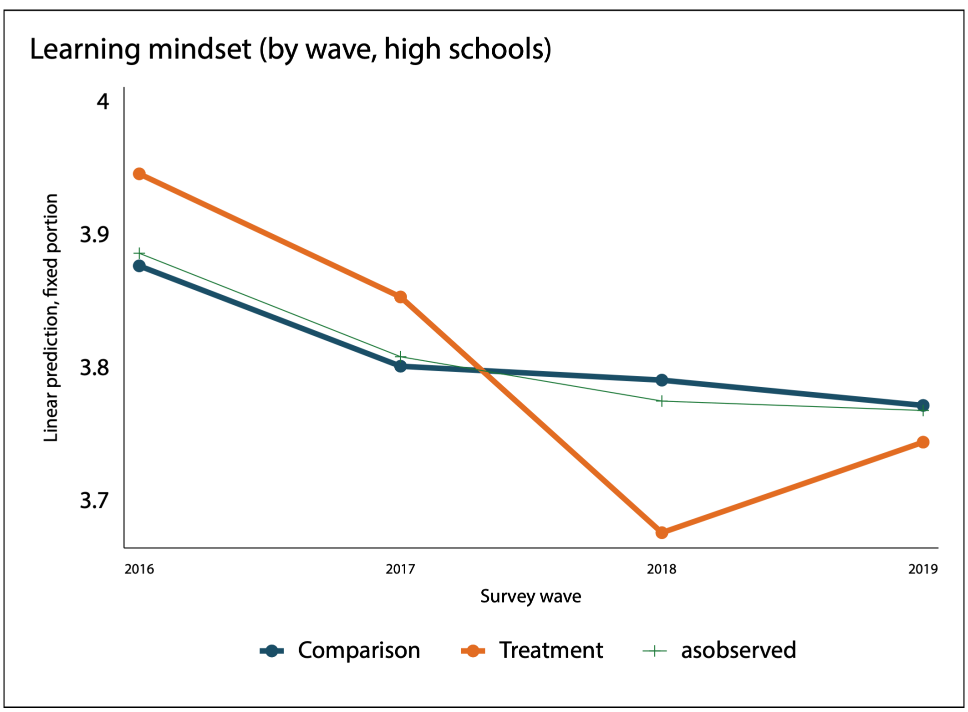


Figure A6: Propensity score matching results, elementary schools

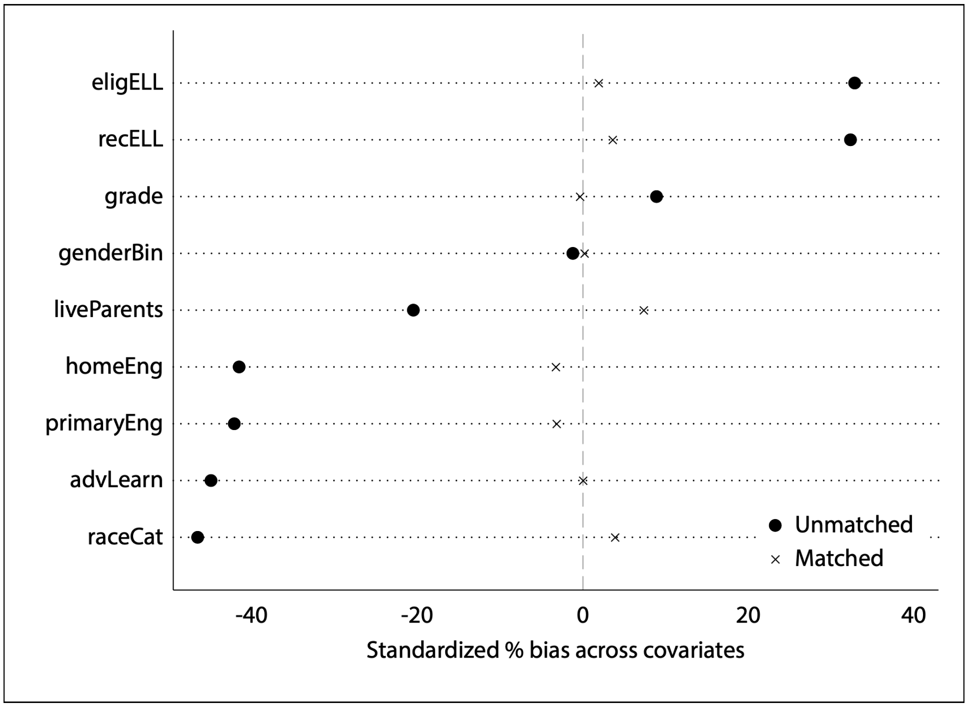


Figure A7: Propensity score matching results, high schools

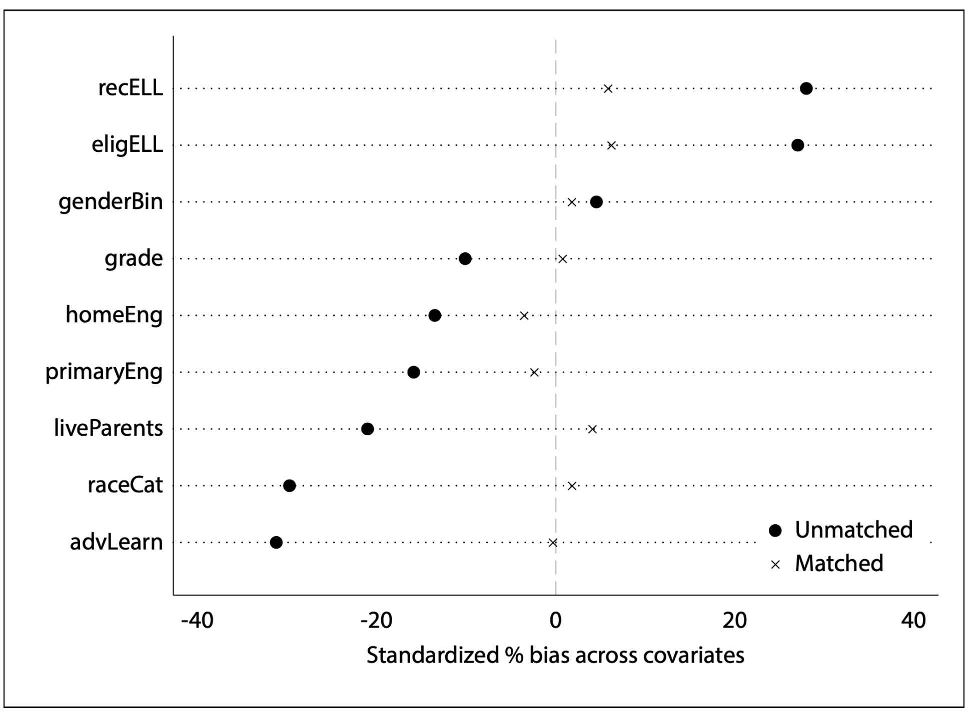


Figure A8: Predicted elementary school state ELA test score by treatment assignment and intervention status

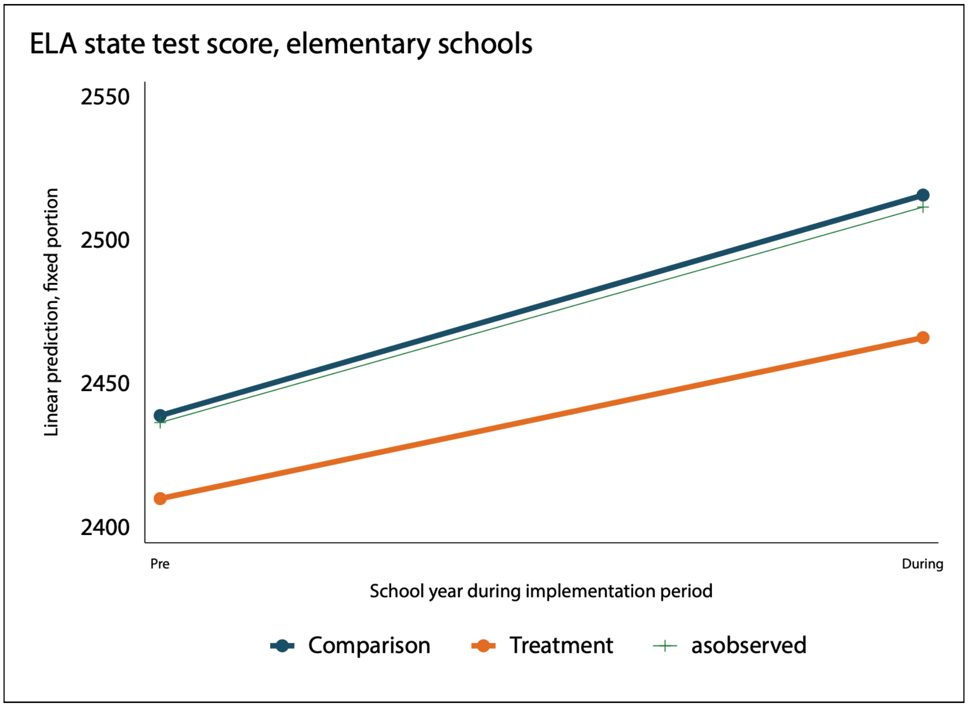


Figure A9: Predicted elementary school state math test score by treatment assignment and intervention status

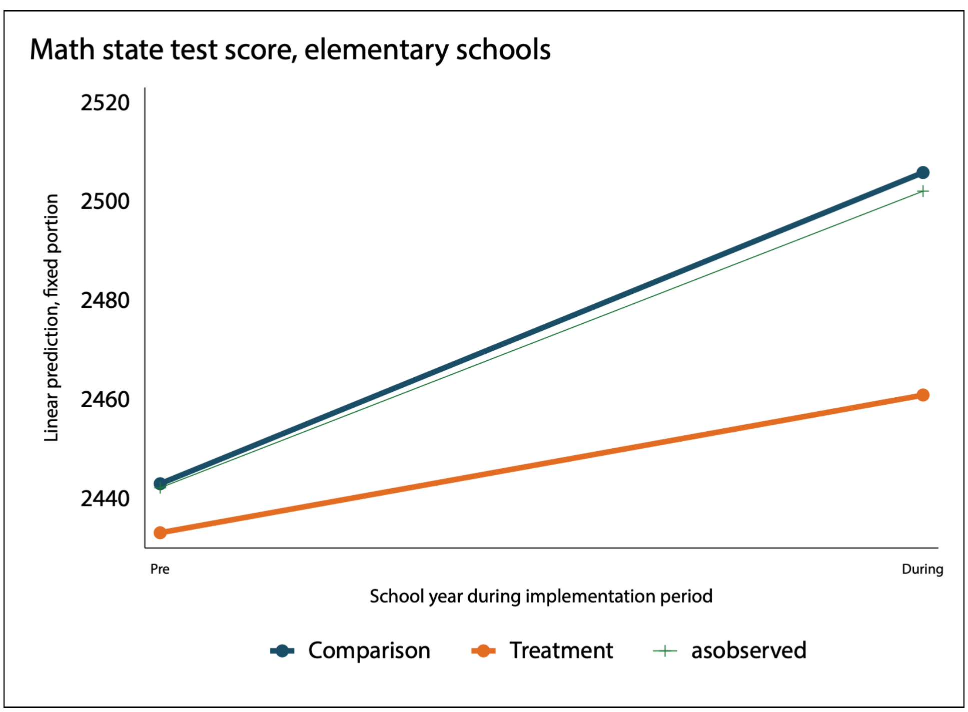


Figure A10: Predicted high school state ELA test score by treatment assignment and intervention status

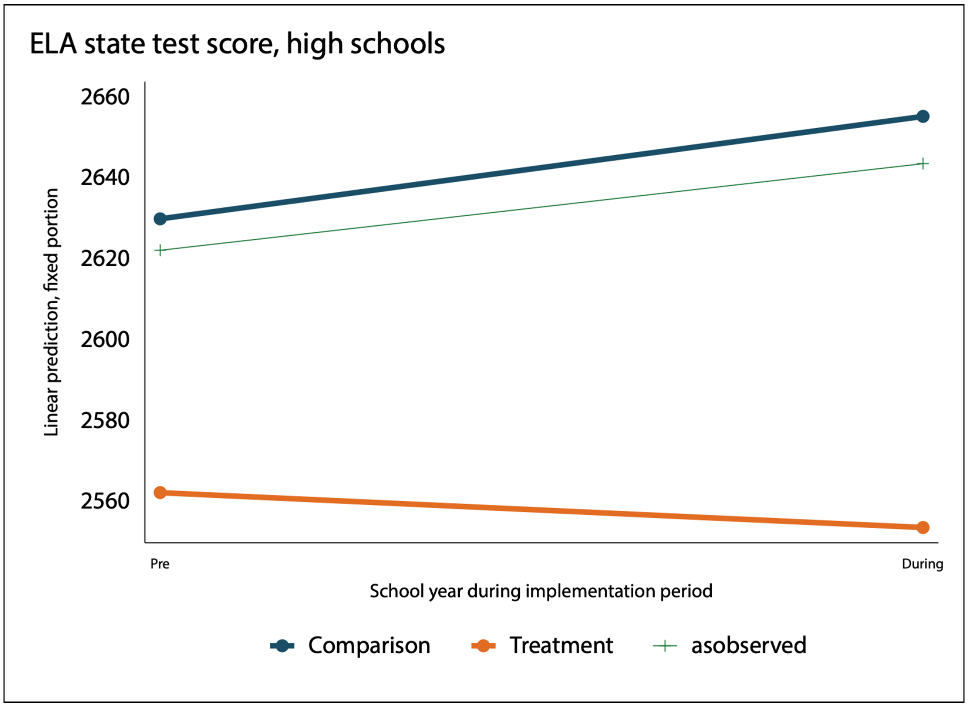


Figure A11: Predicted high school excused absences by treatment assignment and intervention status

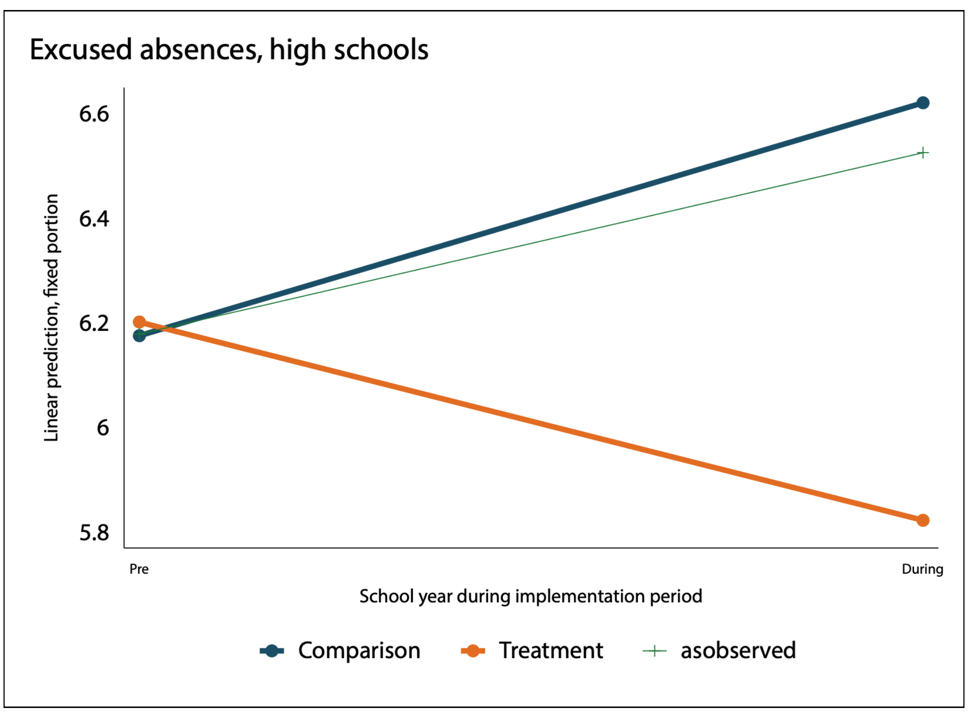


Figure A12: Predicted high school unexcused absence/eligible days ratio by treatment assignment and intervention status

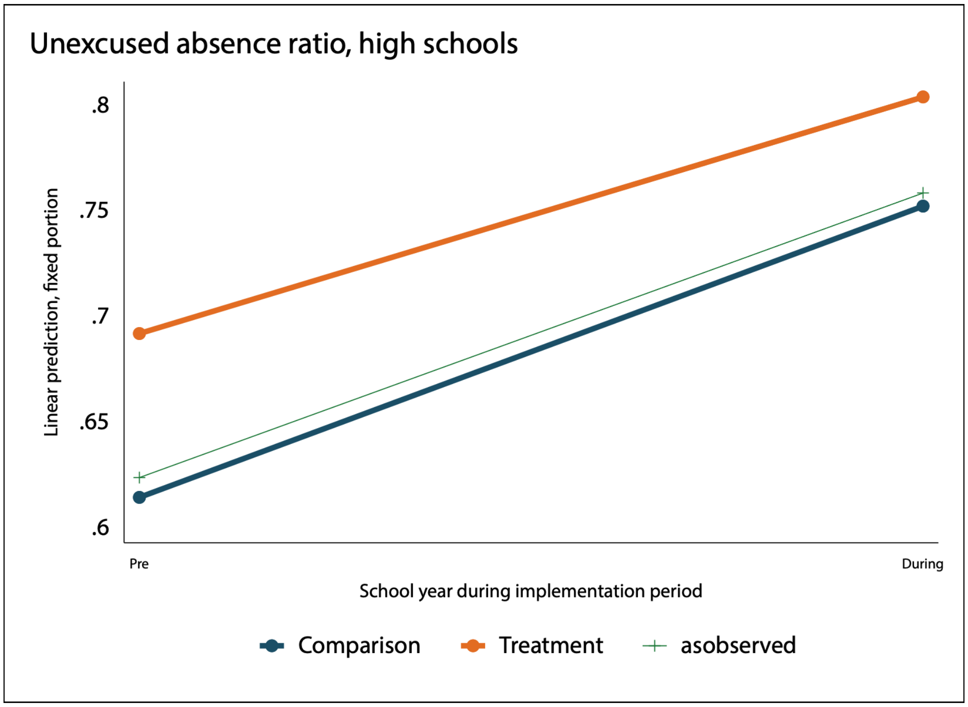
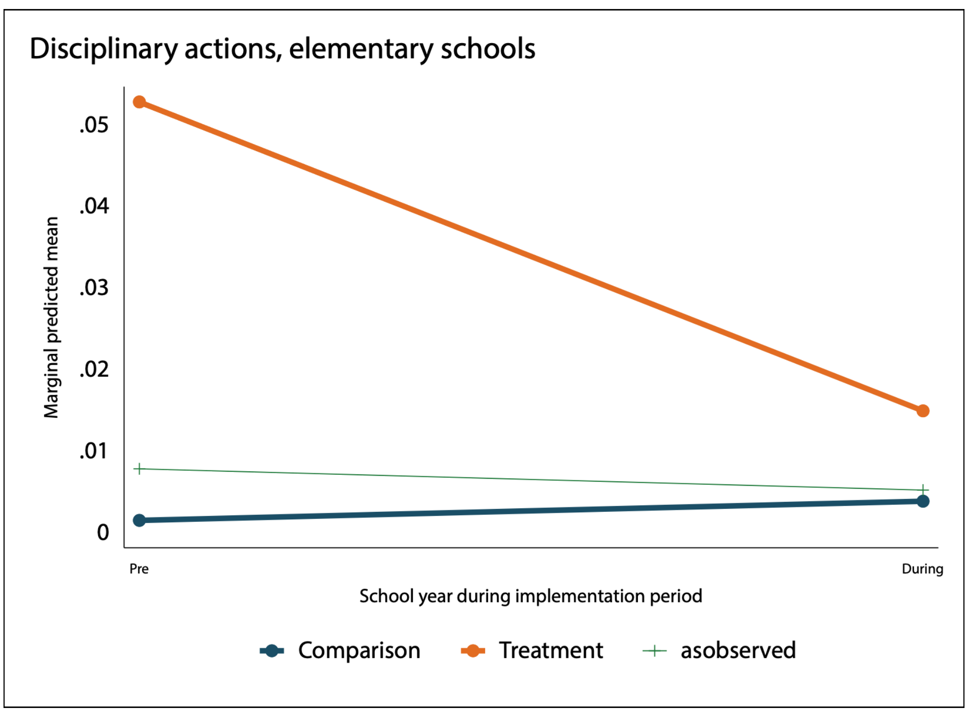


Figure A13: Predicted count of elementary school disciplinary actions by treatment assignment and intervention status



This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Figure A14: Calls for police service in combined treatment and comparison sites, July 2014-June 2022

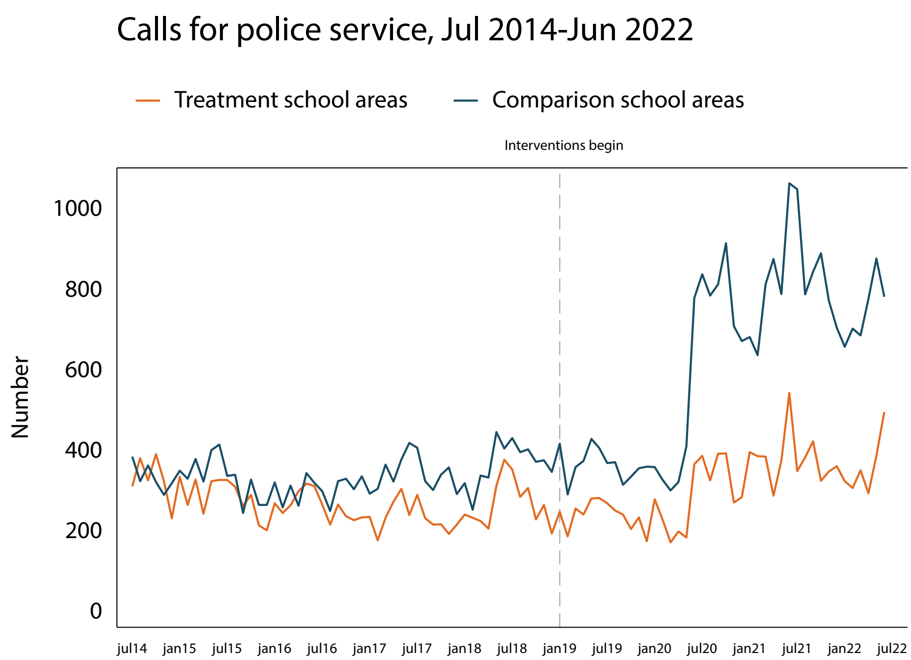


Figure A15: Percent change in calls for service in combined treatment and comparison sites, pre/post 2019

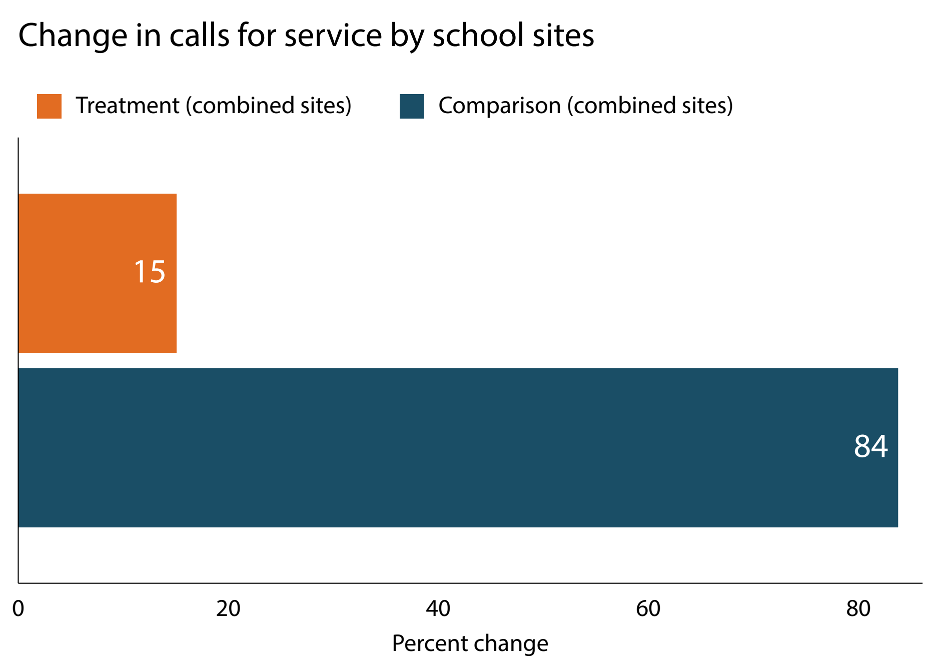
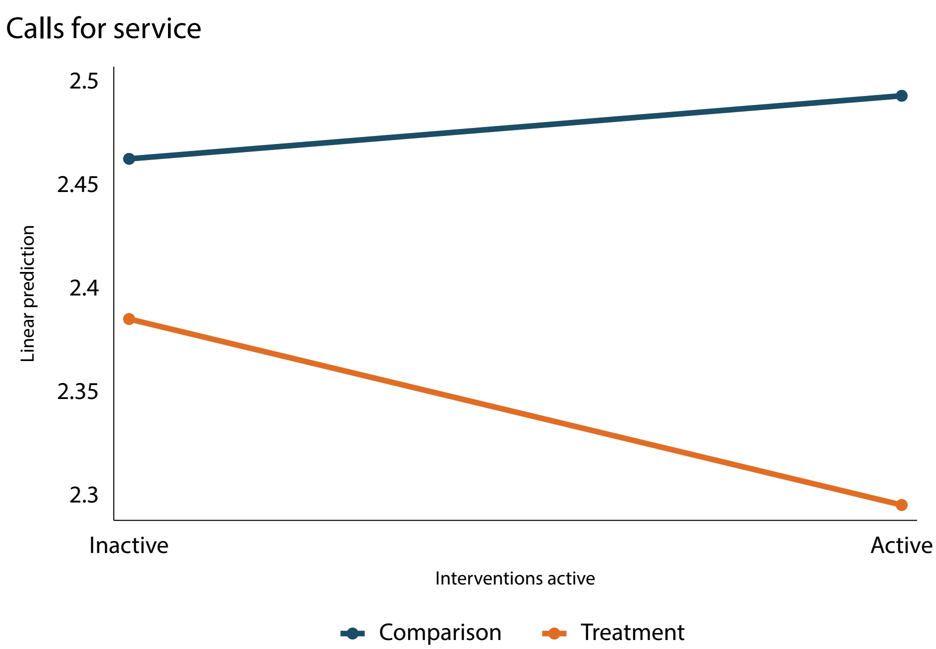


Figure A16: Predicted monthly calls for service by treatment assignment and intervention status



This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Figure A17: Offenses in combined treatment and comparison sites, July 2014-June 2022

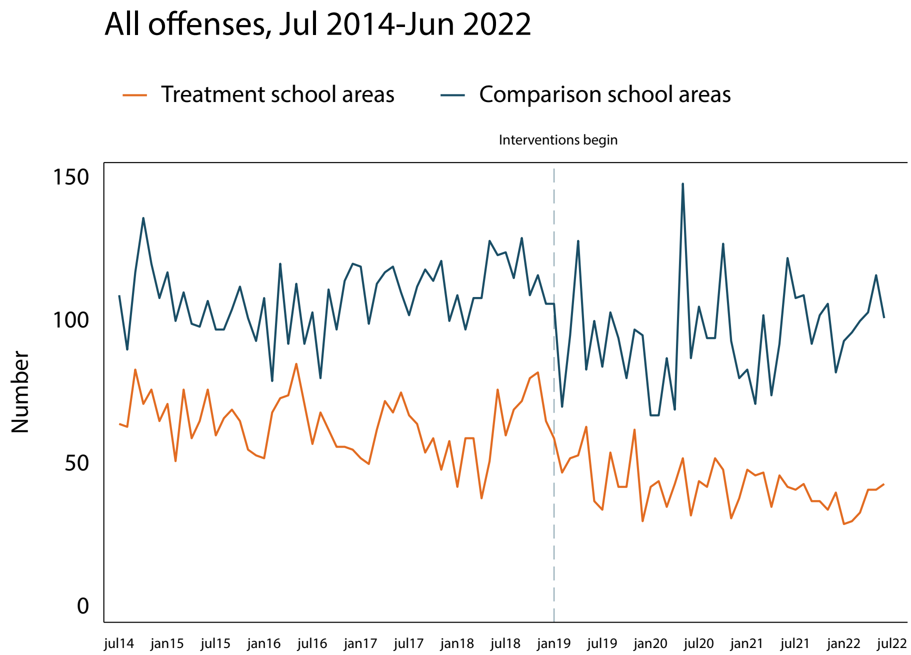


Figure A18: Percent change in offenses in combined treatment and comparison sites, pre/post 2019

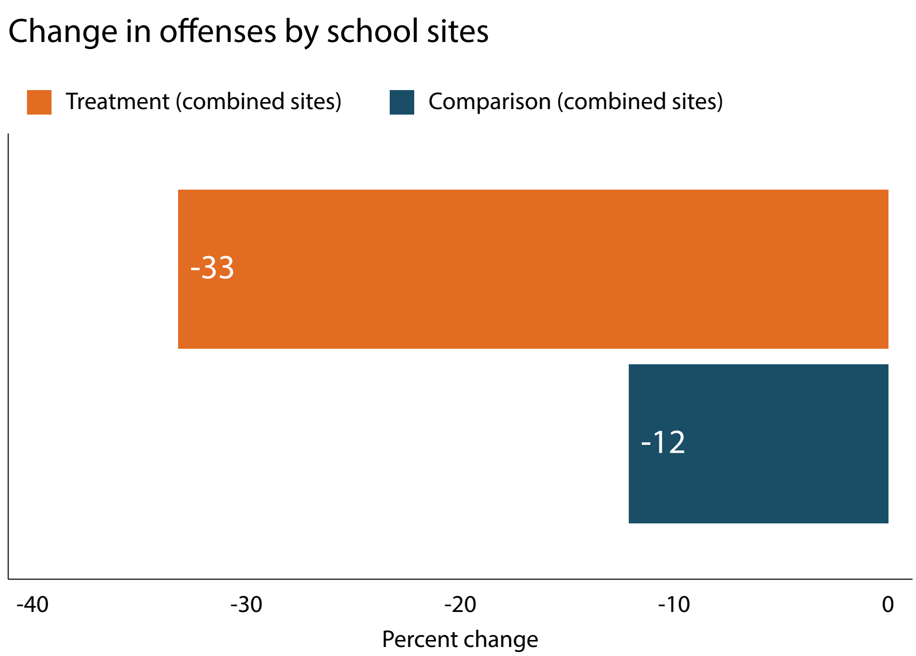
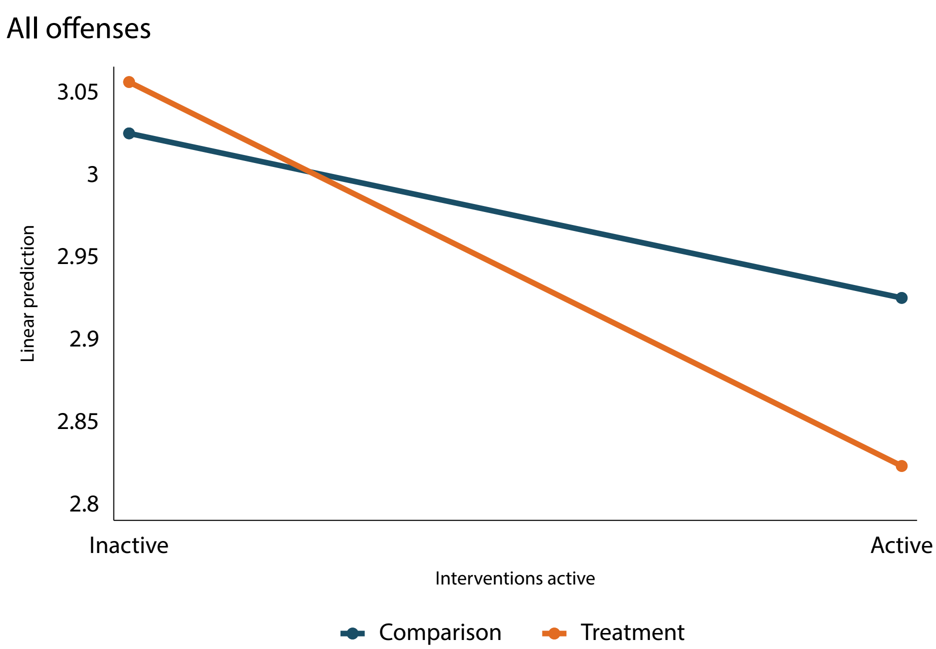


Figure A19: Predicted monthly offenses by treatment assignment and intervention status



This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Figure A20: Offenses involving juvenile suspects and/or victims in combined treatment and comparison sites, July 2014-June 2022

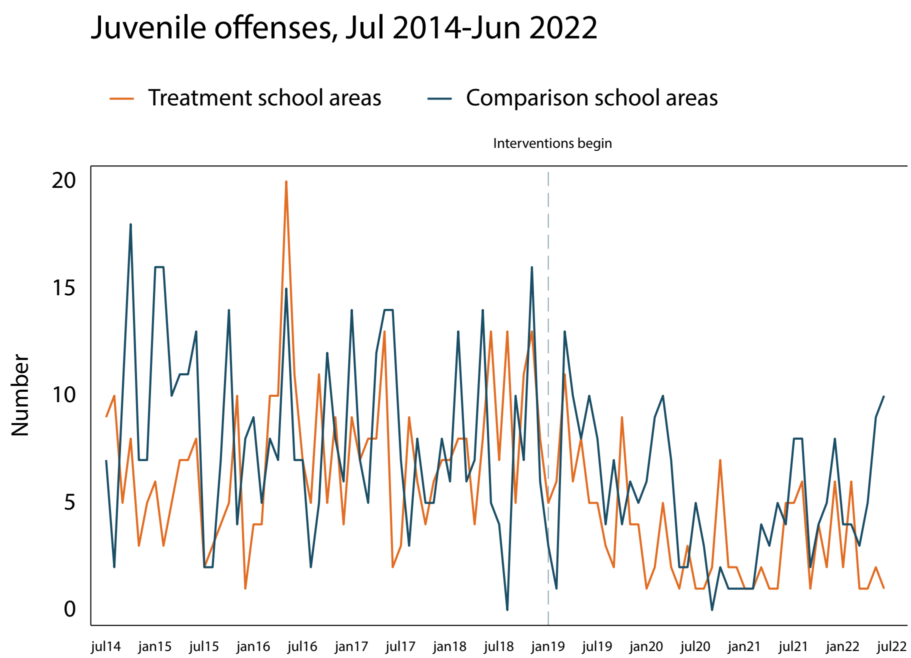
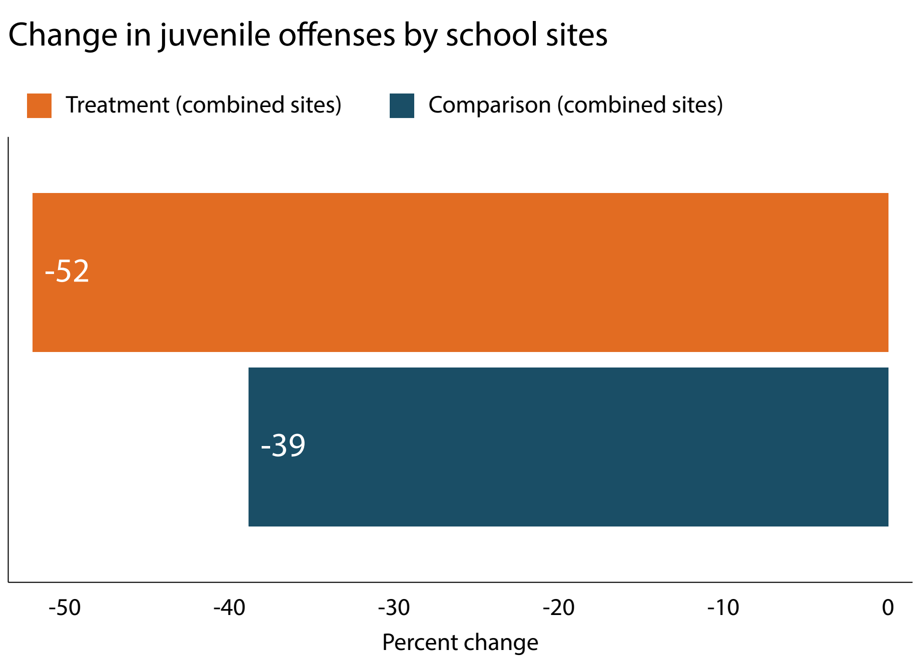


Figure A21: Percent change in offenses involving juvenile suspects and/or victims in combined treatment and comparison sites, pre/post 2019



Appendix B: References

Anderson, C. S. (1982). The search for school climate: A review of the research. *Review of Educational Research*, *52*(3), 368–420. https://doi.org/10.3102/00346543052003368

Arpino, B., & Cannas, M. (2016). Propensity score matching with clustered data. An application to the estimation of the impact of caesarean section on the Apgar score. *Statistics in Medicine*, *35*(12), 2074–2091. https://doi.org/10.1002/sim.6880

Arpino, B., & Mealli, F. (2011). The specification of the propensity score in multilevel observational studies. *Computational Statistics & Data Analysis*, *55*(4), 1770–1780. https://doi.org/10.1016/j.csda.2010.11.008

Battistich, V., Solomon, D., Kim, D.-i., Watson, M., & Schaps, E. (1995). Schools as communities, poverty levels of student populations, and students' attitudes, motives, and performance: A multilevel analysis. *American Educational Research Journal*, *32*(3), 627–658. https://doi.org/10.3102/00028312032003627

Bowen, N. K., & Bowen, G. L. (1999). Effects of crime and violence in neighborhoods and schools on the school behavior and performance of adolescents. *Journal of Adolescent Research*, *14*(3), 319–342. https://doi.org/10.1177/0743558499143003

Bradshaw, C. P. (2013). Preventing bullying through Positive Behavioral Interventions and Supports (PBIS): A multitiered approach to prevention and integration. *Theory Into Practice*, *52*(4), 288–295. https://doi.org/10.1080/00405841.2013.829732

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*(3), 133–148. https://doi.org/10.1177/1098300709334798

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. https://doi.org/10.1080/07418825.2012.673632

Braithwaite, J. (1989). *Crime, shame and reintegration*. Cambridge University Press.

Brantingham, P., & Brantingham, P. (1995). Criminality of place: Crime generators and crime attractors. *European Journal on Criminal Policy and Research*, *3*(3), 5–26. https://doi.org/10.1007/BF02242925

Brunson, R. K., & Miller, J. (2006). Young black men and urban policing in the United States. *The British Journal of Criminology*, *46*(4), 613–640. https://doi.org/10.1093/bjc/azi093

Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review*, *44*(4), 588–608. https://doi.org/10.2307/2094589

Duncan, G., & Raudenbush, S. (1999). Assessing the effects of context in studies of child and youth development. *Educational Psychologist*, *34*(1), 29–41. https://doi.org/10.1207/s15326985ep3401_3

Fabelo, T., Thompson, M. D., Plotkin, M., Carmichael, D., Marchbanks III, M. P., & Booth, E. A. (2011). *Breaking schools' rules: A statewide study of how school discipline relates to students' success and juvenile justice involvement*. Council of State Governments Justice Center. New York. https://csgjusticecenter.org/wp-content/uploads/2020/01/Breaking_Schools_Rules_Report_Final.pdf

Felson, M. (1987). Routine activities and crime prevention in the developing metropolis. *Criminology*, *25*(4), 911–932. https://doi.org/10.1111/j.1745-9125.1987.tb00825.x

Felson, M. (2006). *Crime and nature*. Sage.

Felson, M., & Boba, R. (2010). *Crime and everyday life* (4th ed.). Sage.

Fischer, S., & Argyle, D. (2018). Juvenile crime and the four-day school week. *Economics of Education Review*, *64*, 31–39. https://doi.org/10.1016/j.econedurev.2018.03.010

Fronius, T., Persson, H., Guckenburg, S., Hurley, N., & Petrosino, A. (2016). *Restorative justice in U.S. schools: A research review*. WestEd. San Francisco, CA. https://www.wested.org/resources/restorative-justice-research-review/

Gill, C., & Gross Shader, C. (2020). Building a "Beautiful Safe Place for Youth:" The story of an effective community-research-practice partnership in Rainier Beach, Seattle. In R. J. Stokes & C. Gill (Eds.), *Innovations in community-based crime prevention: Case studies and lessons learned*. Springer.

Gill, C., & Prince, H. (2020). *Rainier Beach: A Beautiful Safe Place for Youth. 2019 evaluation update*. Center for Evidence-Based Crime Policy, Department of Criminology, Law & Society, George Mason University. Fairfax, VA. https://www.rb-safeplaceforyouth.com/wp-content/uploads/2021/01/2019-GMU-ABSPY-Evaluation-Report.pdf

Gill, C., Vitter, Z., & Weisburd, D. (2015). *Identifying hot spots of juvenile offending: A guide for crime analysts*. U.S. Department of Justice, Office of Community Oriented Policing Services. Washington, DC. https://cops.usdoj.gov/RIC/ric.php?page=detail&id=COPS-P298

Gill, C., Vitter, Z., & Weisburd, D. (2016). *Rainier Beach: A Beautiful Safe Place for Youth. Final evaluation report*. Center for Evidence-Based Crime Policy, Department of Criminology, Law & Society, George Mason University. Fairfax, VA. http://www.rb-safeplaceforyouth.com/wp-content/uploads/2018/03/2016-GMU-ABSPY-evaluation-report.pdf

Gill, C., & Wilson, D. B. (2017). Improving the success of reentry programs: Identifying the impact of service–need fit on recidivism. *Criminal Justice and Behavior*, *44*(3), 336–359. https://doi.org/10.1177/0093854816682048

Gottfredson, D. C. (1986). An empirical test of school-based environmental and individual interventions to reduce the risk of delinquent behavior. *Criminology*, *24*(4), 705–731. https://doi.org/10.1111/j.1745-9125.1986.tb01508.x

Gottfredson, D. C., Gottfredson, G. D., & Weisman, S. A. (2001). The timing of delinquent behavior and its implications for after-school programs. *Criminology & Public Policy*, *1*(1), 61–86. https://doi.org/10.1111/j.1745-9133.2001.tb00077.x

Gottfredson, G. D., & Gottfredson, D. C. (1985). *Victimization in schools*. Plenum.

Gottfredson, G. D., Gottfredson, D. C., Payne, A. A., & Gottfredson, N. C. (2005). School climate predictors of school disorder: Results from a national study of delinquency prevention in schools. *Journal of Research in Crime and Delinquency*, *42*(4), 412–444. https://doi.org/10.1177/0022427804271931

Gregory, A., Clawson, K., Davis, A., & Gerewitz, J. (2016). The promise of restorative practices to transform teacher-student relationships and achieve equity in school discipline. *Journal of Educational and Psychological Consultation*, *26*(4), 325–353. https://doi.org/10.1080/10474412.2014.929950

Gross Shader, C. (2020). *Can city government function as a super-controller at crime "hot spots" to create lasting crime control benefits?* [Doctoral dissertation, Hebrew University of Jerusalem].

Hawkins, J. D., Herrenkohl, T. I., Brewer, D. D., Catalano, R. F., Harachi, T. W., Loeber, R., & Farrington, D. P. (1998). A review of predictors of youth violence. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 106–146). Sage.

Henrich, C. C., Schwab-Stone, M., Fanti, K., Jones, S. M., & Ruchkin, V. (2004). The association of community violence exposure with middle-school achievement: A prospective study. *Journal of Applied Developmental Psychology*, *25*(3), 327–348. https://doi.org/10.1016/j.appdev.2004.04.004

Herrenkohl, T. I., Maguin, E., Hill, K. G., Hawkins, J. D., Abbott, R. D., & Catalano, R. F. (2000). Developmental risk factors for youth violence. *Journal of Adolescent Health*, *26*(3), 176–186. https://doi.org/10.1016/S1054-139X(99)00065-8

Horner, R. H., & Sugai, G. (2015). School-wide PBIS: An example of applied behavior analysis implemented at a scale of social importance. *Behavior Analysis in Practice*, *8*(1), 80–85. https://doi.org/10.1007/s40617-015-0045-4

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Horner, R. H., Sugai, G., & Anderson, C. M. (2010). Examining the evidence base for School-wide Positive Behavior Support. *Focus on Exceptional Children*, *42*(8), 1–14. https://journals.ku.edu/FOEC/article/download/6906/6254

Horner, R. H., Sugai, G., & Lewis, T. (2015). *Is School-wide Positive Behavioral Interventions and Supports an evidence-based practice?* U.S. Department of Education, Office of Special Education Programs, Technical Assistance Center on Positive Behavioral Interventions and Supports. Washington, DC. https://www.pbis.org/research

Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A. W., & Esperanza, J. (2009). A randomized, wait-list controlled effectiveness trial assessing School-wide Positive Behavior Support in elementary schools. *Journal of Positive Behavior Interventions*, *11*(3), 133–144. https://doi.org/10.1177/1098300709332067

Kendi, I. X. (2019). *How to be an antiracist*. One World.

Kirk, D. S. (2008). The neighborhood context of racial and ethnic disparities in arrest. *Demography*, *45*(1), 55–77. https://doi.org/10.1353/dem.2008.0011

Kitsantas, A., Ware, H. W., & Martinez-Arias, R. (2004). Students' perceptions of school safety: Effects by community, school environment, and substance use variables. *The Journal of Early Adolescence*, *24*(4), 412–430. https://doi.org/10.1177/0272431604268712

Kochel, T. R., & Weisburd, D. (2017). Assessing community consequences of implementing hot spots policing in residential areas: Findings from a randomized field trial. *Journal of Experimental Criminology*, *13*(2), 143–170. https://doi.org/10.1007/s11292-017-9283-5

Kondo, M. C., Keene, D., Hohl, B. C., MacDonald, J. M., & Branas, C. C. (2015). A difference-in-differences study of the effects of a new abandoned building remediation strategy on safety. *PLOS ONE*, *10*(7), e0129582. https://doi.org/10.1371/journal.pone.0129582

Kubrin, C. E., & Weitzer, R. (2003). New directions in social disorganization theory. *Journal of Research in Crime and Delinquency*, *40*(4), 374–402. https://doi.org/10.1177/0022427803256238

Latimer, J., Dowden, C., & Muise, D. (2001). *The effectiveness of restorative justice practices: A meta-analysis*. Department of Justice Canada. Ottawa, Canada. http://www.justice.gc.ca/eng/rp-pr/csj-sjc/jsp-sjp/rp01_1-dr01_1/index.html

Leuven, E., & Sianesi, B. (2016, January 30). *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing* (Version 4.0.12). http://ideas.repec.org/c/boc/bocode/s432001.html

Limbos, M. A. P., & Casteel, C. (2008). Schools and neighborhoods: Organizational and environmental factors associated with crime in secondary schools. *Journal of School Health*, *78*(10), 539–544. https://doi.org/10.1111/j.1746-1561.2008.00341.x

Lindstrom Johnson, S., Waasdorp, T. E., Cash, A. H., Debnam, K. J., Milam, A. J., & Bradshaw, C. P. (2017). Assessing the association between observed school disorganization and school violence: Implications for school climate interventions. *Psychology of Violence*, *7*(2), 181–191. https://doi.org/10.1037/vio0000045

Losen, D. (2014). *Closing the school discipline gap: Equitable remedies for excessive exclusion*. Teachers College Press.

Margolin, G., & Gordis, E. B. (2000). The effects of family and community violence on children. *Annual Review of Psychology*, *51*(1), 445–479. https://doi.org/10.1146/annurev.psych.51.1.445

McCoy, D. C., Roy, A. L., & Sirkman, G. M. (2013). Neighborhood crime and school climate as predictors of elementary school academic quality: A cross-lagged panel analysis. *American Journal of Community Psychology*, *52*(1-2), 128–140. https://doi.org/10.1007/s10464-013-9583-5

McDougal, L., Atmavilas, Y., Hay, K., Silverman, J. G., Tarigopula, U. K., & Raj, A. (2017). Making the continuum of care work for mothers and infants: Does gender equity matter? Findings from a quasiexperimental study in Bihar, India. *PLOS ONE*, *12*(2), e0171002. https://doi.org/10.1371/journal.pone.0171002

Morenoff, J. D., Sampson, R. J., & Raudenbush, S. W. (2001). Neighborhood inequality, collective efficacy, and the spatial dynamics of urban violence. *Criminology*, *39*(3), 517–558. https://doi.org/10.1111/j.1745-9125.2001.tb00932.x

Mowen, T., & Brent, J. (2016). School discipline as a turning point: The cumulative effect of suspension on arrest. *Journal of Research in Crime and Delinquency*, *53*(5), 628–653. https://doi.org/10.1177/0022427816643135

Osgood, D. W., Wilson, J. K., O'Malley, P. M., Bachman, J. G., & Johnston, L. D. (1996). Routine activities and individual deviant behavior. *American Sociological Review*, *61*(4), 635–655. https://doi.org/10.2307/2096397

Payne, A. A., & Welch, K. (2010). Modeling the effects of racial threat on punitive and restorative school discipline practices. *Criminology*, *48*(4), 1019–1062. https://doi.org/10.1111/j.1745-9125.2010.00211.x

Petrosino, A., Turpin-Petrosino, C., & Guckenburg, S. (2010). Formal system processing of juveniles: Effects on delinquency. A systematic review. *Campbell Systematic Reviews*, *6*(1), 1–88. https://doi.org/10.4073/csr.2010.1

Presser, L., & Van Voorhis, P. (2002). Values and evaluation: Assessing processes and outcomes of restorative justice programs. *Crime & Delinquency*, *48*(1), 162–188. https://doi.org/10.1177/0011128702048001007

Reaves, S., McMahon, S. D., Duffy, S. N., & Ruiz, L. (2018). The test of time: A meta-analytic review of the relation between school climate and problem behavior. *Aggression and Violent Behavior*, *39*, 100–108. https://doi.org/10.1016/j.avb.2018.01.006

Roman, C. G. (2002). *Schools as generators of crime: Routine activities and the sociology of place* [Doctoral dissertation, American University]. Washington, DC. https://www.ncjrs.gov/pdffiles1/nij/grants/201946.pdf

Roman, C. G. (2005). Routine activities of youth and neighborhood violence: Spatial modeling of place, time, and crime. In F. Wang (Ed.), *Geographic information systems and crime analysis* (pp. 293–310). Idea Group.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. https://doi.org/10.1093/biomet/70.1.41

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, *79*(387), 516–524. https://doi.org/10.1080/01621459.1984.10478078

Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services & Outcomes Research Methodology*, *2*, 169–188. https://doi.org/10.1023/A:1020363010465

Ruiz, L. D., McMahon, S. D., & Jason, L. A. (2018). The role of neighborhood context and school climate in school-level academic achievement. *American Journal of Community Psychology*, *61*(3-4), 296–309. https://doi.org/10.1002/ajcp.12234

Sampson, R. J., & Morenoff, J. D. (1997). Ecological perspectives on the neighborhood context of urban poverty: Past and present. In J. Brooks-Gunn, G. J. Duncan, & J. L. Aber (Eds.), *Neighborhood poverty: Policy implications in studying poverty* (pp. 1–22). Russell Sage Foundation.

Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, *277*(5328), 918–924. https://doi.org/10.1126/science.277.5328.918

Shaw, C. R., & McKay, H. D. (1942). *Juvenile delinquency in urban areas*. University of Chicago Press.

Sheldon, S. B., & Epstein, J. L. (2002). Improving student behavior and school discipline with family and community involvement. *Education and Urban Society*, *35*(1), 4–26. https://doi.org/10.1177/001312402237212

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Sherman, L. W., Strang, H., Mayo-Wilson, E., Woods, D. J., & Ariel, B. (2015). Are restorative justice conferences effective in reducing repeat offending? Findings from a Campbell systematic review. *Journal of Quantitative Criminology*, *31*(1), 1–24. https://doi.org/10.1007/s10940-014-9222-9

Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *Urban Review*, *34*, 317–342. https://doi.org/10.1023/A:1021320817372

Strang, H. (2002). *Repair or revenge: Victims and restorative justice*. Clarendon.

Sugai, G., & Horner, R. H. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review*, *35*(2), 245–259. https://doi.org/10.1080/02796015.2006.12087989

Swain-Bradway, J., Eber, L., Sprague, J., & Nelson, M. (2016). *Restorative justice practices in school-wide Positive Behavior Interventions and Support: A model for alignment and evaluation*.

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research*, *83*(3), 357–385. https://doi.org/10.3102/0034654313483907

Tyler, T. R. (1990). *Why people obey the law: Procedural justice, legitimacy, and compliance*. Yale University Press.

Valdebenito, S., Eisner, M., Farrington, D. P., Ttofi, M. M., & Sutherland, A. (2018). School-based interventions for reducing disciplinary school exclusion: A systematic review. *Campbell Systematic Reviews*, *14*(1). https://doi.org/10.4073/csr.2018.1

Weisburd, D. (2015). The law of crime concentration and the criminology of place. *Criminology*, *53*(2), 133–157. https://doi.org/10.1111/1745-9125.12070

Weisburd, D., Bushway, S., Lum, C., & Yang, S.-M. (2004). Trajectories of crime at places: A longitudinal study of street segments in the city of Seattle. *Criminology*, *42*(2), 283–322. https://doi.org/10.1111/j.1745-9125.2004.tb00521.x

Weisburd, D., Gill, C., Wooditch, A., Barritt, W., & Murphy, J. (2021). Building collective action at crime hot spots: Findings from a randomized field experiment. *Journal of Experimental Criminology*, *17*, 161–191. https://doi.org/10.1007/s11292-019-09401-1

Weisburd, D., Groff, E. R., & Yang, S.-M. (2012). *The criminology of place: Street segments and our understanding of the crime problem*. Oxford University Press.

Weisburd, D., Morris, N. A., & Groff, E. R. (2009). Hot spots of juvenile crime: A longitudinal study of arrest incidents at street segments in Seattle, Washington. *Journal of Quantitative Criminology*, *25*(4), 443–467. https://doi.org/10.1007/s10940-009-9075-9

Welsh, W. N. (2000). The effects of school climate on school disorder. *The Annals of the American Academy of Political and Social Science*, *567*(1), 88–107. https://doi.org/10.1177/000271620056700107

Welsh, W. N. (2001). Effects of student and school factors on five measures of school disorder. *Justice Quarterly*, *18*(4), 911–947. https://doi.org/10.1080/07418820100095131

Welsh, W. N., Greene, J. R., & Jenkins, P. H. (1999). School disorder: The influence of individual, institutional, and community factors. *Criminology*, *37*(1), 73–116. https://doi.org/10.1111/j.1745-9125.1999.tb00480.x

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.