Document Title:	PBIS in Challenging Contexts: Evaluating a Replicable Implementation Approach in Philadelphia
Author(s):	Ryan Fink, Katarina Suwak, Karina Diaz, Gwendolyn Lawson, Naomi Goldstein, Joseph Gardella, A. Brooks Bowden, Anika Alam
Document Number:	309966
Date Received:	January 2025
Award Number:	2017-CK-BX-0016

Project Title: PBIS in Challenging Contexts: Evaluating a Replicable Implementation Approach in Philadelphia

National Institute of Justice, Award Number: 2017-CK-BX-0016

Project period: 01-Apr-18 through 30-Sep 2022

Award amount: $2,999,912

Award recipient organization
The Trustees of the University of Pennsylvania
Office of Research Services
5th Floor, Franklin Building
3451 Walnut Street
Philadelphia, PA 19104-6205

Report Authors: Ryan Fink, Katarina Suwak, Karina Diaz, Gwendolyn Lawson, Naomi Goldstein, Joseph Gardella, A. Brooks Bowden, Anika Alam

PI: Ryan Fink, Senior Research Specialist, Consortium for Policy Research in Education, University of Pennsylvania

Co-PI: Gwendolyn Lawson, Assistant Professor of Psychiatry, Children's Hospital of Philadelphia; University of Pennsylvania, Perelman School of Medicine

Contact information
Ryan Fink
ryanfi@upenn.edu
3440 Market Street, STE 560
Philadelphia, PA, 19104
215-573-0780

Positive Behavioral Interventions and Supports (PBIS) in Challenging Contexts: Evaluating a Replicable Implementation Approach in Philadelphia

Ryan Fink[1], Katarina Suwak[1], Karina Diaz[1], Gwendolyn Lawson[2][3], Naomi Goldstein[4], Joseph Gardella[4], A. Brooks Bowden[2], Anika Alam[2]

[1]Consortium for Policy Research in Education
[2]University of Pennsylvania, [3]Children's Hospital of Philadelphia, [4]Drexel University

September 2022

Author Note

Ryan Fink, University of Pennsylvania Graduate School of Education
 https://orcid.org/0000-0002-0019-9510

Gwendolyn Lawson, University of Pennsylvania Perelman School of Medicine and Children's Hospital of Philadelphia

Correspondence concerning this report should be addressed to Ryan Fink, 3440 Market Street – Suite 560, Philadelphia PA 19104. Email: ryanfi@upenn.edu

Table of Contents

Abstract

This report represents the final research report for the National Institute of Justice-funded project, "PBIS in Challenging Contexts: Evaluating a Replicable Implementation Approach in Philadelphia" (#2017-CK-BX-0016). The project was awarded to the University of Pennsylvania's Consortium for Policy Research in Education and included several community and academic partners, including a collaborating school district. The project period began in April 2018 and concluded in September 2022. This report shares the project's main goals and objectives and reports results regarding two different interventions: large scale brief trauma trainings for school safety officers, and an expanded, trauma-informed PBIS implementation model, implemented in four schools. Both interventions were developed and implemented throughout the project by a variety of project partners, including significant contributions from the collaborating school district. Research findings shared within are a result of survey data collected from school safety officers, teacher surveys, school staff interviews, and analysis of school district administrative data. Implementation of the project's components, as well as the timeline for analysis and project completion, were significantly affected by the COVID-19 pandemic and those implications are discussed within.

Positive Behavioral Interventions and Supports (PBIS) in Challenging Contexts:

Evaluating a Replicable Implementation Approach in Philadelphia

1. Summary of the Project

1.1. Major Goals and Objectives

This project aimed to validate a replicable approach to implementing Positive Behavior Interventions and Supports (PBIS), a school-climate improvement program, in a partnering school district. Despite some successes, implementation of PBIS in the partnering school district had proven insufficient for moving most schools beyond initial phases of adoption. By validating a replicable approach to PBIS implementation to improve its effectiveness in a high-poverty urban setting, this project aimed to deliver a locally and nationally useful and policy-relevant approach to addressing pressing issues impacting students' success. The project was a collaboration between the Consortium for Policy Research in Education (CPRE), the partnering school district, and several additional partners, who collaboratively designed and supported the PBIS implementation approach funded from NIJ and supplemental funds from the participating school district.

The goals of the study were to assess the effectiveness of two different interventions with different intensities and resource requirements:

1. Brief trauma trainings for school safety officers. This series of trainings was delivered to school officers in the treatment group during the 2018-2019 and 2019-2020 school years. The impacts of this intervention were assessed via an RCT, where officers from eligible schools were randomly assigned to either receive the trauma trainings or to continue with business-as-usual. Approximately half of the schools in each condition were participants in the partnering school district's PBIS initiative.

2. Expanded climate supports, including intensive whole-staff training and ongoing coaching in trauma and schoolwide PBIS; universal screening for students; and expanded support for Tier 2 implementation. The impacts of this implementation model were assessed via a cluster randomized experiment, with the intervention delivered to classrooms nested in 4 PBIS schools randomly selected from a list of 9 eligible schools. The remaining 5 PBIS schools served as a comparison group.

The research addressed the following objectives:

1. Assess the impacts of trauma-awareness trainings for school safety officers in K-8 schools in the partnering school district on school officers': 1) knowledge and attitudes regarding trauma and trauma-informed practices, and 2) reported behaviors (i.e., use of trauma-informed practices) in their interactions with students.

2. Assess the implementation (including acceptability, appropriateness, and feasibility) and impacts of an expanded, trauma-informed PBIS implementation model in 4 K-5 or K-8 schools on teacher outcomes and student outcomes in the domains of attendance, discipline, and academics, relative to a business-as-usual comparison group of schools.

3. Assess the costs of the expanded, trauma-informed PBIS implementation model components.

1.2. Overview of Expanded, Trauma-Informed PBIS Implementation Model

The expanded, trauma-informed PBIS implementation model (referred to as the "implementation model" throughout this report) integrated trauma-informed approaches with PBIS and included the following components: the collection of school-wide data at the start of the school year (i.e., Baseline School-wide Classroom Observations); individual teacher coaching; a series of trauma-informed PBIS videos for teachers and school staff; universal

5

screening of all students using the Social, Academic, and Emotional Behavior Risk Screener (SAEBRS); expanded support for providing Tier 2 services to students; and individualized coaching for school safety officers. The following section provides a description of each of these components which are referenced throughout this report.

Baseline School-wide Classroom Observations

The baseline school-wide classroom observation component of the model aimed to collect data on teachers' use of specific PBIS classroom practices at the beginning of the school year. Per the implementation model, PBIS coaches conduct classroom observations across all classrooms in the school beginning in late September, and use the data for the following purposes: 1) to inform decision-making about needs for school-wide (e.g., professional development) and individualized (e.g., teacher coaching) support; 2) to provide teachers written feedback about the data collected from their individual classrooms, and 3) to share aggregated school-level data with the whole staff to discuss overall trends. In this project, PBIS coaches used an instrument called the Devereux Classroom Observation Tool (DCOT), developed by one of the project partners. The teacher behaviors included in the DCOT instrument were:

- Delivering "Tokens", Praise Statements, and Corrective Statements
- Providing Opportunities for Students to Respond to Instruction
- Using a Lesson Agenda
- Actively Supervising Students

In addition to these teacher behaviors, the DCOT also included a measure of overall student task engagement.

Individual Teacher Coaching

The individual teacher coaching component was designed to provide support to individual teachers for consistently implementing evidence-based PBIS practices in their classrooms. For this implementation model, teachers are selected for coaching using extant data available to the school as well as the data from the baseline classroom observations. School administrators and coaches are to collaborate in using these data to select teachers for individual coaching. The expectation is that approximately three teachers per school could receive support from the coach during any given time, and that the coaching for each teacher focuses on a specific practice from the DCOT. Per the model, the coach visits the teachers' classrooms once per week to observe and record observational data on the selected skill using the DCOT and the data is then shared back with the teacher via a paper feedback form. The coach also provides strategies or resources for the teacher via a one-on-one meeting to help them improve their implementation, and tracks data from these observations using the DCOT dashboard, an online tool to analyze and archive collected observational data. Following approximately 8-10 of these observations, and in consultation with school administration and the project support team, a determination is made regarding next steps – either ending the coaching relationship, continuing to work with that teacher on the same or different skill, or the school could choose to provide the teacher with more intensive supports.

Trauma-Informed PBIS Videos for Teachers and School Staff

A series of six videos about understanding what trauma is and integrating trauma-informed practices with PBIS were developed to provide teachers and school staff with information and strategies regarding a trauma-informed approach to PBIS. The videos present information about how students' experiences outside of school (such as experiences of traumatic stress) may impact their behavior and emotions in the classroom and provide teachers with

tangible strategies to meet students' needs. The strategies presented in the videos combine

principles of trauma-informed practices with PBIS principles. The titles of the six produced

videos are:

- Fostering Safe, Supportive, Consistent Relationships at School

- Building Connections with Behavior Specific Praise

- Building Regulation Skills in the Classroom

- Managing Emotions in the Classroom

- Regulate, Relate, and Reason

- Empowerment, Voice, and Choice

Along with the videos themselves, a series of documents and resources were created to

support teachers' engagement with the videos. A Teacher Guide was developed for each video to

help teachers reflect on their current practice and distill important content as they viewed the

video. Additional resources (e.g., playlists of music for welcoming students to the classroom;

supplemental ideas for providing opportunities to incorporate student voice) were provided to

support teachers as they implemented the strategies shared in the videos. Per the implementation

model, the videos were designed to be shared with teachers at the rate of one video per month in

large or small groups, and for the coach to facilitate a follow-up discussion during which

teachers could reflect on the content of the video and discuss ideas about how to start

implementing the strategies.

Universal Screener

The universal screening component was designed to provide data regarding needed Tier 1

and Tier 2 supports at the school. The screening instrument used in this project was the Social,

Academic, and Emotional Behavior Risk Screener (SAEBRS). The SAEBRS is a "brief, norm-

referenced tool for screening all students to identify those who are at risk for social-emotional

behavior problems" according to its publisher's website. Per the implementation model,

classroom teachers complete the SAEBRS at two time points during the school year – once in the

early fall and then again in the early spring. According to the model, teachers are trained in how

to complete the SAEBRS and then complete the screener for all students in their class. Following

completion, school teams use the data to address identified needs (e.g., plan additional Tier 1

supports and/or inform the selection of students to receive Tier 2 supports).

Expanded Tier 2 Services

The expanded, trauma-informed PBIS implementation model also includes additional

support for Tier 2 mental health groups, including training for counselors and other relevant staff

in evidence-based mental health group interventions. Per the model, students receive Tier 2

mental health support and are matched to appropriate interventions using data from the SAEBRS

and other available data. According to the model, teachers who had students participating in a

Tier 2 mental health group are expected to complete a daily progress monitoring tool to track that

student's progress over time. An additional feature of this component was expanding an existing

PBIS intervention called Check-In, Check-Out (CICO) which typically focuses on students who

exhibit externalizing behaviors. The implementation model includes a modified version of CICO

with an aim to also provide CICO for students who exhibit internalizing concerns such as

sadness or anxiety.

School Safety Officer Training

This component of the implementation model was designed for PBIS coaches to work

one-on-one with school safety officers throughout the school year to support the officers'

implementation of trauma-informed approaches during their interactions with students. Per the

model, PBIS coaches use a training manual developed specifically for this purpose and meet with officers up to 18 times per school year in 20-minute sessions. Each session in the training manual introduced the officer to trauma-informed skills they were expected to practice with the coach and then use with students with the intention of making their interactions with students more positive and effective.

1.3. Research Questions

This research was designed to answer the following research questions:

RQ1: Assess the impacts of brief trauma-awareness trainings for school safety officers.

1a. Did trainings change officers' reported knowledge and attitudes in the short-term?

1b. What was the impact of the trainings on officers' reported knowledge, attitudes, and behaviors over time?

RQ2: Assess the implementation and impacts of an expanded, trauma-informed PBIS implementation model in four K-5 or K-8 schools.

2a. How acceptable, appropriate, and feasible were the components of the implementation model?

2b. How did teachers in schools that received the implementation model, differ in terms of changes in beliefs, self-efficacy, intentions, and self-reported behavior as compared to teachers in schools that received PBIS as usual?

2c. How did classrooms in schools that received the implementation model differ in terms of changes in attendance, disciplinary outcomes, and academic outcomes as compared to classrooms in schools that received PBIS as usual?

RQ3: Assess the costs of the expanded, trauma-informed PBIS implementation model components.

3a. What resources are required to deliver each component of the implementation model?

3b. What are the associated costs with each component of the model?

1.4. Research Design, Methods, Analytical and Data Analysis Techniques

1.4.1. RQ1: Assess the Impacts of Brief Trauma-Awareness Training for School Safety Officers

Research Design. The impacts of trauma training for school officers were assessed via an RCT of all (N=105) eligible district K-8 schools (including schools that did and did not use school-wide PBIS). The RCT used a blocked design to assign schools to treatment and control conditions within PBIS status (i.e., PBIS school, non-PBIS school), with the officers working in 52 schools randomly assigned to the trauma-awareness training treatment condition and the officers working in 53 schools assigned to the business-as-usual control condition.

Participants. Participants ($N = 97$) most often identified as male (54%), Black or African American (65%), and between the ages of 30 and 60 years (82%).

Procedures. A total of three trainings were delivered to all officers from the schools randomly assigned to the intervention condition.

The research team administered a series of surveys to school officers in the intervention condition before and after each training and at the end of the 2018-19 and 2019-20 school years, for a total of 3 survey administration time points per year. The research team administered surveys to school officers in the control condition at two time points: in September and June of the 2019-2020 school year, at the times that corresponded with survey administration to officers in the intervention condition. In the 2018-19 school year, 32 officers in the intervention condition completed the survey at Time 1, 31 officers in the intervention condition completed the survey at Time 2, and 33 officers in the intervention condition completed the survey at Time 3. In the

11

2019-20 school year, 35 officers in the intervention condition completed the survey at Time 1, 62 officers (38 intervention, 24 control) completed the survey at Time 2, and 48 officers (29 intervention, 19 control) completed the survey at Time 3.

Participants in the two conditions completed the same surveys. All participants received information that the surveys concerned trauma-informed practices that officers may use with students and that each survey would take approximately 10 minutes to complete. Participants voluntarily and confidentially completed surveys administered either via paper survey or electronically using Qualtrics, depending on contact modality (i.e., officers completed all materials electronically in June 2020, when schools were closed at the beginning of the COVID-19 pandemic).

The intervention consisted of a series of three trauma-awareness trainings developed specifically to target the roles and responsibilities of school officers. Training 1 introduced trauma-related concepts and skills (e.g., trauma-sensitive communication, trauma-informed de-escalation) officers could use when working with students who may have trauma histories. Training 2 provided a greater breadth and depth of trauma-related concepts and more intensive focus on skill development. Training 3 served as a booster session one year after the initial training to remind officers of the training content and skills.

Measures. Surveys were used to assess three primary outcome constructs: officers' knowledge (11 items, e.g., impacts of trauma on students' lives), attitudes (6 items, e.g., usefulness of school officers' understanding trauma), and behaviors (12 items, e.g., how often an officer asks students what the officer can do to help them through the day) regarding trauma-related information and skills. Individual response options ranged from 0 to 5 (knowledge: 0 = none at all, 5 = a lot; attitudes: 0 = strongly disagree, 5 = strongly agree; and behaviors: 0 = not

at all, 5 = often). Cronbach's alpha was high (above .925) for each subscale, indicating sufficient internal consistency.

Data Analysis. First, we investigated whether officers' reported knowledge and attitudes changed from pre- to post-training (for each training) by performing a series of repeated measures t-tests on survey constructs; we did not assess changes in behaviors from pre- to post-training, as the frequencies of behaviors performed with students could not change during the course of a single session training. Second, we performed repeated measures ANOVAs across time points to examine changes in officers' knowledge, attitudes, and behaviors over time. Third, we performed an independent samples t-test to examine differences in knowledge, attitudes, and behaviors between the intervention and control conditions at the end of 2019-2020 school year. Although preferable to compare ratings of officers in both conditions over time, the logistical constraints of survey administration to the control group at the same time as administration to officers in the intervention condition prevented reliable analysis and interpretation using such an approach.

1.4.2. RQ2: Assess the Implementation and Impacts of an Expanded, Trauma-Informed PBIS Implementation Model in Four K-5 or K-8 Schools.

Research Design. Impacts of the implementation model were assessed via a cluster randomized experiment. To select schools to receive the expanded supports being provided through this project, the research team collaborated with the partnering school district's school climate administrators and project partners who had been supporting PBIS implementation in the district for several years prior to the start of the project. Criteria were established to determine eligibility for the implementation model, including that eligible schools must have: 1) strong PBIS implementation at Tier 1, 2) attempted implementing some form of Tier 2 supports, and 3)

leadership buy-in. Applying these criteria, nine potential schools were identified as eligible. From this list, four schools (73 classrooms) were randomly selected to receive the implementation model (i.e., Demonstration schools) and the remaining five schools (75 classrooms) served as Comparison schools and conducted business-as-usual PBIS.

The research team collected survey data from K-5 teachers at Demonstration and Comparison schools in the Fall and Spring of the 2019-20 and 2020-21 school years, as well as conducted semi-structured interviews with Demonstration school stakeholders in Spring 2020 and Spring 2021. Survey and interview data from Demonstration school teachers collected each Spring were used to assess the implementation of the expanded, trauma-informed PBIS implementation model (RQ2a). Additionally, survey data from Demonstration and Comparison school teachers across time-points were used to examine changes in teachers' beliefs, self-efficacy, intentions, and self-reported behaviors (RQ2b).

Administrative data was also requested and received from the partnering school district at the classroom level for the four Demonstration and five Comparison schools. These data were used to make comparisons between the two sets of schools on a range of student demographic, attendance, climate, disciplinary, and academic outcomes (RQ2c).

Methods for Evaluating Implementation (Question 2a).

Procedures. In the Springs of 2020 and 2021, K-5 teachers in the Demonstration schools were asked to complete survey questions regarding their experiences with the implementation model (e.g., acceptability of the different model components), and then a select group of teachers, school administrators, and other staff were invited to participate in a qualitative interview. The survey questions were administered as part of the teacher survey (see RQ2b), for which participating teachers provided consent and were compensated with a $25 gift card for

14

their time. Interview participants provided verbal consent to be interviewed and audio-recorded, and participants were offered a $25 gift card for their time. All interviews were audio-recorded and then transcribed by an external agency. Transcripts were reviewed and anonymized by a member of the research team and then were uploaded to an online qualitative data analysis platform, Dedoose, for coding.

Participants. All K-5 teachers in the four Demonstration schools were invited to participate in Surveys regarding their experiences with the expanded, trauma-informed PBIS implementation model in both Spring of 2020 and Spring of 2021. In total 153 surveys were sent, and we received a total of 94 responses, a rate of 61.4%. The response rate for each year's administration was 55.4% for Spring 2020 and 70.5% for Spring 2021.

In the Springs of 2020 and 2021, a total of 51 interviews were conducted with Demonstration school coaches, principals, teachers, and Tier 2 support personnel. The principals of the four Demonstration schools participated in an interview in both years, and all coaches from Demonstration schools participated in interviews in the Spring of each year. Each Spring, 24 teachers, six from each of the four Demonstration schools, were invited via email to participate in an interview. We used purposive sampling to prioritize interviewing teachers who received coaching during the 2019-2020 school year, and otherwise randomly selected teachers from different grade levels to participate. Sixteen teachers (66% response rate, including five of the eight teachers who received coaching) agreed to participate in an interview in Spring of 2020, and 13 teachers (54% response rate) agreed to participate in Spring of 2021.

Quantitative measures. To measure the level of acceptability that teachers reported regarding the components of the implementation model (i.e., the extent to which teachers liked the model components), we used the Acceptability of Intervention Measure (AIM; Weiner et al.,

15

2017). The AIM is a brief, pragmatic measure of acceptability that consists of four items on a 5-point Likert scale (Completely Disagree to Completely Agree); the total score is computed as the mean of the four items. The AIM has shown acceptable reliability (alpha above .82) and test-retest reliability (Pearson correlation above .70; Weiner et al., 2017).

Qualitative interview protocol. A semi-structured interview protocol, developed by the research team, was used to gain an in-depth understanding of the acceptability, appropriateness, and feasibility of the components of the implementation model. The interview protocol was structured around the components of the implementation model with sections for: Baseline school-wide classroom observations, individual teacher coaching, the video series on trauma-informed teaching practices, the universal screening, and expanded support for Tier 2. For each of these components, the protocol also included specific probes regarding overall adoption and, implementation, acceptability and appropriateness, perceived impacts, and barriers and facilitators to implementation. Interview protocols were adapted slightly for each stakeholder group to ensure questions and probes aligned with the participant's role.

Data Analysis.

Quantitative. We used descriptive statistics (i.e., means, standard deviations) to examine teacher-reported acceptability for components of the implementation model. Specifically, descriptive statistics of teacher-reported AIM scores were examined regarding the baseline DCOT, individual teacher coaching, the trauma-informed PBIS videos, and universal screening.

Qualitative. The overall analytic approach for considering the qualitative interview data was guided by an integrated inductive and deductive approach (Bingham & Witkowsky, 2021), using a codebook that members of the research team iteratively developed after initial review of the transcripts and with consideration of the research questions. Specifically, five deductive

16

codes were developed to address constructs of interest (e.g., barriers and facilitators to implementation) for each of the components of the implementation model and for data that referred to the overall implementation of PBIS. Additionally, three inductive codes were added to the codebook based on the interview data: 1) school context, 2) impact of the COVID-19 pandemic, and 3) coach characteristics. Definitions were developed for each code collaboratively by four members of the research team. Before finalizing the codebook, the same two transcripts were then coded by each team member individually, and the team then met to discuss, reconcile codes, and revise code definitions as needed.

The same four members of the research team applied the codebook to code all 51 interview transcripts in Dedoose. Sixty percent of all transcripts were double coded and coding decisions were reconciled through discussion between two research team members (Hill et al., 2005). The remaining transcripts were individually coded. Interview transcripts were coded using a parent code for each component (i.e., DCOT data collection, Individual Teacher Coaching, trauma-informed PBIS Videos, Universal Screener, Tier 2) and then applying child codes corresponding to each of the constructs of interest. DCOT data collection and Universal Screener were not used as codes for Spring 2021 interviews, and the Individual Teacher Coaching code was used to capture general coaching activities during the 2020-21 school year. Excerpt sets for each combination of parent/child codes were then exported for subsequent phases of analysis. Analyses employed data coded as: description of implementation, acceptability and appropriateness, and barriers/facilitators to implementation for each of the implementation model components.

Two members of the research team analyzed the excerpt sets through an iterative process of drafting and revising a series of analytic memos (Saldaña, 2013). The first-stage analytic

17

memo documented an overall description of individual participant responses regarding each component and research question domain. One research team member wrote a series of second-stage analytic memos, which included a summative description of findings for each combination of intervention component/construct (e.g., universal screener – acceptability and appropriateness) across all four schools. This entire process, beginning with the translation of raw excerpts to the first-stage memo, was then vetted by a separate member of the research team. The final memos were discussed as a team prior to converting the memos to the findings shared below.

Mixed method integration. We used a sequential explanatory approach (Quant→QUAL) to mixing quantitative and qualitative data (Creswell & Plano Clark, 2018). Qualitative data from the semi-structured interviews were used to provide additional depth and richness to the quantitative information about teachers' perceptions of acceptability of the implementation model components.

Methods for Evaluating Impacts on Teacher Survey (Question 2b).

Participants. The teacher survey was administered to teachers in both Demonstration and Comparison schools in Fall and Spring of the 2019-20 and 2020-21 school years. At each time point, all K-5 teachers in the Demonstration and Comparison schools were invited by email to participate in the voluntary, online survey. Across both timepoints, 301 surveys were sent, and we received a total of 180 responses, a rate of 59.8%. The response rate for each year's administration was 62.3% (60.6% Demonstration; 63.8% Comparison) for Fall 2019, 53.2% (55.4% Demonstration; 50.7% Comparison) for Spring 2020, 70.8% (71.9% Demonstration; 70.0% Comparison) for Fall 2020, and 68.2% (70.5% Demonstration; 65.3% Comparison) for Spring 2021.

Measures. The teacher survey measured teachers' beliefs about trauma, perceptions of trauma-informed behaviors and systems in their school, self-efficacy for behavior management, and intentions to use specific teaching practices aligned with PBIS and trauma-informed practice. The instrument included five different measures:

1) **Trauma-Informed Beliefs**. This scale was adapted from the Trauma Informed Belief Measure (Brown, Baker & Wilcox, 2012) and was intended to measure teachers' beliefs about trauma. All 10 items were scored on a scale from 1 to 5, where a score of 5 indicates the respondent holds what are considered trauma informed beliefs, while a score of 1 indicates they do not. Six items were reverse-scored according to scale guidance. This scale showed excellent internal consistency (Cronbach alpha at all timepoints above 0.80) during the current administrations (Cronbach, 1951; Cortina, 1993).

2) **Perceptions of Trauma-Informed Behavior in the School**. This scale was adapted from the Staff Behavior in the Treatment Program measure (Brown, Baker & Wilcox, 2012) and aimed to measure teachers' perceptions of trauma-informed behaviors among the teaching staff at the school. This scale showed excellent internal consistency (Cronbach alpha at all timepoints above 0.84) during the current administrations.

3) **Trauma-Informed Systems.** This scale was adapted from the Trauma Informed Systems Change Instrument – Individual Practice (Richardson et al., 2012) and was intended to be a measure of teacher perceptions of their practice in a trauma-informed system. There were three items on the scale scored on a scale from 1 to 5. This scale

showed excellent internal consistency (Cronbach alpha at all timepoints above 0.85) during the current administrations.

4) **Self-Efficacy for Behavior Management.** This scale drew items from the Behavior Management subscale of the Teacher Sense of Self-Efficacy Scale (Tschannen-Moran & Woolfolk-Hoy, 2001) and was intended to be a measure of teacher self-efficacy for behavior management. There were 5 items on the original scale scored on a scale from 1 to 9, where a score of 9 indicates that the respondent feels very confident in their ability to manage behaviors in the classroom without being overwhelmed. An additional item aligned with the scope of the implementation model (i.e., item 5: "How much can you do to build safe, consistent, supportive relationships with your students?") was included here and was also scored on a scale from 1 to 9. This scale showed excellent internal consistency (Cronbach alpha at all timepoints above 0.91) during the current administrations.

5) **Teacher Intentions.** These items used validated stems (Ajzen, 1991) to measure teachers' self-reported intentions to use specific, trauma-informed, PBIS aligned teaching practices (e.g., "Take steps to build safe, consistent, supportive relationships with my students") during the next two weeks. Items were scored on a scale from 1 to 7, where a score of 7 indicated that the respondent believes they are likely to implement the practice and 1 indicated that the respondent believes they are very unlikely to do so. This scale showed excellent internal consistency (Cronbach alpha at all timepoints above 0.91) during the current administrations.

Data Analysis. Composite scores were calculated for each scale by averaging item-level scores. Descriptive statistics were examined to summarize findings for each scale. Paired sample

t-tests for each scale were performed to look for evidence of any significant differences between the Demonstration and Comparison groups at baseline.

The primary analyses examined whether there were any statistically significant differences between treatment groups in the means of the questionnaire scores over time (four time points). First, we evaluated the nested structure of the data by fitting an unconditional model and calculating the amount of variance in teachers' outcomes explained by which school they were in. The Intraclass Correlation coefficients varied between 2.6% (Trauma-Informed Systems) and 13.5% (Self-Efficacy for Behavior Management), suggesting meaningful clustering by school (Kreft & de Leeuw, 1998). Therefore, we considered teachers nested within schools for the primary analyses.

Hierarchical Linear Model (HLM) is frequently used in education research, primarily because of the need to take aggregation levels into account to understand the differences observed between individuals (teachers in this study) (Goldstein, 2003). Further, HLM models can be used to investigate change over time across individuals (Singer & Willet, 2003; Snijders & Bosker, 2012). Particularly, longitudinal three-level data (i.e., repeated outcome measurements over time, Level 1; sampled from various individuals, such as teachers, Level 2; nested within organizational structures, such as schools, Level 3) can be analyzed by implementing HLM. Therefore, for this analysis, we used three-level hierarchical linear models, one for each outcome of interest, with survey waves nested within teachers, which are in turn nested within schools. Traditional maximum likelihood estimation methods for multilevel modeling can lead to biased estimates when the number of clusters -schools and teachers- is below 30. Here, we implemented restricted maximum likelihood (REM) estimation method, which has shown to perform well with ten clusters or fewer (McNeish & Stapleton, 2016).

21

Evaluating a Replicable PBIS Implementation Approach in Philadelphia

Equation 1 specifies the full model that was used to examine the effects.

$$Level\ 1: Y_{ijk} = \beta_{0jk} + \beta_{1jk}TIME_{ijk} + \epsilon_{ijk}$$

(1)

$$Level\ 2: \beta_{0jk} = \gamma_{ook} + \gamma_{ook}Treatment_j + \mu_{0jk}$$

$$\beta_{1jk} = \gamma_{1ok} + \gamma_{1ok}Treatment_j$$

$$Level\ 3: \gamma_{00k} = \delta_{000} + v_{ok}$$

$$\gamma_{10k} = \delta_{100}$$

$$\mu_{0j} \sim N(0, \sigma_\mu^2), \epsilon_{ijk} \sim N(0, \sigma_\epsilon^2), and\ v_{0k} \sim N(0, \sigma_v^2)\ all\ independent,$$

where Y_{ijk} is the ith observation for the jth teacher in the kth school, $\beta_{ijk} ... \beta_{njk}$ refers to the fixed effect (slope) of the TIME level variables ($TIME_{IJk}$), γ_{00k} denotes an average intercept value at the initial time point for each kth school, and μ_{0jk} is a teacher-specific (random) effect on the intercept.

Methods for Evaluating Impacts on Student Outcomes (Question 2c).

Participants. Administrative data was provided by the school district for all K-5 students, aggregated at the classroom level, within the nine participating schools (4 Demonstration, 5 Comparison) for the academic years 2016-17, 2017-18, 2018-19, 2019-20, and 2020-21.

Measures. The school district administrative data provided a range of variables about students, aggregated at the classroom level, including demographics, attendance, climate, disciplinary outcomes, and academic outcomes. The student attendance variable received from the school district indicated the percentage of students with chronic absenteeism, defined as missing 10 or more days during a school year. Students' disciplinary outcomes include the number of Out-of-School suspensions (OSS), number of Office Discipline Referrals (ODRs),

22

count of serious incidents, and the number of recorded bullying reports within each school year. We also received data from the school district from their district-wide survey which measures several constructs related to school improvement. From the school climate construct, we utilized two sub-constructs (Bullying and Belonging) in our analyses. Students' classroom academic outcomes comprise course grades in any English, Math, Science, or Social Studies subject.

Data Transformation. Three of the variables received from the school district (i.e., student grades and the two sub-constructs from the district-wide survey) were transformed prior to the analysis in order to answer research question 2c. For each classroom, we received the number of students receiving each letter grade – A, B, C, D, and F. We used these data to compare the proportion of low-achieving students in each subject by combining the number of students receiving either a D or an F into one category. This "low-achieving" variable was computed using only classrooms for grades 1 through 5, because there were no students receiving grades of D or below among kindergarten students. We also received data from the school climate construct of the school district's district-wide survey. For each of the two climate sub-constructs, we obtained the total number of students, within a classroom, who answered each question. For every item, we also obtained the number of students answering each possible response (Most or all the time, Occasionally, Rarely, and Never). These data were used to calculate sub-construct scores. To that end, we followed the school district's guidelines. First, each possible response was assigned a numeric value, with the most positive response assigned the highest value (10) and the least positive the lowest (0). Later, these values were added for all the items in the sub-construct. Finally, the sum was divided by the total number of survey items among each sub-construct.

Data Analysis. In response to data pertinence and COVID-19 pandemic-imposed limitations, some data were excluded from the present analysis. Mainly, analyses of suspensions and disciplinary outcomes excluded any data points collected after the third quarter of the 2019-20 school year, to avoid any discrepancies due to different COVID-19 school guidelines. Further, school district data reported zero ODRs during 2016-17 and 2017-18 school years. Additionally, no ODRs were recorded during the 2020-21 school year. To account for these data restrictions, ODR outcomes were analyzed using yearly impact analyses for the period of 2018-19 and 2019-20 school years.

Paired sample t-tests for key student-level demographic variables (i.e., sex, ethnicity, Economic Disadvantage status, English Language Learner status, and Special Education status) were performed to determine whether there was baseline equivalence between the Demonstration and Comparison groups using pre-intervention data (i.e., from the 2017-18 and 2018-19 school years). In addition to this test, standardized mean difference Hedges' g, which allows us to account for the small sample of schools included in this study, was calculated to examine the effect size difference between the groups (Hedges, 1981). According to the What Works Clearinghouse Procedures and Standards Handbook, baseline equivalence is defined as differences less than or equal to a Hedges' g effect size of 0.05 standard deviation (What Works Clearinghouse, 2022).

The primary analyses examined evidence of any statistically significant differences between Demonstration and Comparison groups, aggregated at the classroom level, in each relevant outcome over time. For this analysis, we used three-level hierarchical linear models, one for each outcome of interest, with outcome waves nested within classrooms, which were in turn nested within schools. Traditional maximum likelihood estimation methods for multilevel

modeling can lead to biased estimates when the number of clusters (i.e., schools and classrooms) is below 30. Therefore, we instead implemented restricted maximum likelihood (REM) estimation method, which has been shown to perform well with ten clusters or fewer (McNeish & Stapleton, 2016).

Equation 2 specifies the full model that was used to examine the effects.

$$Level\ 1: Y_{ijk} = \beta_{0jk} + \beta_{1jk}TIME_{ijk} + \epsilon_{ijk}$$

(2)

$$Level\ 2: \beta_{0jk} = \gamma_{ook} + \gamma_{ook}Treatment_j + \mu_{0jk}$$

$$\beta_{1jk} = \gamma_{1ok} + \gamma_{1ok}Treatment_j$$

$$Level\ 3: \gamma_{00k} = \delta_{000} + v_{ok}$$

$$\gamma_{10k} = \delta_{100}$$

$$\mu_{0j} \sim N(0, \sigma_\mu^2), \epsilon_{ijk} \sim N(0, \sigma_\epsilon^2), and\ v_{0k} \sim N(0, \sigma_v^2)\ all\ independent,$$

where Y_{ijk} is the ith observation for the jth classroom in the kth school, $\beta_{ijk} ... \beta_{njk}$ refers to the fixed effect (slope) of the TIME level variables ($TIME_{IJk}$), γ_{00k} denotes an average intercept value at the initial time point for each kth school, and μ_{0jk} is a classroom-specific (random) effect on the intercept. When appropriate, log-transformation was used to improve the relative normality of the data.

1.4.3. RQ3: Assess the Costs of the Expanded, Trauma-Informed PBIS Implementation Model Components

To inform future implementation, the cost study examined the resources that would be required to enact the expanded, trauma-informed PBIS implementation model described in this report. The research questions guiding this work were: 1) What resources are required to deliver

25

each component of the implementation model? 2) What are the associated costs of those resources? To address these questions, we applied the ingredients method (Levin et. al, 2018) to examine the resources and their costs of delivering each program component to additional schools.

There were two main goals for this work. First, we aimed to inform the partnering school district's future delivery of PBIS and school climate strategies, especially through an expansion of the schools that receive these supports. Second, we aimed to inform school systems more broadly about the resource requirements or costs involved in providing this framework of approaches and supports.

The cost study is a district-level analysis that provides cost information on each component of the implementation model. To inform stakeholders broadly across the country, we used average U.S. prices reflected in 2022 constant dollars. When a resource's cost varied by the number of schools served by a district, we estimated the total cost for serving 10 schools implementing PBIS. Two of the components require multiple years of time and investment to reach fidelity. Thus, following the design of these components, school officer training and the universal screener costs are provided across multiple years.

We collected data through observation of team meetings and interviews with project partners at the Devereux Center for Effective Schools, Drexel University, the University of South Florida, and Jefferson University. Interviews were iterative in nature and focused on the resources required to deliver each component of the program with fidelity according to design.

1.5. Expected Applicability of the Research

Positive Behavioral Interventions and Supports (PBIS) is a behavioral framework that has been demonstrated to improve school climate by reducing behavioral incidents and improving

26

students' emotional and academic outcomes (Bradshaw, Mitchell, & Leaf, 2010; Madigan, Cross, Smolkowski, & Strycker, 2016; Horner, Sugai, Fixsen, 2017). PBIS has been adopted widely, both within the partnering school district and nation-wide. However, large urban school districts often face challenges implementing PBIS consistently and addressing the extensive needs of their student populations and school communities (Gray, Sirinides, Fink, Flack, DuBois, Morrison, & Hill, 2017; Molloy et al., 2013).

In response to this identified need, the current project has developed and evaluated a replicable approach to implementing PBIS within the context of the partnering school district. The project assesses the impacts of the implementation model with a rigorous approach and provides crucial insights regarding implementation and impacts. The larger RCT isolates the innovate training to support the work of school safety officers including increasing positive and beneficial interactions with students. The evaluation of the implementation model in the four Demonstration schools provides other districts and schools key findings around implementation, including barriers and facilitators, potential impacts on teachers' knowledge and self-efficacy, and disciplinary and academic outcomes for students.

Lessons learned from this research are expected to be applicable to inform similar implementation efforts in other large, urban school districts aiming to improve school climate and reduce negative disciplinary actions. Additionally, the components of this project related to school safety officers – the trainings and manual developed and implemented – have the potential to influence how school safety officers are trained and supported, addressing the need to bridge the gap between training and practice by scaffolding learning to help officers develop new skills and use them in their day-to-day work with students. Lastly, the work in this project

Evaluating a Replicable PBIS Implementation Approach in Philadelphia

has continued to be built upon and expanded to many more schools by our partners at the school district.

2. Participants and Other Collaborating Organizations

The research team that carried out this study was led by Dr. Ryan Fink, Ed.D., Senior Researcher at the Consortium for Policy Research in Education (CPRE) at the University of Pennsylvania's Graduate School of Education and Dr. Gwendolyn Lawson, Assistant Professor of Psychiatry at the Children's Hospital of Philadelphia (CHOP) and University of Pennsylvania's Perelman School of Medicine. Dr. Lawson collaborated closely with CPRE regarding planning and executing the implementation and effectiveness evaluations of the project. Dr. Fink and Dr. Lawson were supported by colleagues Dr. Karina Diaz, Katarina Suwak, and Tesla DuBois at CPRE. The initial grant was awarded to PI Dr. Abigail Gray who left CPRE in August 2019. Dr. Philip Sirinides served as the project's methodologist until his departure from CPRE in December 2019.

The design and implementation of the supports provided to schools throughout this project were a result of a close collaboration between CPRE and the partnering school district, as well as a number of other collaborating organizations. Leadership and staff from the partnering school district, particularly the offices that support school climate and mental health prevention and intervention, were integral to the project's success. Other partnering organizations included: 1) the Devereux Center for Effective Schools, 2) Drexel University's Juvenile Justice Research and Reform Lab, 3) Jefferson University; and 4) the University of South Florida.

The Devereux Center for Effective Schools collaborated on the classroom observation and teacher coaching aspects of the project, including providing training and support for PBIS coaches to use the Devereux Classroom Observation Tool and to develop skill guides for

28

classroom management supports for school PBIS teams. Devereux also supported the development of the trauma-informed PBIS video series and accompanying resources.

Drexel University's Juvenile Justice Research and Reform Lab (JJR&R) team led the training for school safety officers in K-5/K-8 treatment schools. This team also led the related data collection, analysis, and presentation of the findings related to those trainings. The JJR&R Lab also developed the Positive School Safety Curriculum and trained PBIS coaches to implement the program with school safety officers in the four Demonstration schools. Partners at the University of South Florida were responsible for supervision of the universal screening consultation and implementation, as well as creating support materials for SAEBRS data interpretation and conducting meetings with school leaders. Trauma-experts at Jefferson University, contracted by the partnering school district, supported the development and implementation of the trauma-informed PBIS video series and accompanying resources. These experts also provided specialized training in trauma and trauma-informed practices for Demonstration school PBIS coaches throughout the project.

3. Changes in Approach From Original Design and Reason for Change

3.1. Change from QED Design to RCT

The initial proposal planned for a quasi-experimental design, in which all classrooms in the four Demonstration schools would be matched to comparison classrooms from within 22 PBIS schools which were to be randomly assigned to the business-as-usual SWPBIS condition in the larger RCT using propensity score analysis. A total of 80 demonstration classrooms, from the four Demonstrations schools, would then be matched to 80 comparison classrooms within the 22 PBIS schools.

To select the four Demonstration schools, a total of 9 PBIS schools were identified as eligible based on a set of criteria including readiness and leadership buy-in. Because the pool of PBIS schools identified as eligible was larger than the required Demonstration group of four, schools were randomly assigned to Demonstration and Comparison from within this group. A randomized block designed approach was used to minimize the effects of systematic error. The two binary blocking variables were derived for two partnering school district survey scale outcomes: climate and process. The study was then carried out based on a sample of nine schools (four Demonstration, five Comparison). The sample includes a total of 148 teachers, 75 Demonstration and 73 Comparison.

Given that Demonstration and Comparison schools were selected randomly from the pool of eligible schools, this study was treated as a cluster randomized trial, where schools are randomly assigned to a treatment group and the unit of study is teachers/classrooms. This design supports strong causal inferences about program impacts by avoiding bias, achieving balance of confounding factors between treatment groups, and providing valid comparisons of intervention strategies. This change does not impact the initially proposed research questions.

3.2. Changes Due to the COVID-19 Pandemic

As a result of school closures and virtual schooling beginning in March of 2020 and continuing through the 2020-21 school year, it was necessary to make adaptations to the expanded, trauma-informed PBIS implementation model during this time. In March of 2020, schools closed for in-person instruction due to the COVID-19 pandemic and did not reopen for the remainder of the school year. This abbreviated school year affected implementation during the 2019-2020 school year in a number of ways. First, all six of the trauma-informed PBIS videos were not shared with schools. Second, individual teacher coaching was underway, however no teachers had completed a full cycle of coaching. Third, the SAEBRS universal

30

screening was only completed one time by classroom teachers in the Fall of 2019. Fourth, Tier 2 mental health groups were in progress, however no groups were able to be implemented in their entirety. Finally, support for non-instructional staff including school safety officers was unable to be completed.

The 2020-21 school year began completely virtual for all of the partnering school district's schools and remained virtual until March 2021 when some schools reopened with an option to attend in person, although few students elected to attend in person. As a result, the following impacts on implementation occurred during the 2020-2021 school year: baseline classroom observation data collection was not conducted; individual teacher coaching was not implemented as designed; the SAEBRS universal screening instrument was not administered; Tier 2 mental health groups were not administered in some schools; and no school safety officers were trained. These changes are accounted for and discussed in the sections that follow.

Additionally, the planned cluster RCT comparing classrooms in the Demonstration schools to classrooms in the Comparison schools was designed to rely on observational data collected by our research team as well as administrative data acquired from the partnering school district following the 2019-2020 school year. As a result of school closures beginning in March 2020, our research team was unable to conduct classroom observations as planned. Additionally, school district administrative data was not available following the closure of in-person schooling in March 2020. Given the incomplete implementation of the intervention prior to school closures and the lack of complete data, we requested and received approval for a change of scope to shift the timeline for the matched comparison QED analysis until following the 2020-2021 school year. Following a second extension request due to a delay in receiving administrative data from the school district, a revised end date for this project was set for September 30, 2022.

3.3. Change in Scope From Original RCT Intervention

The original proposal for the study included two separate evaluations of impact. The first component was a quasi-experiment (later shifted to a cluster RCT) to examine the impacts of a set of expanded, trauma-informed PBIS implementation supports on student and classroom level outcomes in the partnering school district. The second component of the research, as originally planned, was an RCT including approximately 160 schools to examine the impacts of two interventions on the treatment schools: 1) brief trauma training for teachers; and 2) trauma training for school safety officers. We submitted a change in the scope of this portion of the project in July of 2019. The change in scope eliminated the brief trauma training for teachers and instead focused the RCT on the trauma training for school safety officers. The reason for this proposed change was that, since the time of the award, the partnering school district expedited its plans to train all teachers in trauma to meet schools' requests for this training. The partnering school district prioritized training for all teachers in trauma, which eliminated the availability of a control group of schools for the RCT.

4. Outcomes

4.1. Activities/Accomplishments

The scope of this project included the design and development, implementation, and evaluation of a varied set of approaches and resources to improve school climate and reduce negative disciplinary incidents in K-5 classrooms in the partnering school district. This work was organized within two main areas: 1) the development of approaches to training school safety officers to become more trauma-informed in their interactions with students, and 2) the development of a set of expanded trauma-informed implementation supports to strengthen PBIS implementation.

The development of approaches to training school safety officers to become more trauma-informed in their interactions with students was led by Drexel University's Juvenile Justice Research & Reform Lab in collaboration with the other partners. The team developed and delivered a set of trainings for school safety officers to help them understand the impacts of trauma on students in their schools and how they can interact with students in more positive ways. Additionally, a manualized program for working more intensively with school safety officers was developed and implemented. This provided a foundation for future efforts to expand the manual to other school climate staff and to implement it more widely.

The expanded, trauma-informed implementation model developed for the four Demonstration schools included both newly-developed as well as existing resources and approaches. The 6-part video module series and its accompanying implementation supports were created entirely during this project via a collaboration between CPRE, CHOP, Devereux, Jefferson, and the partnering school district. These videos have now been embedded into the district's virtual professional development platform and continue to be viewed by teachers and staff across the district. The Devereux Classroom Observation Tool (DCOT) was developed prior to the start of this project, but the project allowed for coaches and other climate staff in the district to be trained on using the DCOT to conduct classroom observations, and to analyze and share that data with school teams. Use of this tool also allowed observed teachers to receive feedback regarding their implementation of various PBIS practices and some also received individualized support from a PBIS coach. This project also represented an initial effort in the school district to have schools complete and use a universal screening instrument. For instance, during project activities, project partners improved processes and methods for training teachers and using universal screening data; the participating school district has continued to build on that

experience as they expand the use of universal screening instruments to more schools. Through the project, Demonstration school counselors and other school staff received training in the Tier 2 intervention Calm Cat along with the implementation materials to conduct that intervention in their schools. During school closures due to the COVID-19 pandemic, Calm Cat video trainings were developed and continue to be utilized by the district's staff.

4.2. Results and Findings

4.2.1. RQ1: Assess the Impacts of Brief Trauma-Awareness Training for School Safety Officers

Results. Officers' reported knowledge improved from pre- to post-training for each training in the 2018-2019 school year (see Table 1). For example, trauma knowledge ratings increased significantly from before the first training (M = 3.15, SE = 0.15) to after the first training (M = 4.31, SE = 0.10), $t(24) = -5.79$, $p < 0.01$). Knowledge improved significantly from pre- to post-training with four of the five trainings provided during this study—all except training 2 in year 2. See Table 1.

In 2018-2019, officers' reported attitudes improved significantly from pre- to post-training with Training 1, but they did not improve significantly from pre- to post-training with any of the other four trainings provided. Importantly, though, mean attitude scores remained higher at all future time points than they were at pre-training at Time 1 in Year 1. See Table 1 and Figure 1.

Table 1

Changes in Officers' Reported Knowledge and Attitudes Following Trainings

	Pre-training		Post-training		*t*	95% CI	
	M	*SE*	*M*	*SE*		LL	UL
8/22/2018 Training 1							
Knowledge	3.15	0.15	4.31	0.10	*t*(24)=-5.79*	-1.57	-0.75
Attitudes	4.11	0.11	4.72	0.08	*t*(27)=-6.90*	-0.80	-0.43
11/6/2018 Training 2							
Knowledge	3.67	0.15	4.10	0.16	*t*(20)=-3.14*	-0.71	-0.14
Attitudes	4.51	0.11	4.49	0.11	*t*(25)=0.22	-0.13	0.17
8/27/2019 Training 1							
Knowledge	3.34	0.34	4.18	0.22	*t*(6)=-2.47*	-1.68	-0.07
Attitudes	4.30	0.19	4.18	0.23	*t*(7)=0.51	-0.46	0.71
8/27/2019 Training 3							
Knowledge	3.82	0.18	4.14	0.14	*t*(17)=-3.68*	-0.50	-0.14
Attitudes	4.53	0.13	4.51	0.13	*t*(17)=0.24	-0.18	0.22
9/27/2019 Training 2							
Knowledge	3.73	0.38	3.82	0.16	*t*(4)=-0.28	-0.98	0.80
Attitudes	4.37	0.27	4.43	0.22	*t*(6)=-0.33	-0.48	0.37

* *p* < .05

Figure 1

Changes in Officers' Reported Attitudes, Knowledge, and Behaviors: 2019 – 2019

Changes in Officers' Reported Attitudes, Knowledge and Behaviors: 2018-2019
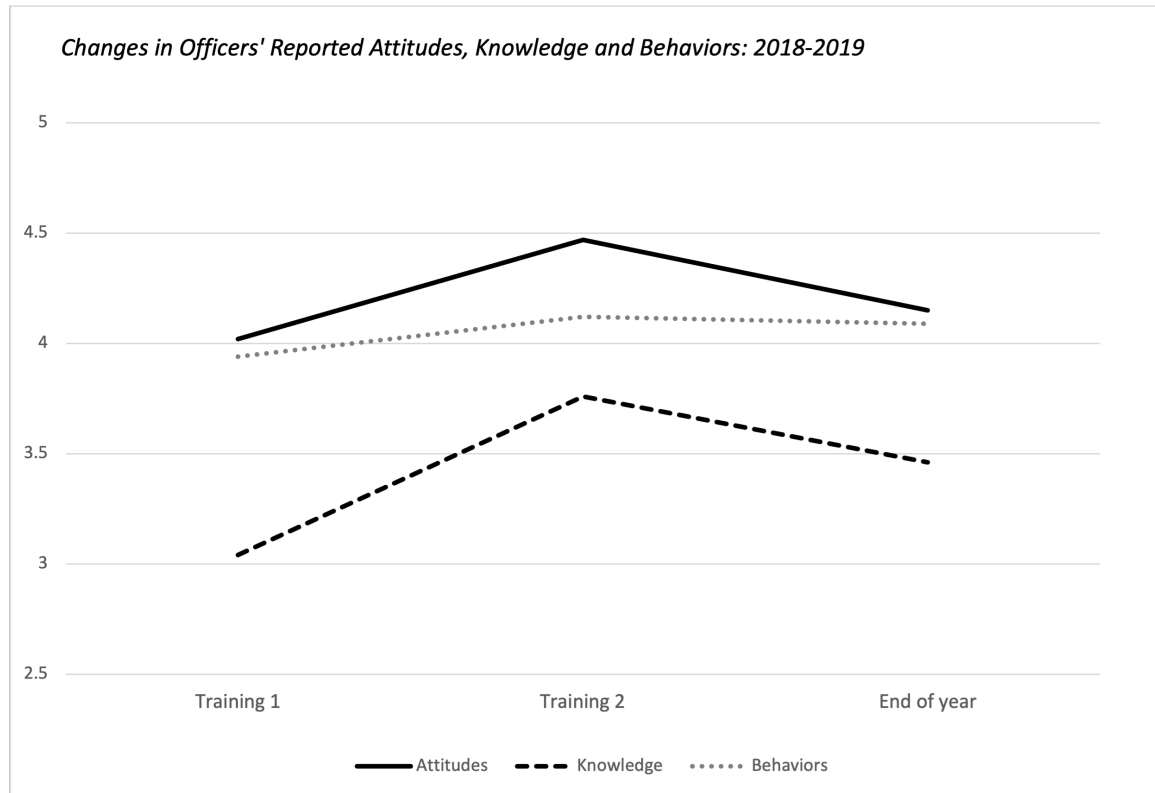
Analyses of knowledge, attitudes, and behaviors among officers in the intervention condition who participated in trainings in the 2018-2019 school year revealed significant positive change in officers' attitudes over the course of one school year $F(2, 36) = 4.82$, $p < 0.05$. See Figure 1. However, we found no significant change in knowledge, $F(2, 26) = 2.91$, $p > 0.05$, or behaviors, $F(2, 24) = 0.29$, $p > 0.05$. Although new 2019-2020 intervention officers completed surveys that provided data about knowledge, attitudes, and behaviors, too few officers (n = 9) were added to the intervention schools during this school year to reliably or meaningfully analyze the changes in the three outcomes across Year 2 of the study.

Additionally, comparing across conditions, at the conclusion of the final training or equivalent time period, officers in the intervention condition did not report more favorable knowledge ($M = 3.84$, $SE = 0.17$), $t(34) = -1.56$, $p > 0.05$; $r = .26$), attitudes ($M = 4.36$, $SE = 0.14$), $t(36) = 0.35$, $p > 0.05$; $r = .06$, or behaviors ($M = 4.15$, $SE = 0.13$), $t(34) = 0.24$, $p > 0.05$; $r = .38$) than those officers in the control condition. See Table 2 for these results.

Table 2

Intervention and Control Group Differences at the Conclusion of All Training Activities

	Control		Intervention		t	95% CI	
	M	*SE*	*M*	*SE*		LL	UL
Knowledge	3.43	0.16	3.84	0.17	$t(34)=-1.56$	-0.95	0.13
Attitudes	4.43	0.12	4.36	0.14	$t(36)=0.35$	-0.34	0.48
Behaviors	4.20	0.14	4.15	0.13	$t(34)=0.24$	-0.36	0.46

*$p < .05$

Discussion. School safety officers' reported knowledge regarding trauma and trauma-informed practices in schools generally improved from pre- to-post training, and attitudes improved at the initial training. Additionally, changes in attitudes improved significantly throughout the school year, suggesting the value of trauma and trauma-informed training for generating departmental culture change. Officers reported use of trauma-informed practices with students did not change, however, with their participation in this training. These findings are consistent with extant research indicating that trainings are often useful mechanisms to convey factual information and influence attitudes and beliefs, but they are far less successful in influencing behaviors (e.g., Beidas & Kendall, 2010; Frank et al., 2020). More in-depth skills

training, including ongoing coaching, may be necessary to meaningfully change officers'
behaviors with students and help them put knowledge into practice.

4.2.2. RQ2: Assess the Implementation and Impacts of an Expanded, Trauma-Informed PBIS Implementation Model in Four K-5 or K-8 Schools

Implementation Results (Question 2a). These findings are organized by component of
the implementation model, including findings regarding the extent to which implementation
occurred as intended, and the perceived acceptability and appropriateness of each component.

Baseline School-wide Classroom Observations Using the DCOT.

Description of Implementation. Interview participants reported that the baseline DCOT
observations took place in each of the four participating schools in Fall 2019, but not in Fall
2020 because schools at that time were operating virtually due to the pandemic. Based on Fall
2019 observations, the extent to which DCOT data were utilized varied between schools.
Although individual feedback forms were distributed to teachers in all schools in 2019, only two
of the four schools provided an opportunity for the aggregated school-level data to be shared
with the entire staff. One of these two schools considered the data as a whole group and decided
on which of the observed skills they, as a staff, needed to focus on to make improvements.
Interview respondents from two schools recalled that the data were also used by coaches and
administrators to identify teachers for individual or small group teacher coaching and additional
supports.

Acceptability and Appropriateness. On the teacher survey administered in Spring 2020,
27 Demonstration teachers indicated being aware of the baseline DCOT and those teachers
reported a mean acceptability score of 3.37 ($P = .86$; out of a possible score of 1-5) regarding the
baseline DCOT component of the implementation model.

In interviews, most respondents across the four schools described the practice of a coach collecting data via classroom observation at the start of the school year as acceptable. Interview participants noted several reasons for finding the baseline DCOT to be an acceptable practice including allowing the coach and school administration to gain an overall sense of what was happening inside of classrooms at the beginning of the school year. Teachers also reported liking that a coach, rather than a principal or administrator, conducted the observations. Teachers shared that they viewed the coach as a supportive resource and as a source of specialized knowledge in classroom management and student engagement, and therefore did not perceive the observation as evaluative. Teachers also appreciated the opportunity to receive feedback and improve their practice. For example, some teachers shared that the observation data encouraged them to become more intentional about how frequently they were offering praise to students.

Although the level of acceptability of this component was high overall, a few teachers mentioned aspects of the baseline DCOT which they did not find acceptable – mainly related to a lack of clarity about the purpose and timing of the observations. For example, one teacher noted that they would have preferred to know the purpose of the observation in advance so that they could have a chance to prepare. Other teachers mentioned that the coach came to their classrooms to observe at inconvenient times, and they believed that this timing influenced what the coach saw. Relatedly, several teachers shared that they would have liked to receive notification about the planned observation or have some input into when the observation took place.

Individual Teacher Coaching.

Description of Implementation. Coaches established individual coaching relationships with teachers across the four participating schools during the 2019-20 school year; however, the

39

individual teacher coaching component was implemented with varying levels of consistency. All coaching ended in March of 2020 due to the COVID-19 pandemic and minimal individualized coaching took place during the 2020-21 school year due to the virtual teaching format. Teachers were selected for coaching in the 2019-20 school year using the baseline DCOT data collected by the coach, as well as input from school administration. In one school, the principal contacted the identified teachers to inform them that they would be receiving coaching, while in the other schools, initial communication came directly from the coach.

In some cases, the coach collaborated with the selected teachers to identify the skill to focus on, and in other cases, the coach selected the skill they thought would be most impactful for that teacher. Some of the teachers who were coached reported in interviews that the coach would come into the classroom and observe and take part in the class by modeling the identified skill or interacting with students, while other teachers reported that the coach observed in the classroom and left feedback via paper or email but did not engage in any skill modeling or follow-up discussion.

During the 2020-21 school year, when schools were operating predominantly virtually due to the COVID-19 pandemic, the implementation of individual teacher coaching occurred differently to accommodate the virtual format, in which classroom observations and coaching in virtual classrooms were perceived as disruptive. Instead of implementing individual teacher coaching as initially intended, teachers, coaches, and principals reported in interviews that coaches supported teachers in other ways. Mainly, they helped to prepare and lead professional development for teachers around PBIS implementation, supported teachers in running daily community meetings for their students, and occasionally provided individualized support to a few teachers across the four schools. During the first few months of the school year, coaches

supported teachers to lead daily community meetings by conducting observations and providing resources and supports for teachers. At one school, this took the form of the coach working with three or four teachers individually over several weeks to improve the teachers' implementation of community meetings and PBIS strategies in the virtual teaching and learning environment.

Acceptability and Appropriateness. On the teacher survey administered in the Spring of 2020, eight Demonstration teachers indicated that they received individual teacher coaching, and those teachers reported a mean acceptability score of 3.75 ($SD = .81$; out of a possible score of 1-5) regarding the individual teacher coaching component of the implementation model. On the Spring 2021 survey, 23 teachers reported receiving coaching (reflecting the type of support described above) and those teachers reported a mean acceptability of 3.38 ($SD = .65$ out of a possible score of 1-5).

In interviews, respondents overall found individual teacher coaching in the 2019-20 school year to be an acceptable practice, although teachers varied in their enthusiasm for receiving coaching. Interview respondents liked that the coaches reduced caseload of schools allowed them to have an opportunity to work with individual teachers. Some principals, for example, described this affordance as addressing an important need to provide more intensive supports to some teachers regarding classroom management. One coach reported that the opportunity to support individual teachers was their favorite part of their job. The coaching process (i.e., coaches observe, leave feedback, and meet with teachers to discuss when possible) was also generally acceptable to teachers who received the coaching. Some teachers shared that they found this coaching helpful and valuable; other teachers reported that it was acceptable, and they appreciated the additional support, however, they did not believe they gained much from the

41

coaching. Scheduling limitations on the part of both the teachers and coaches was the primary concern cited by respondents who found this component to be less acceptable.

During the 2020-21 school year, teachers overall found the support from coaches to be highly acceptable in that they perceived the coach as being available to support them and their interactions with their coach to be helpful. No teachers explicitly reported that they perceived the coaching to be unacceptable, though some stakeholders thought it was less effective in the virtual setting. Additionally, one coach perceived teachers to be less receptive to coaching than they would be in the in-person setting. Despite the differences in the way that coaching was implemented in the mostly virtual school year, teachers found this component of the implementation model acceptable.

Trauma-Informed PBIS Video Series.

Description of Implementation. Stakeholders reported in interviews that the series of videos about trauma-informed PBIS were made available to all teachers during the 2019-20 school year; they reported that teachers were typically invited to view the videos individually, and less frequently had opportunities to view and discuss them in groups. The coaches shared that grade group meetings, where they had originally planned to view and discuss the videos with teachers, were already dedicated to other priorities. Instead, the coaches and project team developed a strategy to share the videos with teachers via email. Coaches sent links via email to teachers throughout the school year and encouraged them to watch the videos on their own time; additionally, coaches offered times for "watch parties" where teachers could come together to watch the video as a group and participate in discussion. To encourage more teachers to watch the videos, an incentive was offered (an opportunity to win a gift card via a raffle) for teachers who clicked the link and reported watching the videos. Data from the teacher survey show that

42

over 75% of teachers reported watching at least two of the videos during the 2019-20 school year.

During the 2020-21 school year, there were more opportunities to view and discuss the videos in large and small groups due to an increase in district-allocated professional development time. Coaches, principals, and teachers reported that coaches presented the videos, often along with supporting slides and discussion points, to school staff during professional development time that had not been available in previous years. In addition to showing the videos and facilitating discussion, teachers reported that the coaches followed-up via email with additional resources, including linking to an online library of resources developed by the coaches and project team.

Acceptability and Appropriateness. On the teacher survey administered in the Spring of 2020, 38 Demonstration teachers indicated that they watched one or more videos; those teachers reported a mean acceptability score of 3.87 (*SD* = .50 out of a possible score of 1-5). On the Spring 2021 survey, 35 teachers reported watching at least one video and those teachers reported a mean acceptability of 3.84 (*SD* = .61 out of a possible score of 1-5).

According to analysis of interviews conducted during the 2019-20 school year, the content of the videos was largely perceived as acceptable, although some teachers reported that the unrealistic scenarios made them less acceptable. Additionally, some teachers reported that the lack of opportunities to discuss and process the videos with colleagues made their experience with them less acceptable. In interviews, teachers who watched at least one or more videos expressed that they found the content to be helpful and relevant given the needs and experiences of their students. Although the videos aimed to demonstrate strategies teachers could implement

43

in their classrooms, some teachers still expressed a desire for more ways to apply the knowledge about how trauma affects their students.

Teachers who watched the videos on their own often expressed a desire to have the opportunity for discussion with their colleagues to process the content of the video and share ideas about ways to implement the strategies. When teachers did have the opportunity to discuss the videos in a group, they shared that this opportunity for discussion was the aspect of this component that they liked the most and found most useful. One teacher shared,

> The discussion with the video was really good. I think that's better than when you just look at it yourself with no feedback. The discussion piece is really helpful especially when you start hearing the thoughts of others, and then you can start sharing viewpoints, and that might even change your perspective. "Okay, well you know what? I didn't think of it like that, but okay." But if you took that and you watch it yourself, you could tend to go into your own little shell on the world and be like, "Okay, you know what? It's just one more," with nobody you could bounce off of."

A few teachers also mentioned that they did not like that the videos were "staged" – meaning they were not filmed in actual school district classrooms and did not portray real-life scenarios. Even for teachers who identified this as an issue, most still reported that the videos were worthwhile and overall acceptable.

Teachers also found the videos to be acceptable during the 2020-21 school year in which instruction was mostly virtual. Though some teachers again reflected that the scenarios depicted in the videos were not applicable to the virtual setting, most teachers found the content and strategies to be useful and appropriate for their students' exposure to trauma (both from their

44

neighborhoods in general as well as regarding the pandemic). In addition to the content of the videos, teachers again found value in the discussions that took place following whole- or small-group viewing of the videos. One teacher shared, "I found [the videos] very useful. The language, how to approach the student, I found that useful. Then we had discussions later on after the video...in the beginning I was a little reluctant, but once I watched a video and I could relate to some of the scenarios and see how they would handle, I really appreciated it – the exposure - and really giving me what it looks like and how to handle it."

Many teachers who participated in interviews talked about the value of these discussions that took place, and found that having these conversations with their colleagues, even across grades, helped them to better understand how to take what they learned from the videos and put it to use in their own classrooms. One teacher shared:

> I really like that aspect. I like watching them and having the discussions because it's an eyeopener, or some things you're aware of, but then you kind of forget because day to day goes on and that kind of makes you sit and reflect on what that module or the trauma is. I do like having that refresher and having those discussions, like I said, cause I'm kindergarten, but people of different, having discussions with different grades, different people that are part of this [school] team, they have different views, have different experiences. That part I really do like. And I feel like it's effective.

Two coaches shared that they perceived the videos to be a good fit for their role. One coach reflected on how the trauma-informed aspect of the videos helped to ground the PBIS work that the school was engaging in and that having a trauma-informed approach and using the videos was one of their favorite aspects of their role. Another coach also commented on how the

trauma-informed lens fit well with their own career background and was "very necessary" for their school, as well as something that they enjoyed doing as part of their coach role.

Universal Screening.

Description of Implementation. Each school completed one administration of the SAEBRS universal screening instrument during the 2019-20 school year, although there was variation in the extent to which the screening data were used by school teams. Each school's staff received training regarding how to complete the screener, either directly from a project partner overseeing the implementation of the screener, or from school district coaches, including the PBIS coach, supporting the school. Some schools also allocated time for teachers to complete the screener following this training, while other schools did not allocate that time. Of the 42 teachers who completed the teacher survey in Spring 2020, 86% (i.e., 36/42) recalled being asked to complete the screener. Respondents at one school reported using the screening data to identify and place students in appropriate Tier 2 interventions; however, most other interview respondents reported not being aware of how the data were being used. The universal screening component of the implementation model was not carried out during the 2020-21 school year due to the format of virtual schooling.

Acceptability and Appropriateness. On the teacher survey administered in the Spring of 2020, 36 Demonstration teachers indicated that they received training on the universal screener and those teachers reported a mean acceptability score of 3.53 (*SD* = .9; out of a possible score of 1-5) regarding the universal screening component of the implementation model.

Interview participants found the training and completion of the SAEBRS to largely be acceptable. However, they expressed overall dissatisfaction with this component because they perceived no follow-up or sharing of the data after the screener was completed. Most teachers

46

shared in interviews that the training and completion of the screener was an acceptable practice for them. They felt the training adequately prepared them to complete the screener and they liked that the screener was able to be completed online. Additionally, teachers also liked when they were allocated time to complete the screener. Those teachers who reported that they perceived the universal screening process as unacceptable typically mentioned that they found it tedious or uninformative.

Expanded Tier 2 Services.

Description of Implementation. In the 2019-20 school year, some schools began to implement Tier 2 mental health groups for students prior to school closures but those groups were interrupted by the pandemic; no schools reported using the modified version of Check-In, Check-Out (CICO). Counselors and/or supervised counseling interns at each school received training in at least one Tier 2 mental health group in January of 2020 and three of the four schools reported enrolling students in the intervention for which they had been trained prior to March 2020. However, these schools were not able to complete the full intervention prior to pandemic-forced school closures. All schools reported using CICO, although none reported using the "modified" version for students with internalizing concerns.

In the 2020-21 school year, all four schools utilized some form of CICO, and two of the four schools were able to implement mental health groups virtually. Check-In, Check-Out was generally less structured when implemented in the virtual space. For example, none of the schools had teachers complete the student cards which are traditionally used to monitor a student's progress towards behavior goals throughout the school day. Instead, it was often a less formal arrangement where a school counselor or other climate support staff would check-in with a particular student as needed. The behaviors for which students were referred to CICO also

differed from previous years. The behavior issue most frequently identified in this virtual school year was attendance and/or students not keeping their cameras on during instruction. Two of the four schools reported conducting virtual mental health groups using the Calm Cat intervention.

Acceptability and Appropriateness. Interview respondents during the 2019-20 school year found the Tier 2 interventions, including both mental health groups and CICO, to be acceptable and appropriate interventions for their schools, although some teachers expressed dissatisfaction regarding the task of completing CICO point cards throughout the day. Most interview respondents stated that the Tier 2 interventions offered at their school were acceptable to them and they believed the interventions were necessary and appropriate for their students, however, teachers expressed doubts about the effectiveness of the CICO cards and the consistency with which these cards were completed.

Overall, stakeholders found the expanded Tier 2 aspect of the implementation model to be acceptable during the 2020-21 school year. For the most part, CICO and mental health groups were seen as welcome and beneficial for the students who were involved. One coach perceived the way CICO was implemented at their school to be unacceptable because they believed the school attempted to include too many students in the program. However, a teacher from this school did report CICO to be acceptable for both them and their students who were enrolled.

Overall Feasibility, Barriers, and Facilitators to Implementation. Across components of the expanded, trauma-informed PBIS implementation model, five factors emerged as barriers and/or facilitators to implementation. The extent to which a particular factor hindered or facilitated implementation varied by component as well as by school and school year. Specifically, time, staff buy-in, training and implementation support, staff turnover, and

48

alignment with other priorities, were identified as key factors that impeded and/or facilitated implementation across components and schools.

Time to engage with the implementation model was consistently identified as both a barrier and facilitator to implementation of the different model components across schools. For example, teachers found it helpful when time was allocated in their schedules to engage with a particular intervention component (such as completing the SAEBRS or watching and discussing the videos) and challenging when time was not allocated. Relatedly, staff buy-in (e.g., teacher buy-in to engage with coaching or Tier 2 staff buy-in to implement the group interventions) was identified as both a barrier and a facilitator, and some interview participants noted that staff were more willing to allocate time to the intensive PBIS model when there was strong buy-in about the importance of the model. The availability (or lack of availability) of training and implementation support was identified as another key barrier and facilitator that varied between model component and school; for example, Tier 2 staff identified training and materials as an important facilitator to implementing Tier 2 groups, while a perceived lack of training about interpreting universal screener data was identified as a barrier. Additionally, staff turnover was identified as a barrier to implementation at some schools. For example, two of the four schools implementing the intervention experienced turnover of their principal during the two years of implementation. This caused certain routines to be changed and priorities to be reset at the school-level and took additional time for the project and school support teams to gain buy-in from the new principal. Finally, the alignment between the implementation model and other school initiatives and identified student needs (e.g., the extent to which the Tier 2 interventions were perceived as meeting student needs) was identified as an important barrier and facilitator across components and schools.

Impacts on Teacher Survey (Question 2b).

Descriptive. Teachers self-report survey responses were collected during the 2019-20 and 2020-21 school years. In the 2019-20 school year, 86 teachers (40 Demonstration, 46 Comparison) from nine schools completed the survey in the Fall and 90 teachers (42 Demonstration, 48 Comparison) completed the survey in the Spring. In the 2020-21 school year, 90 teachers (41 Demonstration, 49 Comparison) completed the survey in the Fall and 86 teachers (40 Demonstration, 46 Comparison) completed the survey in the Spring.

Table 3 presents the demographic characteristics of the teachers who participated in both the Fall 2019 and Spring 2020 administration of the survey. Most teachers were female (91%) and white (76%); this was consistent across the Demonstration and Comparison groups. On average, teachers had approximately 13 overall years of teaching experience and approximately six years of teaching experience at their current school.

Table 3

Fall 2019 & Spring 2020 Teachers Survey - Demographics

Teachers Demographics	Demonstration	Comparison	Total
Female	30	38	68
Male	2	3	5
Prefer Not to Say	1	1	2
Black	8	5	13
White	25	32	57
Prefer Not to Say	2	2	4

Baseline Equivalence. Paired sample t-tests found no statistically significant differences at baseline between the Demonstration and Comparison group in three of the five scales measured by the survey (Trauma-informed Beliefs, Trauma-informed Behavior, and Trauma-informed Systems). However, two scales showed statistically significant baseline differences between the Demonstration and Comparison group. As shown in Table 4, schools in the Comparison group reported higher baseline levels of Self-efficacy for Behavior Management (t =-2.33, *df* =84, *p* = 0.022) and on Intentions to use specific, trauma-informed, PBIS aligned teaching practices (t =-2.13, *df* =84, *p* = 0.036).

Table 4

Descriptive Statistics by Scale

	Demonstration		Control			
	Pre	Pre SD	Pre	Pre SD		
	μ^1	σ	μ^1	σ	t(84)	p
Trauma-informed beliefs	3.48	0.40	3.52	0.45	-0.53	0.598
Trauma-informed behavior	3.46	0.55	3.53	0.45	-0.67	0.502
Trauma-informed systems	3.53	0.92	3.69	0.70	-0.89	0.379
Self-efficacy for behavior management	6.37	2.25	7.26	1.18	-2.33	0.022*
Intentions	5.84	0.78	6.13	0.49	-2.13	0.036*

Teacher Self-Report Outcomes. We then explored potential impacts of the intervention on teachers' beliefs and perceptions of trauma-informed behaviors and systems over time (four time points), by fitting hierarchical linear models (HLM) for each teacher outcome. Time (level 1

51

Evaluating a Replicable PBIS Implementation Approach in Philadelphia

variable) is considered nested within teachers (level 2 variable), which in turn are nested within schools (level 3 variable).

HLM models, in contrast to the repeated measures ANOVA approach, do not assume that all individuals - teachers in this study - have complete data for the same number of waves. Thus, the impact analysis included a total of 75 teachers who completed at least three of the four survey administrations. This approach reduces the amount of drop-out bias while ensuring a retention indicator (the number of responded follow-ups) and a last response factor (the timing of last response) of 0.75 (Wærsted et al., 2018). Table 5 presents model results.

The effect of interest in these models is the Time*Treatment interaction, which indicates if and how scores for the two treatment groups (i.e., Demonstration and Comparison) differ across time. Results showed no significant interaction effects across most survey scales. However, there was a significant Time*Treatment interaction for the Self-efficacy for Behavior Management scale. Although both Demonstration and Comparison group teachers showed decreased scores over time on this scale, there was a positive Time*Treatment interaction, indicating that the decrease among teachers in the Demonstration group was smaller compared to the decrease among teachers in the Comparison group (see Table 5).

Table 5

Impact analysis results 2019-2021 Survey Administrations

Scale	Effect	Estimate	SE	95% CI		p
				LL	UL	
Trauma-informed beliefs	Fixed effects					
	Intercept	3.557	0.063	3.433	3.681	<0.001*
	Time	0.079	0.018	0.043	0.115	<0.001*
	Treatment (Demonstration)	-0.098	0.093	-0.280	0.083	0.2903
	Time*Treatment	0.005	0.027	-0.049	0.058	0.0854

52

Trauma-informed behavior	Fixed effects					
	Intercept	3.56	0.096	3.371	3.748	<0.001*
	Time	0.06	0.022	0.019	0.107	0.005
	Treatment (Demonstration)	-0.14	0.142	-0.420	0.135	0.3401
	Time*Treatment	0.01	0.033	-0.056	0.075	0.7692
Trauma-informed systems	Fixed effects					
	Intercept	3.816	0.124	3.572	4.059	<0.001*
	Time	0.055	0.034	-0.012	0.122	0.1096
	Treatment (Demonstration)	-0.395	0.182	-0.751	-0.038	0.0608
	Time*Treatment	0.087	0.051	-0.013	0.187	0.0883
Self-efficacy for behavior management	Fixed effects					
	Intercept	7.663	0.257	7.159	8.166	<0.001*
	Time	-0.555	0.082	-0.714	-0.395	<0.001*
	Treatment (Demonstration)	-0.854	0.378	-1.594	-0.113	0.0487
	Time*Treatment	0.280	0.121	0.043	0.517	0.0216*
Intentions	Fixed effects					
	Intercept	6.288	0.092	6.107	6.468	<0.001*
	Time	-0.078	0.032	-0.140	-0.016	0.015*
	Treatment (Demonstration)	-0.178	0.135	-0.443	0.086	0.189
	Time*Treatment	0.070	0.047	-0.022	0.163	0.139

Impacts on Student Outcomes (Question 2c).

Descriptive. The school district administrative data provided information for a total of 137 K-5 classrooms (64 Demonstration, 73 Comparison) during the 2019-20 academic year and 130 K-5 classrooms (58 Demonstration, 72 Comparison) during the 2020-21 academic year. Table 6 presents baseline demographic students characteristics, aggregated at the classroom level, all of which were found to be equivalent.

Table 6

Students Demographics at Baseline

Students Demographics	Demonstration	Comparison	Total
Female	49%	49%	49%
Male	51%	51%	51%
Black	64%	63%	63%
Hispanic	30%	28%	29%
White	2%	3%	3%
Other	4%	6%	5%
Economic Disadvantage	83%	82%	83%
EL	10%	10%	10%
SPED	12%	14%	13%

Potential Impact on Students' Outcomes. We explored potential impact on students' outcomes over time by fitting hierarchical linear models for each academic and disciplinary outcome. These three-level models consider Time (level 1 variable) to be clustered within Classrooms (level 2 variable), which are clustered within schools (level 3 variable). The impact analysis included a total of 511 classrooms (240 Demonstration, 261 Comparison) over the 2016-17, 2017-18, 2018-19, 2019-20, and 2020-21 school years.

As with teachers' outcomes, the effect of interest in these models is the Time*Treatment interaction. This effect indicates if and how students' outcomes for the two treatment groups (i.e., Demonstration and Comparison) differ across time. Results suggest no significant

54

interaction effects across most academic, disciplinary, attendance, and climate survey outcomes. Results of models for each individual outcome measure are summarized in Table 7. We did find a significant Time*Treatment interaction for the social studies' academic outcome. For this outcome, Demonstration group students showed decreased proportion of students scoring withing the low achievement category over time (i.e., obtained F or D course grade).

To account for data restrictions discussed in the previous section, Office Discipline Referrals (ODRs) were analyzed using separate yearly hierarchical models for the 2018-19 and 2019-20 school years. Table 8 summarizes these results. We found that Demonstration classrooms tend to record a lower number of ODRs than Comparison classrooms across time, however, as suggested by model results, this difference is not statistically significant.

Table 7

Impact analysis results administrative data

Outcome	Effect	Estimate	SE	95% CI		p
				LL	UL	
Academic						
Math	Fixed effects					
	Intercept	2.277	0.460	1.376	3.178	<0.001*
	Time	2.119	0.223	1.682	2.555	<0.001*
	Treatment (Demonstration)	-0.817	0.656	-2.104	0.469	0.226
	Time*Treatment	0.174	0.325	-0.463	0.811	0.592
Science	Fixed effects					
	Intercept	0.468	0.598	-0.704	1.641	0.4509
	Time	1.207	0.174	0.867	1.548	<0.001*
	Treatment (Demonstration)	-1.158	0.881	-2.884	0.569	0.2188
	Time*Treatment	0.452	0.253	-0.043	0.948	0.0742
English	Fixed effects					
	Intercept	1.580	0.428	0.740	2.420	<0.001*
	Time	2.413	0.230	1.963	2.863	<0.001*
	Treatment (Demonstration)	-0.433	0.607	-1.622	0.756	0.4758
	Time*Treatment	0.147	0.335	-0.509	0.803	0.6600
Social Studies	Fixed effects					
	Intercept	0.395	0.457	-0.501	1.290	0.4037
	Time	1.331	0.163	1.011	1.651	<0.001*
	Treatment (Demonstration)	-0.218	0.667	-1.524	1.089	0.75
	Time*Treatment	-0.502	0.238	-0.968	-0.037	0.0349*

Table 7. *Impact analysis results administrative data (continued)*

Disciplinary

OSS	Fixed effects					
	Intercept	1.289	0.332	0.639	1.939	0.0049
	Time	-0.065	0.016	-0.096	-0.034	<0.001*
	Treatment (Demonstration)	-0.682	0.494	-1.650	0.285	0.2062
	Time*Treatment	0.038	0.023	-0.007	0.084	0.0973
Bullying	Fixed effects					
	Intercept	0.097	0.040	0.018	0.177	0.0265*
	Time	-0.006	0.005	-0.014	0.003	0.2160
	Treatment (Demonstration)	0.076	0.059	-0.039	0.191	0.2102
	Time*Treatment	-0.008	0.006	-0.021	0.005	0.2096
Serious Incidents	Fixed effects					
	Intercept	1.678	0.145	1.263	2.229	<0.001*
	Time	0.968	0.030	0.913	1.026	0.2687
	Treatment (Demonstration)	0.887	0.216	0.581	1.354	0.5933
	Time*Treatment	1.014	0.043	0.932	1.103	0.7470

Attendance

Chronic Absenteeism	Fixed effects					
	Intercept	13.500	1.111	11.32	15.678	<0.001*
	Time	-1.033	0.242	-1.508	-0.558	<0.001*
	Treatment (Demonstration)	-0.542	1.650	-3.776	2.693	0.7508
	Time*Treatment	0.590	0.351	-0.098	1.278	0.0933

Table 7. *Impact analysis results administrative data (continued)*

District-Wide Survey

Bullying	Fixed effects					
	Intercept	8.397	0.180	8.044	8.750	<0.001*
	Time	-0.582	0.074	-0.728	-0.437	<0.001*
	Treatment (Demonstration)	0.083	0.243	-0.393	0.559	0.7330
	Time*Treatment	-0.104	0.103	-0.306	0.097	0.3120
Belonging	Fixed effects					
	Intercept	7.146	0.205	6.745	7.547	<0.001*
	Time	0.096	0.065	-0.031	0.224	0.1390
	Treatment (Demonstration)	0.008	0.288	-0.557	0.572	0.9790
	Time*Treatment	-0.055	0.090	-0.232	0.122	0.5440

Table 8

Impact analysis results yearly ODRs

Scale	Effect	Estimate	SE	95% CI		p
				LL	UL	
ODRs 2018-19	Fixed effects					
	Intercept	13.699	0.579	12.56	20.628	0.003*
	Treatment (Demonstration)	-0.268	0.868	-1.43	10.649	0.173
ODRs 2019-20	Fixed effects					
	Intercept	9.266	0.511	8.266	10.267	0.003*
	Treatment (Demonstration)	-0.521	0.764	0.116	2.019	0.422

58

4.2.3. RQ3: Assess the Costs of the Expanded, Trauma-Informed PBIS Implementation Model Components.

This project took place within the context of a school district where an existing PBIS model was already being delivered. The expanded, trauma-informed PBIS implementation model described here builds on existing PBIS implementation in the partnering school district. Here we describe the added costs for the expanded, trauma-informed implementation model including specially trained PBIS coaches who work 0.5 full-time equivalent (FTE) per school, or as one full-time coach supporting two schools per school year. This cost study builds on this staffing structure and does not include the costs of recruiting, hiring, and scaling up PBIS coaches, as the costs associated with this process can largely vary by district needs. Another factor to consider while hiring or scaling up is overhead costs, or the district's costs to hire new personnel outside of basic salary. These can include healthcare, benefits, employment taxes, physical space, and training. Districts should be prepared to bear overhead costs that are 25-40% of the base salary (Weltman, 2019).

Implementation Model Components and Costs

Positive School Safety Program (PSSP) for School Officers. The Positive School Safety Program is a manualized approach to support school safety officers in implementing trauma-informed and PBIS approaches during their interactions with students. The manual is designed to be delivered to officers in schools by a trained coach. The manual used in the project described in this report consisted of 18 interactive sessions, each lasting approximately 15-20 minutes and designed for implementation over the course of a single school year; however, following implementation in this project's demonstration schools, several sessions were combined to avoid redundancy, and the PSSP now consists of only 16 sessions. Although the PSSP was carried out

59

by PBIS coaches in the project described in this report, the Juvenile Justice Research & Reform (JJR&R) Lab at Drexel University found that the PSSP could be more effectively implemented by officer coaches. "Officer coaches" refer to current school safety officers who become trained as PSSP coaches to provide peer-to-peer PSSP coaching to other school officers, so those officers can implement the program in their respective schools. Compared to PBIS coaches, officer coaches have more familiarity and experience with the roles, responsibilities, requirements, and challenges of school safety officers, enhancing their ability to relate to officers and provide meaningful coaching. Therefore, the following discussion presents an implementation model that relies on officer coaches rather than PBIS coaches.

The activities and hours detailed in Table 9 reflect the costs borne by the district to implement the fully developed PSSP including time for district leaders to oversee the program implementation, as well as coaching and training of officer coaches. It does not include costs for the time for officers to receive the coaching, because that would take place during their regular duties.

Training and implementation of the PSSP for the district is overseen by a senior level person from the district's office responsible for overseeing school safety. This role includes collaborating with an Implementation Lead and Project Director from Drexel University's JJR&R Lab. Cost estimates are based on an average hourly rate of $180 per hour, accounting for differences in salary levels of the Senior SD Point, Implementation Lead, and Project Director. These costs reflect training and supporting officer coaches to implement the PSSP.

The full implementation of PSSP described here took two school years. The first year of implementation focused on training officer coaches and preparing for the launch of the PSSP program that took place the following year and included seven components. The first component

60

involved planning for local implementation, with a cost of $7,200, and included meetings between Drexel and district administration, understanding the local role of school officers, and adapting materials for local use. The second component involved recruiting, interviewing, and selecting existing officers to serve as PSSP coaches, with an estimated cost of $3,600. The third component is a one-time stipend of $400 each for 30 officers ($12,000) who have been selected to be trained as coaches. The third, fourth, and fifth components included three types of training that officer coaches completed – coaching skills training, trauma training, and PSSP curriculum training, which cost $17,910. A breakdown of cost by type of training can be found in Table 9. The final component involved 20 sessions of ongoing support for officer coaches, which was provided by Drexel and cost $4,500. The total estimated cost for the first year of implementation (i.e., the preparation and training year) was $45,210

The second year of implementing PSSP had seven components that focused on the implementation stage of the PSSP. The first component was a district leadership information session to provide district leadership with a plan for the coming school year, which cost about $720. The subsequent components in the second year included a refresher training, facilitated by Drexel University's JJR&R Lab for officer coaches on the manual ($2,520); it also included the launch of the PSSP program and trauma training for school safety officers ($2,520). Monthly check-ins with the officer coaches were also included in the second year ($9,360), as well as biweekly check-ins with school district leadership ($4,500). There was also a mid-year professional development for school safety officers provided by Drexel's JJR&R Lab ($5,760). Finally, the Year 2 cost estimate also included Drexel's JJR&R Lab providing administrative support to the district to increase program fidelity, including support for tracking session completion, sending emails, and providing updates to partners throughout the school year

($10,500). The total cost associated with the second year of PSSP was $36,180.

Of note, since the completion of this project's implementation, Drexel's JJR&R Lab, in collaboration with the partnering school district, developed a plan to reduce implementation time for the PSSP to one school year (i.e., by providing coaching training during professional development days and PSSP implementation in schools throughout the same school year). At the time this report was written, the abbreviated implementation approach had just begun, and, thus, associated cost estimates for the abbreviated implementation timeline were not included in the current project.

The cost of the manual, workbook, and training templates were estimated at a flat cost of $15,000 for grades K-8 and $15,000 for upper grades, 9-12. Although the material for both age groups covers the same skills and conceptual session content, they differ in that each version scaffolds examples and scenarios to match the experiences and interactions appropriate for their grade level. Because the manual and workbooks are digitalized, the estimates do not include the costs of printing materials. The overall cost for implementing PSSP would cost a district $111,390.

The implementation of the PSSP may require minor revisions based on the number of officer coaches, level of training and support needed, and timely updates to program materials. Additional costs of program implementation beyond those detailed below may include costs for officer coach leads, officer coaches, and officer trainees, should a district decide to establish a different compensation structure for work related to the PSSP.

Table 9

Ingredients for PSSP

Activities	Year	Staffing	Administrative Time (Person Hours)	Execution Time (per activity)	Execution Time (Person Hours)	Total Time	Cost
Planning for local implementation (e.g., meetings with school district, understanding local role of school officers, adapting materials for local use)	1	Senior SD Point, Implementation Lead (Drexel), Project Director (Drexel)	40	0	0	40	$7,200
Recruitment, Interview, and Selection of Officer Coaches (= existing officers)	1	Senior SD Point & Implementation Lead	5	15	15	20	$3,600
Stipend for Officer Coaches	1	30 Officer Coaches	Flat Rate of $400 per year, per coach				$12,000
Coaching Skills Training for officer coaches	2	Project Director, Implementation Lead & Senior SD Point	3	4	12	15	$2,700
Coaching Trauma Training	1	Project Director, Implementation Lead & Senior SD Point	2	2.5	7.5	9.5	$1,710
Coaching Curriculum Training - 10 sessions	1	Project Director, Implementation Lead & Senior SD Point	30	15	45	75	$13,500
Coaching Skill Hours - 20 sessions	1	Implementation Lead	5	20	20	25	$4,500
Leadership PSSP Info Session	2	Project Director, Implementation Lead & Senior SD Point	1	1	3	4	$720
Coaching Skills Refresher	2	Project Director, Implementation Lead & Senior SD Point	2	4	12	14	$2,520
PSSP Program Launch & Trauma Training	2	Project Director, Implementation Lead & Senior SD Point	5	3	9	14	$2,520
Monthly Coaches' Meetings	2	Project Director, Implementation Lead & Senior SD Point	7	15	45	52	$9,360
Biweekly Leadership Check-Ins	2	Implementation Lead & Senior SD Point	5	10	20	25	$4,500
Midyear Professional Development	2	Project Director, Implementation Lead & Senior SD Point	30	2	2	32	$5,760
Administrative Support to Increase Fidelity to Program (e.g., tracking session completion, sending emails/checking in with coaches, providing updates to partners)	2	Implementation Lead & Senior SD Point	40	20	20	60	$10,800
		Totals				**386.5**	**$81,390**

To purchase Digital Access and Unlimited Copies to K-8 Package (Manual & Workbook) $15,000

To purchase Digital Access and Unlimited Copies to 9-12 Package (Manual & Workbook) $15,000

Final Cost   **$111,390**

Trauma-informed PBIS Videos. The 6-part video series was created to provide teachers and school staff with background knowledge and tangible strategies for implementing a trauma-informed approach to PBIS. Each video shares information about how exposure to traumatic events may impact students' behavior and emotions in the classroom, while also providing teachers with strategies to address students' needs.

The video series was used to support PBIS coaches in integrating a trauma lens to the existing PBIS framework. Jefferson Trauma Education Network, housed at Jefferson University, provides trauma training to support coaches in working with teachers via three components: an implementation guide, the trauma-informed PBIS video series, and the provision of ongoing coaching support.

The resources involved in this trauma training are shown in Table 10 and discussed in detail below. PBIS coaches and other relevant staff first receive the developed implementation guide to provide an understanding of the purpose and intended uses of the video series, and to offer additional strategies for integrating trauma-informed information and practices into their work with teachers. Each of the six videos have been embedded into a self-guided learning module and, therefore, there is no timeline on completing the trauma training. The implementation guide, video modules, and accompanying resources are currently digitally available for all districts at no cost. All accompanying handouts and resources are digitalized, and the estimates do not include printing costs. The only cost associated with the trauma training is optional ongoing support for coaches. Jefferson offers consultation with a trauma expert at an hourly rate of $200. The extent to which districts emphasize or require use of the videos and associated materials, including purchasing ongoing consultation with Jefferson experts, can vary. Ongoing support for coaches could include 2-hour monthly trainings that: (1) introduce each

64

module and available resources; (2) reinforce relevant trauma-sensitive principles and provide

rationale for novel strategies; (3) introduce strategies prioritized in the module; and (4) support

coaches to plan for implementation. As an estimate, a district with 10 schools receiving trauma

training for the first time would require 20 hours of support throughout the school year ($4,000).

Table 10

Ingredients for Trauma Training

Activities	Staffing	Executive Time (Person Hours)	Hourly Rate	Cost
Implementation Guide	Not needed	-	-	$0
Video Modules	Not needed	-	-	$0
Ongoing support for coaches (optional)	1 Trauma consultant	20	$200	$4,000

Total cost of ingredients: $4,000

Note: Costs reflect average U.S. prices in constant 2022 dollars, rounded to the nearest ten.

Devereux Classroom Observation Tool. The Devereux Center for Effective Schools

provides training and ongoing support to PBIS coaches on how to use the Devereux Classroom

Observation Tool (DCOT) to collect baseline school-wide classroom observational data

regarding the implementation of several PBIS strategies, as well as to inform individual teacher

coaching. Collecting these data for all classrooms at baseline allows for the intentional

deployment of Tier I supports as well as to assist with identifying teachers who might benefit

from individual teacher coaching. The activities and costs discussed in the following paragraphs,

therefore, address both the baseline school-wide classroom observations and the individual

teacher coaching components of the implementation model.

Table 11 displays the cost estimate for the DCOT and includes the following

components: training of PBIS coaches, a fidelity matching process, training on how to use the

DCOT dashboard, and ongoing support for PBIS coaches.

Devereux's model includes a 3-hour foundational training for PBIS Coaches on how to use the DCOT for both school-wide baseline data collection, as well as for individual teacher coaching. The training cost ($450) would remain the same regardless of how many PBIS coaches were being trained. The DCOT tool and feedback forms are included in the cost of the training and are provided digitally; however, the cost for printing these materials is not included.

Devereux also provides a one-time, 1-hour training on the DCOT dashboard ($150) to assist coaches with using this online tool to analyze and archive collected observational data. A 3-hour fidelity matching process is also included in the costs and involves conducting double-observations to assure reliability with using the DCOT tool. The fidelity-matching process takes about three hours on average; however, this could fluctuate depending on how long it takes the individual to reach reliability. Assuming there are 5 PBIS coaches in a district, the annual cost of fidelity matching is $2,250.

Devereux provides ongoing support and supervision for PBIS coaches in their individualized teacher coaching. The support covers reviewing the data using the data dashboard, creating evidence-based decisions, and offering resources and strategies to support teachers. This support is typically about one hour per month per coach ($7,500).

The total cost of ingredients to use the DCOT for baseline school-wide classroom observational data and individual teacher coaching is $10,350. Assimilating the DCOT in different school districts may require minor revisions to the level of ongoing monthly supervision or fidelity matching required to implement this component. The costs provided in Table 11 estimates for a district that has 10 schools (therefore has 5 PBIS coaches) who will each receive 1 hour of training each month. Additional factors such as teacher compensation are not included in the estimate.

Table 11

Ingredients for DCOT

Activities	Staffing	Administrative & Executive Time (Person Hours)	Hourly Rate	Cost
DCOT training for PBIS coaches	Multiple Devereux experts	3	$150	$450
Fidelity matching (3 hours per PBIS coach)	Multiple Devereux experts	15	$150	$2,250
DCOT dashboard-specific training (one-time)	Devereux expert	1	$150	$150
DCOT tool & feedback forms*	-	-	-	$0*
Ongoing support for coaches (1 hour per coach month)	Multiple Devereux experts	50	$150	$7,500
			Total cost of ingredients:	**$10,350**

Notes: Costs reflect average U.S. prices in constant 2022 dollars, rounded to the nearest ten. Only digital access to the material is included; the estimates do not include printing costs.

Tier 2 Supports. Part of the expanded, trauma-informed PBIS implementation model includes training and support for Tier 2 mental health group interventions. These evidence-based interventions are carried out by school-based counselors and mental health providers to support Tier 2 mental health groups. One of the several interventions for Tier 2 mental health groups is Calm Cat, a group-based relaxation skills training for elementary students with anxiety. Devereux provides training and dissemination of the Calm Cat curriculum to districts for a fee of $1,500 (Table 12). This cost covers the 3-hour recorded training and unlimited digital access to required materials such as the implementation manual, the student activity packets, and the daily progress monitoring tool. Because the materials are digitalized, the cost estimates do not include the cost of printing the material. Assuming that school mental health providers have internet access and a computer, the cost of implementing Calm Cat does not fluctuate by district size or

67

by the number of counselors or mental health providers. An optional 1-hour question-and-answer session for providers following the training is offered for an additional $150.

Table 12

Ingredients for Tier 2 Mental Health Groups

Activities	Staffing	Administrative Time (Person Hours)	Hourly Rate	Cost
Calm Cat Training – 1 session	Not needed - virtual recording	-	-	$1,500
(Optional) Question & Answer Session post-training – 1 session	1 Devereux Specialist	1	$150	$150
Implementation manual, the student activity packets, and the daily progress monitoring tool	-	-	-	$0*
			Total cost of ingredients:	**$1,650**

Notes: Costs reflect average U.S. prices in constant 2022 dollars, rounded to the nearest ten. Only digital access to the material is included; the estimates do not include printing costs.

Universal Screening. The Social, Academic, and Emotional Behavior Risk Screener (SAEBRS) is a screening tool that supports a district's efforts to ensure all students have access to appropriate levels of school-based mental health services through screening and risk identification (Fastbridge, 2022). The plan presented in this section is offered by the School Mental Health Collaborative (SMHC) at the University of South Florida. SMHC provides comprehensive training, consultation, and technical assistance to districts to support evidence-based practices in mental health screening and service delivery. SMHC partners with schools for a minimum contract of 1.5 to 2 years. Implementing the SAEBRS tool involves five components – first-year consultation, teacher training, school Tier 2 team training, screening students, and second-year consultation. Table 13 provides a breakdown of these cost components.

The first step in installing the SAEBRS tool is to have consultations pre-installation,

which include SMHC consultants reviewing the screening readiness checklist (provided by SMHC), planning for local implementation (e.g., meetings with the partnering school district, understanding role of key stakeholders, adapting materials for local use), and creating pre-assessment development. Because districts can vary in the level of readiness prior to installing SAEBRS, the length and depth of consultation to district leaders can range from 5 to 10 hours per week with an hourly rate of $75-100 (for the first-year).

The second step is teacher training which can cost $5-10k for the entire district. This cost can vary by two factors: 1) the number of teachers needing to be trained; and 2) the need (if at all) to develop customized training for the district rather than using a pre-existing and more general training. A one-time Tier 2 training is then provided, typically lasting two hours and includes four, 30-minute follow up virtual consultations. This training includes data reviews in Fall, Winter, and Spring, as well as an end of year program evaluation and is charged at a flat rate of $2,000 for each school. Assuming there are 10 schools in a district, the estimated cost of Tier 2 training is $20,000.

The cost of screening students in a district is between $1-2 per student. SMHC offers bulk pricing. The cost of these packages varies depending on the different versions of SAEBRS that the district can choose from. The last cost component of installing the SAEBRS tool is consultation during the second year of the contract, which primarily focuses on maintenance and program improvement. This also varies by the depth of consultation needed and costs the same as consultation pre-installation. Year 2 consultation covers program evaluation including post assessment development, data management, data analysis, report creation, and presentation of results.

Table 13

Ingredients for Universal Screening

Activities	Staffing	Administrative and Executive Time (Person Hours)	Hourly Rate	Cost
Year 1 Activities				
Consultation (5-10 hours per week over 36 weeks) - readiness and installation phase	Various SAEBR expert consultants	180-360	$75-$100	$13,500-$36,000
Teacher training – 1 session (e.g., tailoring to fit district's needs, developing videos, content materials, etc.)	2 personnel – SAEBR expert consultants	-	-	$5,000-$10,000
Training schools' Tier 2 Teams – 1 session per school ($2,000 per school)	1 personnel – SAEBR expert consultant			$20,000*
Screening students 3 times a year at $1-2 per student	Not needed – teachers will administer	-	-	$10,000-$20,000*
Year 2 Activities				
Consultation (5-10 hours per week over 36 weeks) on maintenance and program improvement	Various SAEBR expert consultants	180-360	$75-$100	$13,500-$36,000
				Total cost of ingredients: $48,000 to $86,000

Notes: Costs reflect average U.S. prices in constant 2022 dollars, rounded to the nearest ten. The school Tier 2 training is an estimate for a district with 10 schools implementing PBIS

The cost estimates in Table 13 do not capture additional costs that may incur during the second year of the contract, such as material development. SMHC can provide (upon request) decision flow charts, trainings, and additional data reports.

Summary. As shown in Table 14, the total cost of the expanded, trauma-informed PBIS implementation model is between $175,390 - $213,390. This includes the optional expenses outlined in earlier tables. All components except for PSSP and universal screening are for a projected 1-year timeline.

Table 14

Ingredients for Expanded, Trauma-Informed PBIS Implementation Framework

Components	Cost
	$111,390
PSSP for school officers (2 years)	
Trauma training & implementation	$4,000
DCOT training & ongoing support	$10,350
Tier 2 training & implementation	$1,650
Universal screening (2 years)	$48,000-$86,000
Total cost of Expanded, Trauma-Informed PBIS Implementation Framework	**$175,390 - $213,390**

Note: Costs reflect average U.S. prices in constant 2022 dollars, rounded to the nearest ten.

There are several limitations to this cost analysis. First, the estimate does not include the cost of printing digitalized material and accompanying handouts. Second, the PSSP training for school officers, trauma training, and DCOT training do not account for the time required for school officers, school Tier 2 teams, and teachers. These estimates solely include the implementation of the component. Thirdly, the presented estimates may vary by district size and capacity. The cost estimates presented in this chapter are not approximated to a specific district; a district may incur slightly lower or higher costs, depending on its needs. When a particular cost varies by district size, the estimate assumes the cost for a district with 10 schools (or 5 PBIS coaches). The assumptions suggested are to represent realistic, average values in the absence of more precise estimates. The goal of this analysis is to provide guidance on future practice and decision-making. This chapter represents a starting point for discussion around the expanded, trauma-informed PBIS implementation model.

4.3. Limitations

The results described here should be interpreted in the context of several limitations. First, the COVID-19 pandemic substantially impacted the scope of intervention activities that

71

could be feasibly implemented. For instance, it was not possible for individual teacher coaching to occur during the time when schools were operating virtually. Moreover, the context of the pandemic impacted the availability and meaning of school-district administrative data. For example, office disciplinary referrals were not recorded during the 2020-2021 school year. Furthermore, schools may have varied in their pandemic restrictions (e.g., timeline on which they returned to a hybrid model during the 2020-2021 school year), and we were not able to account for this variation in analyses.

Second, analyses using school district administrative data were conducted at the classroom level, rather than individual student level, due to incompatibilities between school district data sharing permissions and NIJ archiving requirements. This limited analyses by reducing the statistical power to detect a difference between the Demonstration and Comparison groups. Further, the aggregate data could not be used to test subgroup analyses (e.g., test whether the Demonstration was more effective for students who were low achieving at baseline compared to those who were higher achieving at baseline). Finally, it was not possible to follow individual students over time to understand the longer-term impact of the implementation on students (Robin, 2016).

Additionally, the school staff members (e.g., teachers, safety offers, Tier 2 providers) who agreed to participate in surveys and interviews may not have been representative of the broader population of staff members in participating schools. The response rate was approximately 60% for teacher surveys and 54-66% for teacher interviews, which is comparable to response rates obtained on other voluntary surveys (Baruch & Holtom, 2008), but nonetheless may have biased results.

5. Artifacts

5.1. List of Products

Throughout the course of this project, project leadership, partners, and personnel from the collaborating school district developed and implemented a range of products to support the expanded, trauma-informed PBIS implementation model and the brief trauma trainings for school safety officers. Although the collaborating school district had preexisting efforts to implement PBIS in many of its schools, including the four Demonstration schools, and to support the use of trauma-informed practices, efforts through this project brought to bear new resources that enabled the development of a model encompassing both approaches. The project's partners as well as administrators and coaches from the collaborating school district played a leading role in the development of many of these resources. Many of the resources and products developed through this project are in continued use throughout the school district, in schools and with personnel beyond those that initially participated in the project.

Through work to support Research Question 2 (i.e., developing and implementing an expanded, trauma-informed PBIS implementation model), the following products were developed:

- A 6-part video module series about trauma-informed practices within PBIS. This video series was written, filmed, and edited during this project. These videos presented teachers with tangible information about the potential impacts of trauma on students and how a trauma-informed approach was in many ways aligned with PBIS principles.

- Implementation support resources that accompany the video module series providing links and additional tips to promote teachers' adoption of the strategies shared in the video modules.

73

- An approach for supporting small groups and individual teachers with implementing the skills shared in the videos. This approach relied on ongoing teacher observations and graphing of data over time to chart progress. PBIS coaches working on the project provided significant input on the development of this approach.

- Resources to strengthen Tier 2 implementation were also developed, including a structured plan for implementation of Tier 2 interventions and supports, as well as the development of online training modules for specific Tier 2 interventions used across the district.

- Finally, a manual to guide the training of school safety officers to adopt a trauma-informed approach in their interactions with students was developed during this project. This effort was led by a project partner and included significant feedback from PBIS coaches.

Through work to support Research Question 1 (i.e., brief school safety officer training), the following products were developed:

- A series of trainings for school safety officers were designed at the outset of this project to support officers in understanding trauma and the principles of PBIS. The development of these trainings was led by a project partner along with input from the collaborating school district. These included:

  - Two workshops on understanding and responding to students' trauma for school safety officers.

  - A workshop on PBIS for school safety officers.

5.2. Data Sets Generated

1) **School Safety Officer Surveys:** De-identified dataset at the school officer level. The dataset includes quantitative and qualitative responses from the school safety officers survey administered at the end of the 2018-19 and 2019-20 school years.

2) **Teacher Surveys:** De-identified dataset at the teacher level. The dataset includes quantitative responses from the teacher survey administered in Fall and Spring of the 2019-20 and 2020-21 school years.

3) **Administrative data:** De-identified dataset at the classroom level. The dataset includes administrative data provided by the school district for all k-5 students within the nine participating schools for the academic years 2016-17, 2017-18, 2018-19, 2019-20, and 2020-21. Dataset variables comprise demographic, attendance, climate, disciplinary, and academic outcomes.

4) **Qualitative interview data:** De-identified dataset at the individual interview participant level. The dataset includes transcripts of interviews with teachers, principals, coaches, and Tier II support personnel that were conducted in the Springs of 2020 and 2021.

5.3. Dissemination Activities

- The research team provided the Tier 2 leadership team at the partnering school district with preliminary implementation study findings regarding practitioners' experiences with the Tier 2 intervention components in the summer of 2020. This information was used to plan for Tier 2 supports during the current school year and is also providing a direction for the development of future Tier 2 supports across the district.

- Our video "Fostering Safe, Supportive, Consistent Relationships at School" was a finalist at the 2020 PBIS Film Festival at the Association for Positive Behavior Support conference.

- The project's implementation and evaluation approach was shared at NIJ conference in February 2021: *"Implementation of a Trauma-Informed Approach to PBIS."*

- Preliminary findings were shared during a symposium session at the Society for Research on Educational Effectiveness in September 2021: "*Evaluating and Implementing a Trauma-Informed PBIS Model in Philadelphia*". Three presentations were included in the symposium:

  o Teacher Outcomes from a Randomized Experiment of a Trauma-Informed PBIS Model

  o Implementation Outcomes from a Randomized Experiment of a Trauma-Informed PBIS Model

  o Evaluation of a Training on Trauma-Informed Practice for School Police Officers

- The PI presented, along with school district and other partners, at the National Association of School Psychologists conference in February 2022: "*Trauma-Informed PBIS in Philadelphia*".

- The PI shared project findings with all PBIS coaches in the collaborating school district in July 2022.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*. 50, 179-211.

Baruch, Y., & Holtom, B. C. (2008). Survey response rate levels and trends in organizational research. *Human Relations*, *61*(8), 1139–1160.

Beidas, R. S., & Kendall, P. C. (2010). Training therapists in evidence-based practice: A critical review of studies from a systems-contextual perspective. *Clinical Psychology: Science and Practice*, 17(1), 1–30. https://doi.org/10.1111/j.1468-2850.2009.01187.x

Bingham, A. J., & Witkowsky, P. (2021). Deductive and inductive approaches to qualitative data analysis. *Analyzing and interpreting qualitative data: After the interview*, 133-146.

Brown S. M., Baker C. N., & Wilcox P. (2012). Risking connection on trauma training: A pathway toward trauma-informed care in child congregate care settings. *Psychological Trauma: Theory, Research, Practice, and Policy*, 4(5), 507–515.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology*, *78*(1), 98.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334. https://doi.org/10.1007/BF02310555

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research* (3rd ed.) SAGE.

Fastbridge (2022). *SAEBRS (The Social, Academic, Emotional Behavior Risk Screener).* https://www.illuminateed.com/products/fastbridge/social-emotional-behavior-assessment/SAEBRS/

Frank, H. E., Becker-Haimes, E. M., & Kendall, P. C. (2020). Therapist training in evidence-

based interventions for mental health: A systematic review of training approaches and outcomes. *Clinical Psychology: Science and Practice*, 27(3),e12330. https://doi.org/10.1111/cpsp.12330

Goldstein, H. (2003). Multilevel Modelling of Educational Data. In: Courgeau D. (eds) Methodology and Epistemology of Multilevel Analysis. Methods Series, vol 2. Springer, Dordrecht.

Gray, A., Sirinides, P., Fink, R., Flack, A., DuBois, T., Morrison, K., & Hill, K. (2017). Discipline in context: Suspension, climate, and PBIS in the School District of Philadelphia (CPRE Research Reports).

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics*, 6(2), 107–128. https://doi.org/10.3102/10769986006002107

Hill, C. E., Knox, S., Thompson, B. J., Williams, E. N., Hess, S. A., & Ladany, N. (2005). Consensual qualitative research: An update. *Journal of Counseling Psychology*, *52*(2), 196–205. https://doi.org/10.1037/0022-0167.52.2.196

Kreft, I., & de Leeuw, J. (1998). Introducing statistical methods: Introducing multilevel modeling. London: SAGE Publications, Ltd. https://doi.org/10.4135/9781849209366

Levin, H. M., McEwan, P. J., Belfield, C. R., Bowden, A. B., & Shand, R. (2018). Economic evaluation in education: Cost-effectiveness and benefit-cost analysis (3rd ed.). SAGE.

McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: a review and illustration. Educational Psychology Review, *28*, 295–314.

Molloy, L.E., Moore, J.E., Trail, J., Van Epps, J.J., & Hopfer, S. (2103). Understanding real-world implementation quality and "active ingredients" of PBIS. *Prevention Science,* 14, 593–605. https://doi.org/10.1007/s11121-012-0343-9

Richardson, M. M., Coryn, C. L. S., Henry, J., Black-Pond, C., & Unrau, Y. (2012). Development and evaluation of the Trauma-Informed System Change Instrument: Factorial validity and implications for use. Child and Adolescent Social Work Journal, 29(3), 167–184

Robin, J. (2016). Using aggregate administrative data in social policy research, OPRE Report #2016-91, Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Saldaña, J. (2013). The Coding Manual for Qualitative Researchers. 2nd ed. SAGE.

Shand, R. & Bowden, A. B. (2022). Empirical Support for Establishing Common Assumptions in Cost Research in Education, Journal of Research on Educational Effectiveness, 15:1, 103-129.

Singer, J. D., & Willett, J. B. (2003). Applied longitudinal data analysis: Modeling change and event occurrence. New York: Oxford University Press.

Snijders, T. A. B., & Bosker, R. J. (2012). Multilevel analysis: An introduction to basic and advanced multilevel modeling. Thousand Oaks: Sage Publishing.

Tschannen-Moran, M., & Woolfolk-Hoy, A. (2001). Teacher-efficacy: Capturing an elusive construct. Teaching and Teacher Education, 17, 783–805.

What Works Clearinghouse (2022). What Works Clearinghouse Procedures and Standards Handbook, Version 5.0. Washington, DC: U.S. Department of Education, Institute of

Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). This report is available on the What Works Clearinghouse website at https://ies.ed.gov/ncee/wwc/Handbooks.

Wærsted, M., Børnick, T. S., Twisk, J., & Veiersted, K. B. (2018). Simple descriptive missing data indicators in longitudinal studies with attrition, intermittent missing data and a high number of follow-ups. *BMC research notes*, *11*(1), 123. https://doi.org/10.1186/s13104-018-3228-6

Weiner, B. J., Lewis, C. C., Stanick, C., Powell, B. J., Dorsey, C. N., Clary, A. S., Boynton, M. H., & Halko, H. (2017). Psychometric assessment of three newly developed implementation outcome measures. *Implementation Science*, 12(1). https://doi.org/10.1186/s13012-017-0635-3

Weltman, S. (2019). *How Much Does an Employee Cost you? U.S. Small Business Administration.* https://www.sba.gov/blog/how-much-does-employee-cost-you#:~:text=There's%20a%20rule%20of%20thumb,little%20harder%20to%20pin%20down.