Document Title:	**School Safety and School-Based Mental Health Services in a Large Metropolitan School District**
Author(s):	**Anna Yaros, James Trudeau, Jason Williams, Antonio Morgan-Lopez, Alex Buben, Stefany Ramos, Lissette Saavedra, Terri Dempsey, Alan Barnosky, Alex Cowell, Sherri Spinks**
Document Number:	**304408**
Date Received:	**March 2022**
Award Number:	**2015-CK-BX-0010**

November 2021

School Safety and School-Based Mental Health Services in a Large Metropolitan School District

Final Report

School Safety and School-Based Mental Health Services in a Large Metropolitan School District

Final Report

November 2021

Prepared by

Anna Yaros, James Trudeau, Jason Williams, Antonio Morgan-Lopez, Alex Buben, Stefany Ramos, Lissette Saavedra, Terri Dempsey, Alan Barnosky, Alex Cowell, Sherri Spinks
RTI International
3040 East Cornwallis Road
Research Triangle Park, NC 27709

_____

Contents

ii

Exhibits

Number		Page

Executive Summary

Evidence-based secondary and tertiary prevention programs for mental health in schools have the potential to impact an entire school population by reducing aggression and victimization and improving overall climate for students and staff (Ballard et al., 2014). RTI International partnered with Charlotte Mecklenburg Schools (CMS) to study school safety using a school-randomized controlled trial (RCT) of three types of school-based mental health (SBMH) services and a quasi-experimental study that compared each of the three SBMH arms to a set of propensity score-matched, nonrandomized, non-SBMH comparison schools (n = 34 schools).

ES-1. CENTRAL FEATURES OF SCHOOL-BASED MENTAL HEALTH

To help improve school safety, we examined the role of both perpetrators and victims of school violence as both are at risk for negative outcomes later in life. Victims of school violence often show depression (Greene, 2005), lower academic achievement (Nakamoto & Schwartz, 2010), and skipping school/truancy (Gastic, 2008). Perpetrators may go on to be at risk for dropout, alcohol or drug use, and other criminal behavior in adulthood (Hymel & Swearer, 2015; Osgood et al., 1996; Yeager et al., 2015). Violence perpetration and victimization in schools is preventable through school-based interventions.

School-based mental health (SBMH) refers to a range of services potentially provided by mental health professionals in the school building. We targeted SBMH because evidence suggests that students whose mental health needs are not met may be more likely to perpetrate school infractions or violence (Fabelo et al., 2011). The benefits of SBMH are sometime overlooked because implementation of evidence-based approaches of SBMH can be challenging within the school environment.

Our project was a partnership with Charlotte Mecklenburg Schools (CMS), which had existing SBMH services in place for their students through community partners who employed licensed therapists (1) to work in the schools, meeting with students weekly and as needed. Other providers in the schools who address student mental health needs included CMS-employed school counselors (2), school social workers (3), and school psychologists (4). Each of these 4 provider types played a role in our study.

Our study examined three different types of SBMH – Enhanced Treatment, Expanded Treatment, and Treatment as Usual. Funds from the grant were used in all three treatment conditions to support therapy provided in schools by community providers. These funds were used if students did not have Medicaid, private insurance, or the ability to pay for services out of pocket. Treatment As Usual (TAU) schools received standard SBMH with community therapists as well as services regularly provided by CMS-employed school counselors, social workers, and psychologists. Expanded (EX) Treatment schools received all of these services, plus support from Student Services Facilitators who tried to reduce administrative burden on school counselors by taking over Section 504 plan management and other administrative duties of school counselors. EX schools also received a changed allotment for psychologists' time such that psychologists only had 2 schools rather than 3 to serve. The Enhanced (EN) treatment schools received the services for TAU and EX schools, and added on training and implementation of evidence-based practice in dialectical behavior therapy (DBT) and in the Structured Psychotherapy for Adolescents Responding to Chronic Stress (SPARCS). Each of these interventions proved to have unique implementation findings and outcome effects.

ES-2. STUDY METHODOLOGY

The primary study goals were to

1) Implement and support staff in implementing 2 evidence-based programs in EN schools;

2) Conduct a process evaluation of this implementation and of school providers' experiences;

3) Use complementary experimental and quasi-experimental study arms to test school-level changes in student well-being and school safety;

4) Assess the cost and cost effectiveness of these SBMH models; and

5) Develop recommendations for the use of SBMH, DBT, and SPARCS in schools.

We used a design with both experimental and quasi-experimental arms to test our three types of SBMH and compare them with schools not receiving any SBMH. Thirty-four middle schools in CMS began participating in our project during the 2016-2017 school year and continuing through the 2018-2019 school year. Some schools in CMS already had SBMH services prior to the start of study (n = 25) and others did not (n = 9). The 25 schools in the SBMH arm of the study were randomly assigned to TAU, EX, and EH treatment conditions.
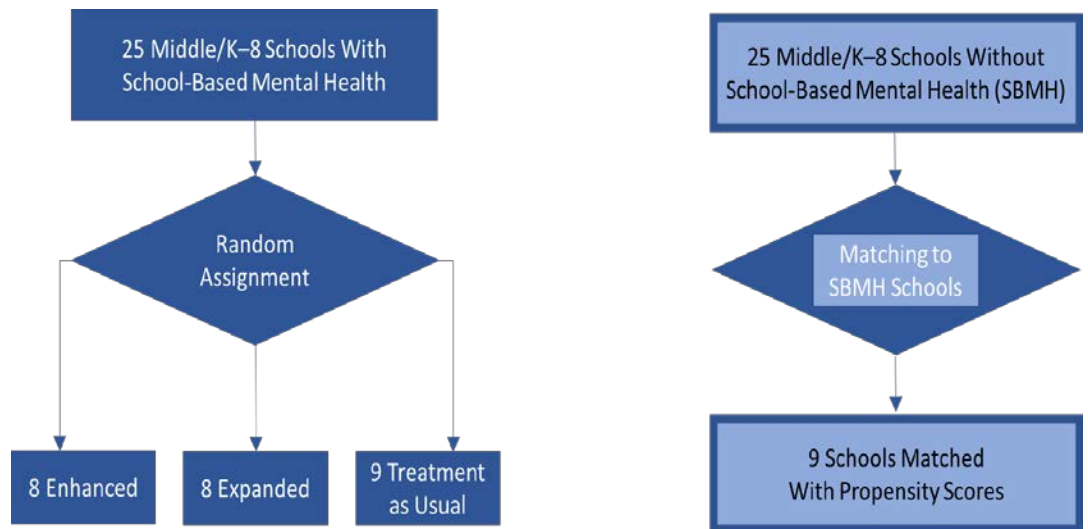
We met the study goals through carefully-collected implementation data, outcome data, and cost data. Implementation data consisted of provider surveys, provider logs, and provider interviews. Data collection proved somewhat challenging for each of these, but was very fruitful in content. Over the three years of data collection, providers completed a total of 181 surveys (some of which were completed at different time points by the same person), accounting for 91 School Counselor, 48 School Psychologist, 16 School Social Worker, and 26 Therapist surveys. Provider logs

recorded the number of students seen and the number of sessions with each student for therapists, psychologists and counselors. To better understand these quantitative data, qualitative data were also collected in the form of 22 interviews completed by providers who answered questions about barriers and supports for the implementation of these programs.

Outcome Data

Outcomes were measured using student and staff survey data as well as administrative data from schools. They were collected at the beginning of the study, which began in the 2016-2017 school year and each subsequent spring thereafter. Both staff and student surveys represented the whole staff and student bodies at their schools, not just those involved in SBMH. As a such, topics in the survey included school safety, violence, and victimization and reflected the experience of the people in the school as a whole.

Administrative data were analyzed by our research team after being exported by CMS to include grades, academic achievement tests, attendance, and discipline infractions. Given the focus of this study, we analyzed discipline infractions data primarily.

Cost evaluation tools were created for this study to measure the cost of resources including training, support, and staffing for each of the three SBMH groups.

Analytic Approach

To study implementation, we examined provider logs as a key measure of "service levels". We converted data in logs to create measures of SBMH by counting (1) the number of students that providers served, and (2) the amount of time providers spent with students. Because the current study focused on SBMH effects for the whole school, we converted these numbers into rates for each school and provider type based on enrollment of that school resulting in "minutes per student" and "percentage of students" served by therapists, psychologists, and counselors.

We analyzed outcome survey data using multilevel models with individual staff or students (depending on the survey) nested within schools, which were nested within treatment conditions. Administrative data were similarly used along with propensity

x

score weights to balance treatment groups at baseline. It is important to note, that students and staff were not tracked over time, but rather sampled at each timepoint independently from previous timepoints. This resulted in a serial cross-sectional design.

One group of multilevel models included treatment group and timepoint as categorical variables to estimate the effect of each treatment group. Because we found that service-levels as measured by implementation data from provider logs did not correlate as expected with treatment groups, we removed the treatment group distinction and ran other models using service levels from provider logs.

Cost data were used for three analyses – start-up costs, ongoing costs, and cost-effectiveness analyses. The cost-effectiveness analysis combined ongoing estimates for costs over three years and compared it with changes observed in student-rated victimization.

Lastly, qualitative data from provider interviews were coded and combined with service level data to understand barriers and supports for high, moderate, and low implementing providers. This mixed methods model allowed us to understand some of the variability seen between schools in service levels.

ES-3. STUDY FINDINGS

Service-level Findings

Examination of service levels, specifically "minutes per student" and "percentage of students" served by therapists, psychologists, and counselors, revealed that service levels did not differ by treatment group as we had predicted. For example, both EX and EN groups included more time for psychologists to spend at each school, but provider logs did not show a significant difference in either minutes per student or percentage of students between those conditions and the TAU and comparison condition. Instead, service levels varied among schools within the same treatment group.

Staff and Student Survey Results

Results from the staff survey showed mixed results when examining school safety and school climate outcomes between

treatment groups. Most staff members reported feeling safe before, during and after school, but most did report observing unsafe behaviors during the past 30 days including bullying, which are more concrete observations. Some trends toward showing statistically significant improvement over time in staff-reported bullying, fighting, and disruptive behavior were present through Time 3, but did not last into Time 4.

Student reported overall low levels of aggression and victimization, but as many as 50% of students reported feeling unsafe before or during school. Results for these and other indicators of school safety were mixed when comparing treatment groups with the SBMH groups showing improved effects over the non-SBMH group in some variables but not others.

Infraction Data

Infraction data included records of student discipline events that were coded as binary, indicating whether each student had an infraction in each of six categories, or count, measuring the number of infractions each student had in each of six categories. Discipline infraction data indicated that insubordination and disrespect were reduced in the Enhanced and Expanded groups relative to the TAU group. Findings in infractions of fighting, bullying, aggressive behavior and disruption were mixed depending upon whether the variable was binary or an overall count of the infractions.

When looking at all three outcome measures – staff survey, student survey, and infraction data, we conducted additional analyses using five measures of service levels rather than between treatment groups. Of the 9 staff survey outcomes, most significant findings were favorable, showing that therapist and counselor service levels had the largest number of significant favorable relations to staff survey outcomes, compared to unfavorable outcomes. For the 7 student survey outcomes, higher service levels were significantly related to favorable and unfavorable outcomes at about the same rate. Significant findings in the relations between service levels and favorable and unfavorable associations showed that higher service levels were more frequently related to more behavioral infractions than they were to fewer infractions.

Cost Evaluation

Cost analyses revealed that both start-up and ongoing costs of the Enhanced group were more than both the Expanded and TAU groups, as might be expected given the costs of evidence-based treatments. In pairwise comparisons of Enhanced, Expanded, and TAU treatment groups, the Expanded treatment showed the lowest Incremental cost-effectiveness ratio (ICER) in comparison to TAU, suggesting that it is the most cost-effective option for reducing student victimization in comparison to other types of SBMH.

Provider Survey and Interviews

The provider survey results showed that providers tended to be willing to practice evidence-based interventions. Results also showed that attitudes toward evidence-based interventions were significantly correlated with organizational readiness assessing their school's capacity to adopt evidence-based interventions.

Provider interviews supplemented these quantitative findings by illustrating the barriers and facilitators of implementation experienced by providers. Many providers reported significant and overwhelming mental health needs of students as a barrier to meeting their needs. They described believing that SBMH, DBT, and SPARCS were effective in addressing some students needs, but identified issues that make implementation challenging as lack of time and competing administrative responsibilities.

Mixed Methods Findings

The process evaluation used a sequential mixed methods design characterized by use of qualitative data to explain quantitative findings. Quantitative data from provider logs were combined to categorize schools into low, moderate, and high implementing schools. Providers at high implementing schools reported several more facilitators of SBMH than were reported by low implementing schools, though high and low implementing schools reported similar barriers.

ES-4 CONCLUSIONS AND RECOMMENDATIONS

Our study sought to implement and understand three levels of school-based mental health (SBMH) in a large metropolitan school district. Our measurement of SBMH intervention showed

that our interventions did not produce the systematic changes to the levels of service provided to students that we predicted. Schools did not systematically vary in their levels of service between treatment groups. Our qualitative data suggest that one reason for this variability may have been the intensive training and time commitment required to implement DBT and SPARCS treatments which may have taken away from the amount of time providers could spend with students.

Overall, there was partial support of our hypotheses that Enhanced SBMH would result in improved school safety and school climate above Expanded SBMH and TAU, but there were also results that did not support these hypotheses. Improvement in staff-rated safety and climate due to Enhanced and Expanded SBMH was found and was typically seen in the same school year as the intervention rather than in future years. In student-rated safety and climate, more unfavorable than favorable patterns were observed suggesting that, in concurrent measurement, higher levels of safety and climate problems as reported by students relate to simultaneously higher levels of services administered by providers in response to that need. We were not able to determine whether services were causing safety and climate issues, or, as is more likely, service levels reflected overall higher levels of student mental health needs which also related to safety and climate issue.

Analyses of student-rated victimization showed improvement across time in the Enhanced and Expanded interventions in comparison with the TAU group. Taking these effects into account, cost-effective analyses suggest that the costs associated with Expanded interventions provide the best cost-effectiveness to reduce student victimization, rather than Enhanced SBMH.

Challenges and Limitations

Our findings can be best understood by considering several limitations and challenges. Most of these relate to implementation of SPARCS and DBT which were found to be time-intensive in a school setting. Barriers that were also noted related to implementation were provider turnover and mixed attitudes toward engaging in such an intensive treatment.

In our design, we had some changes to our sample, which was mixed with K-8 and 6-8 schools. Some schools were dissolved due to redistricting and others moved conditions, though these were excluded from the analyses. We also experienced challenges in not be able to measure detailed implementation and fidelity data about SPARCS and DBT which might have explained variations in both implementation and in outcomes. The additional information we were able to get from provider interviews was still from a small number of providers who were likely the most engaged.

Our longitudinal design which sampled students and staff independently each time they were surveyed meant that students and staff were not tracked over time. Because of this we were not able to account for repeated measures variability, but instead treat it as unexplained variability leading to lower statistical power. We also were not able to track whether individual SBMH students improved given a lower number of consented students.

Recommendations for Future Research

This study provides evidence of efficacy and cost-efficacy of some elements of SBMH on some school safety and climate outcomes. This appears to be most endorsed by surveyed staff rather than surveyed students. Our study also illustrates challenges with introducing and implementing intensive evidence-based treatments in real-world schools. Meeting the needs of the number of students with mental health concerns and improving safety in their schools may require more providers with more time to devote to student service provision.

1 Introduction

Evidence-based secondary and tertiary mental health programs in schools have the potential to impact an entire school population by reducing aggression and victimization and improving overall climate for students and staff (Ballard et al., 2014). RTI International partnered with Charlotte Mecklenburg Schools (CMS) to study school safety using a school-randomized controlled trial (RCT) of three types of school-based mental health (SBMH) services and a quasi-experimental study that compared each of the three SBMH arms to a set of propensity score-matched, nonrandomized, non-SBMH comparison schools (n = 34 schools). Findings from staff surveys, student surveys, and administrative data did not show reliably improved school safety between treatment arms. Examination of implementation levels suggested that variability within treatment arm in levels of SBMH received by students predicted staff report, and to a lesser degree student report, of increased school safety. Specifically, the percentages of students seen for services by a school psychologist, school counselor, or SBMH therapist were related to increased feelings of safety and fewer unsafe incidents. A cost effectiveness analysis revealed that two levels of increased SBMH services were both more costly and more effective than SBMH treatment as usual.

1.1 IMPROVING SCHOOL SAFETY ON A SCHOOL LEVEL BY PREVENTING INDIVIDUAL MENTAL HEALTH PROBLEMS

School safety and violence prevention programs have targeted entire school populations, and evaluations of these universal programs typically evaluate the impact on the whole school. In contrast, evaluations of selective interventions have typically

measured outcomes that were specific to those students at risk for violence perpetration, victimization, or both. However, the current study assessed interventions that targeted and treated selected youth with specific mental health issues and evaluated the subsequent impact on the entire school population on reductions in aggression and victimizatio. This approach distinguishes the current study from prior studies. We partnered with Charlotte-Mecklenburg Schools (CMS) to create a robust study of several types of school-based mental health (SBMH) based on recommendations from the peer-reviewed literature.

1.1.1 School Violence Victimization and Perpetration

Exposure to violence in schools can have significant concurrent and long-lasting impacts on victims. Continuing victimization is associated with higher rates of internalizing and externalizing behaviors (Elliott et al., 1998; Hawker & Boulton, 2000; Sullivan et al., 2006). It has also been shown to lead to depression and anxiety disorders (Greene, 2005), relate to lower academic achievement (Nakamoto & Schwartz, 2010), and be associated with skipping school/truancy (Gastic, 2008). The more immediate academic consequences of violence in schools include disruption of the overall school environment— which, for victims, can include minor to serious physical injury from the violence itself, avoidance of the school environment, decreased classroom participation, and increased disconnections from academic pursuits (Buhs et al., 2006; Hymel & Swearer, 2015).

Perpetrators of school violence, particularly in middle school, are at risk for later increases in displays of aggressive behavior and weapon carrying (Hawker & Boulton, 2000). The longer-term risks for victims and perpetrators alike include school dropout, suicide, alcohol and other drug use disorders, and future involvement with delinquency and adult criminal behavior (Hymel & Swearer, 2015; Osgood et al., 1996; Yeager et al., 2015).

1.1.3 Addressing School Violence Victimization and Perpetration

In light of the concurrent and long-term risks that violence victimization and perpetration pose to youth, considerable research has been devoted to addressing and preventing school violence. These efforts initially started in the form of universal

prevention approaches that target peer violence perpetration during the middle school years, the critical point during the formative years at which peer-based aggression increases most steeply (Farrell et al., 2005). Advocates for universal prevention programming that is specific to peer victimization in middle school have suggested that several factors—including peer relationships, student-teacher relationships, and individual social-emotional skills—should be the primary targets of intervention (Portnow et al., 2018; Sullivan et al., 2012).

Reviews of the universal intervention literature across the spectrum of violence prevention suggest limited efficacy, as the impact is limited for youth who already display identifiable aggressive behavior (Bradshaw, 2015). These youth are more likely to benefit from selective intervention, an observation underscored by the facts that a large proportion of disciplinary infractions and school safety problems are perpetrated by a small minority of students within schools (Fabelo et al., 2011; Hoagwood et al., 2012) and that addressing their needs can improve the school climate for the entire student body (Ballard et al., 2014).

SBMH approaches have been a primary approach to serving the needs of the minority of students in school settings who have shown an early propensity for violence perpetration (Bruns et al., 2004; Cowan et al., 2013; U.S. Department of Education, 2014). This SBMH approach goes beyond traditional school mental health services provided by counselors and social workers to include community mental health providers within schools. Positive outcomes have been seen in academic performance and school attendance, although existing research has typically not employed experimental designs with random assignment. Despite evidence of efficacy, unaffordability of services creates a major barrier to implementation of services for many students with behavioral and emotional needs across the country.

An evidence-based approach to SBMH is consistently recommended (Ringeisen et al., 2003; Weist et al., 2014), including targeted or selective (Tier 2) prevention programs for youth at risk for aggression and violence and indicated (Tier 3) interventions for youth with behavior problems to improve school safety. Real-world implementation of evidence-based approaches in SBMH programs can prove difficult, with

limitations of time and school personnel, a lack of ongoing support for evidence-based implementation, and competition with established programs and priorities in the school system resulting in a gap between recommendations and actual implementation. Additional research is needed to test adoption and implementation of multitiered, evidence-based SBMH services, especially in middle schools. CMS, which has a large, existing SBMH program, provides an ideal setting for a research design that will test how improving implementation and increasing evidence-based therapies affects school safety and student well-being.

1.2 SCHOOL-BASED MENTAL HEALTH AND STUDENT SUPPORT SERVICES IN CHARLOTTE-MECKLENBURG SCHOOLS

CMS is the public educational system for students in Mecklenburg County, North Carolina. With more than 156,000 students and 178 school buildings, CMS ranks 17th in enrollment size among U.S. school districts.

School safety problems in the middle grades are prevalent around the country and in CMS. A 2017 survey of CMS students in grades 6 through 12 (Youth Risk Behavior Survey) indicated an urgent need for student support services:

- 42% of middle school students reported having been bullied at school

- 27% of district middle school students reported having carried a weapon at least once

- 50% of middle schoolers and 22% of high schoolers reported having been in a physical fight that year

Many CMS students experience maltreatment and negative occurrences not only in a school environment but also outside school, in the home, or in their neighborhood. As a result of not having the appropriate coping mechanisms, children may suffer difficulties in relationships or school performance. If left untreated, these difficulties can have a long-term impact on a child's educational attainment, employment, and relationships and can lead to incarceration. Although CMS has attempted to address these student needs in the past, individual results have been inconsistent.

Although CMS strives to address the social-emotional needs of all students, the size of its enrollment causes the district to struggle to meet the appropriate student-to-staff ratios for school counselors, psychologists, and social workers. Lack of availability of student services personnel greatly affects schools with high concentrations of students with low socioeconomic status. These schools are least likely to provide preventive behavioral measures and most likely to operate in a school culture of daily crisis. To address the mental health needs of students, the schools often need additional supports and services.

When the project began in 2015, CMS already had a robust system of student services providing mental health supports, including four primary types of providers: school counselors (two or three per middle school), school psychologists (one for each two or three middle schools), social workers (assigned to high-needs schools), and SBMH therapists (four agencies serving 25 middle schools, with one therapist per school). Therapist services were provided by community agencies who placed agency therapists at one or more schools to provide individual therapy and occasionally family or group therapy, if needed. Therapists were typically concentrated in Title I schools and primarily serve Medicaid-eligible students. Therapists work collaboratively with teachers and staff, attend individual student meetings, and provide consultation and education to school staff.

Exhibit 1.1   Types of Student Services Providers
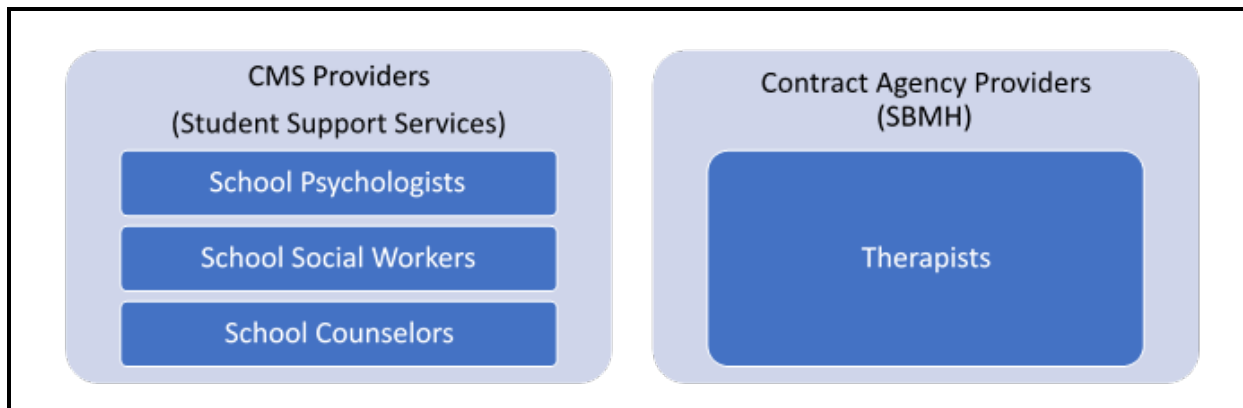
1.3 STUDY INTERVENTIONS

Three treatment conditions were used in our study to better understand the role of SBMH in school safety and student mental health in a real-world setting.

1.3.1 Treatment as Usual

Treatment as usual (TAU; described above) implemented the existing standard of practice in CMS using community therapists to provide **SBMH services** and school staff—counselors, social workers, and psychologists—to provide **student support services**. Two additional types of SBMH were tested: Enhanced and Expanded.

In schools randomized to receive TAU, Expanded, or Enhanced services, CMS used project funds to increase student access to SBMH for students whose families were unable to afford services. During the grant, the total number of students receiving SBMH per school was about 10 to 50 per school per year, depending on referral amounts.

1.3.2 Expanded Treatment for SBMH-Enhanced and SBMH-Expanded Schools

To allow school counselors and social workers to focus on their SBMH and student support roles, Enhanced and Expanded schools added a new itinerant position of student services facilitator. This person's role was to administratively support the tasks and duties typically managed by school counselors and social workers, such as intervention team facilitation and attendance monitoring. In addition, school psychologist staffing allotments were increased in Enhanced and Expanded schools so that each school would have at least a half-time school psychologist, rather than a one-third-time school psychologist, as in TAU schools.

1.3.3 Evidence-Based Interventions in SBMH-Enhanced Schools

Schools in the SBMH-Enhanced condition received all the interventions listed above along with Dialectical Behavior Therapy (DBT) and Structured Psychotherapy for Adolescents Responding to Chronic Stress (SPARCS). Our study used DBT and SPARCS services for two distinct groups of students on the basis of their demonstrated need and referral according to the CMS Response to Intervention Behavior Multi-Tiered Systems of

Support and the existing SBMH criteria described in
Exhibit 1.2.

Exhibit 1.2   Response to Intervention Behavior Multi-Tiered Systems of Support

Tier	Description of Referral Behavior	SBMH Services
Tier 1 problem student behavior	Sporadic occurrence of problem behavior with little disruption to the classroom learning environment. Only informal or brief consultation necessary from behavior support resources at the school site, such as classroom behavior support plans.	None
Tier 2 problem student behavior	Acute demonstration of the problem behavior. Tier 1 efforts have been exhausted. Problem behavior may consistently disrupt the learning environment or impede academic performance of the student. Potential for behavioral skill deficits.	SPARCS recommended if child has chronic stress or other needs for coping skills
Tier 3 problem student behavior	Tier 2 efforts have been exhausted or acute demonstration of problem behavior exists with severe spikes. Behavioral displays are internalized or externalized. Problem behavior may be the immediate and predominant concern over academic performance.	DBT recommended if child has emotion regulation or interpersonal problems

In eight schools randomly selected for training and implementation of Enhanced services, rigorous training in DBT and SPARCS was provided. The DBT treatment team, comprising SBMH private mental health providers, their supervisors, and the school psychologists, received Dialectical Behavior Therapy Intensive Training™ by Behavioral Tech, LLC, the leading DBT trainer in the United States. Training was conducted at CMS in two parts—a 5-day training (Part 1) and a 3-day training (Part 2), separated by 6–9 months for clinicians to apply what they had learned. DBT implementation in Enhanced schools consisted of three components: *individual therapy* conducted by the SBMH private mental health providers, *group skills training* conducted by the school psychologists, and weekly DBT *treatment team meetings*. After completion of the DBT training, members of the DBT treatment team could contact trainers at Behavioral Tech for ongoing consultation or technical assistance throughout the study.

In the same eight Enhanced schools, school counselors and school social workers, along with their supervisors from CMS, completed training in SPARCS from Duke University's National Center for Child Traumatic Stress (NCCTS). The NCCTS training consisted of three 2-day trainings held at CMS over the course of 10 months. Support provided during those 10 months by the

NCCTS also included twice-monthly consultation calls and monthly practice video conference calls, as well as intensive fidelity assessment consisting of ratings of session adherence and skills mastery, with feedback for trainees to facilitate improvements. After training, the NCCTS was available for continued consultation as needed. Student services staff in non-Enhanced schools had access to professional development available to all CMS or agency staff members. Students referred for SPARCS completed the 16-week program at the pace of one group session per week. For SBMH students, the typical length of individual sessions with a private mental health provider was 12.9 weeks with one session per week, but treatment plans were determined on the basis of clinical need.

Exhibit 1.3   Enhanced Treatment and Multi-Tiered Systems of Support

Dialectical Behavior Therapy (DBT; Linehan, 2014)

- Administered by therapists and school psychologists

Structured Psychotherapy for Adolescents Responding to Chronic Stress (SPARCS)

- Administered by school counselors and school social workers

Tier 3 Tertiary Prevention (Intensive)

Tier 2 Secondary Prevention (Targeted)

Tier 1 Primary Prevention (Universal)
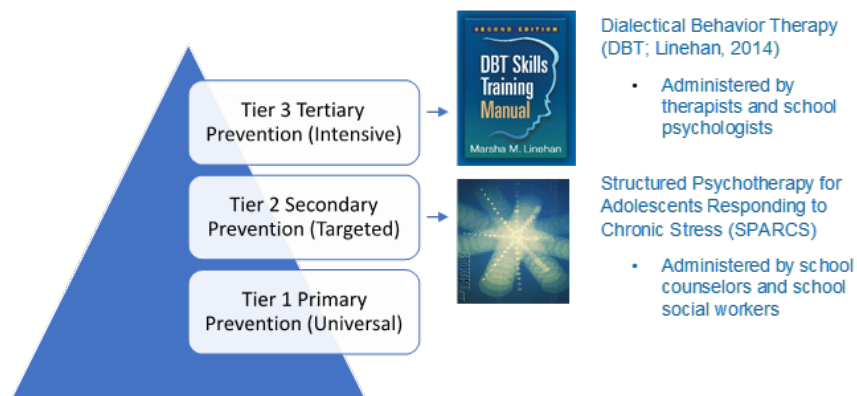
2 Study Goals, Design, and Methods

2.1 STUDY GOALS

The purpose of this study was to rigorously evaluate the implementation, outcomes, and cost of three levels of SBMH services as methods for improving school safety and student well-being. The specific goals of the study were as follows:

The purpose of this study was to rigorously evaluate the implementation, outcomes, and cost of SBMH services as a method for improving school safety and student well-being in middle and K–8 schools in a large, culturally diverse school system. The specific goals of the study were as follows:

- **Goal 1:** Implement rigorous training and support of evidence-based therapies for SBMH providers, while expanding student support services in a randomized sample of middle and K–8 schools.

- **Goal 2:** Conduct a thorough process evaluation of training and implementation of evidence-based interventions in a real-world SBMH setting.

- **Goal 3:** Use complementary experimental and quasi-experimental designs to evaluate the efficacy of SBMH enhancements and expansions in improving student well-being and school safety.

- **Goal 4:** Assess the cost and cost-effectiveness of SBMH enhancements and expansions using economic analysis.

- **Goal 5:** On the basis of findings, develop recommendations on the use of SBMH, DBT, and SPARCS in schools; disseminate recommendations through publications and presentations. Create a toolkit for schools to evaluate their own SBMH programs.

2.2 STUDY DESIGN

The study included two complementary components. One component was a three-arm, school-randomized controlled trial (RCT) comparing three treatment conditions: SBMH-TAU programming, Expanded programming, and Enhanced programming. The second component was a supplementary quasi-experimental study that compared each of the SBMH arms to a set of propensity score-matched, nonrandomized, non-SBMH comparison schools. **Exhibit 2.1** summarizes the randomized and nonrandomized components of the evaluation design.

Exhibit 2.1   Randomized and Nonrandomized Study Components
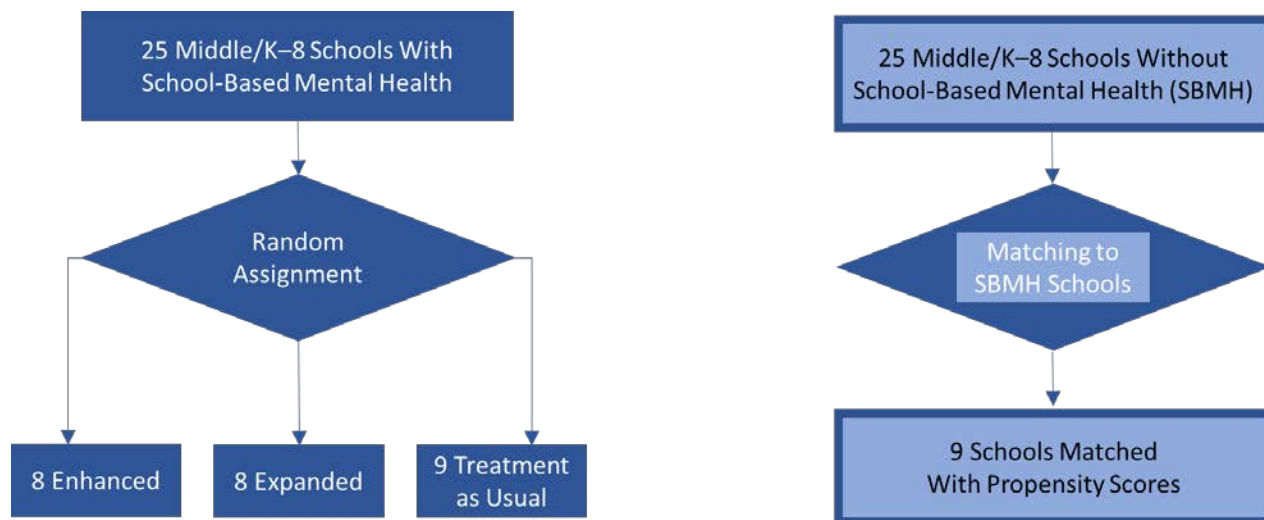
Exhibit 2.2 summarizes the mental health services that were available in schools in each study condition. To avoid denying SBMH services to students in need, CMS started SBMH in two non-SBMH schools, after we had randomized schools. Later, we describe how we accommodated this change analytically (**Section 2.6**) and its effect on results (**Section 3**).

Exhibit 2.2   Treatment Condition Student Support Services

Condition at Randomization	Non-SBMH (n = 7)	Formerly Non-SBMH Now TAU (n = 2)	TAU (n = 9)	Ex-panded (n = 8)	En-hanced (n = 8)
School counseling, school psychology, & social work	X	X	X	X	X
CMS TAU SBMH program therapists		X	X	X	X
Added *pro bono* time for SBMH therapists		**X**	**X**	**X**	**X**
Student services facilitator				**X**	**X**
Added school psychologists and increased coverage				**X**	**X**
Training in evidence-based treatments (SPARCS, DBT)					**X**

Note: CMS, Charlotte-Mecklenburg Schools; DBT, Dialectical Behavioral Therapy; EN, Enhanced; EX, Expanded; SBMH, school-based mental health; SPARCS, Structured Psychotherapy for Adolescents Responding to Chronic Stress; TAU, Treatment as usual.

2.3   LOGIC MODEL

As shown in ***Exhibit 2.3***, the project built on complementary *inputs*, including extant CMS services (SBMH, other student support services, and school safety efforts) in middle and K–8 schools and the well-established training programs for DBT and SPARCS available from Behavioral Tech, LLC, and the Duke University NCCTS, respectively. CMS leveraged these inputs to implement *activities*, including training of selected therapists and school staff and enhanced therapies provided to students. The evaluation measured *outputs* of these activities, such as the number of students who received enhanced therapies and the fidelity of the therapy, as well as *outcomes* at the school level (e.g., school climate) and the individual level (e.g., behavior).

Exhibit 2.3   Logic Model

INPUTS	ACTIVITIES	OUTPUTS	OUTCOMES
CMS Student Support Services • Short-term individual counseling • Group counseling • Parent training **CMS School-Based Mental Health (SBMH)** • 4 provider agencies serving 25 middle or K–8 schools **CMS School Safety Efforts** • CMS Police • School resource officers from local law enforcement jurisdictions **Clinical Training Expertise** • Behavioral Tech, LLC • Center for Child and Family Health (CCFH)	**Enhanced Training of Therapists** • Training in Structured Psychotherapy for Adolescents Responding to Chronic Stress (SPARCS) by Duke University's National Center for Child Traumatic Stress (NCCTS) • Training in Dialectical Behavior Therapy (DBT) provided by Behavioral Tech, LLC **Enhanced Therapies** • SPARCS for students who have experienced trauma or chronic stress • DBT for students with mental health disorders **Expanded Availability of Therapy** • Regardless of ability to pay • Include undocumented students	**Training Outputs** • # SBMH counselors and social workers trained in SPARCS • # SBMH therapists and school psychologists trained in DBT • Training satisfaction **Implementation Outputs** • DBT fidelity measures • SPARCS fidelity checklist • Client satisfaction questionnaire • Evidence-Based Practice Attitude Scale (EBPAS) **Service Outputs** • # of students who receive SBMH services • # of students who receive DBT • # of students who receive SPARCS • # of sessions—by treatment, by clinician, by school	**School Level** • Student survey—school safety, perpetration, victimization, school climate • School staff survey—school safety, school climate • Discipline infractions • Suspensions • Attendance • Academic performance **Individual Level** • Strengths and Difficulties Questionnaire (SDQ) • Youth Outcomes Questionnaire (YOQ) • Discipline infractions • Attendance • Academic performance

Process Evaluation
- Logs of implementation of enhanced therapies
- Provider survey: attitudes, organizational readiness, collaboration, and self-efficacy
- Qualitative interviews with high- and low-implementation clinicians

Outcome Evaluation
- School records
- Surveys of students and staff
- Individual measures

2.4 EVALUATION QUESTIONS

In this section, we present the evaluation questions (EQs) addressed in the study and briefly explain how we addressed each one.[1] We describe the data sources in **Section 2.5** and the analytic approaches in **Section 2.6.**

EQ-1. How did outcomes change for schools in the different study groups?

To answer this question, we used administrative data provided by CMS and staff and student survey data collected in the evaluation.

EQ-2. What levels of SBMH and other student support services were provided in each school each year?

---

[1] These evaluation questions have been streamlined from those in the proposal to speak more directly to what the evaluation was designed to study and provide information on.

To answer this question, we used data compiled from psychologist, therapist, and counselor service logs. Psychologists and school counselors were represented in all study schools; therapists provided services only in SBMH schools. For these types of services, we calculated for each school and each school year the number of students served and the hours of service provided. (Counselor logs indicated the number of students served but not the hours of service provided.) For all measures, we controlled for school size by dividing by the number of students enrolled.

EQ-3. What were the barriers and supports to implementation of different SBMH services?

The challenges and facilitators for SBMH service providers were measured using qualitative and quantitative tools. All service providers—therapists, psychologists, school counselors, and social workers—were asked to complete a survey in each year of the study to describe their beliefs and attitudes about providing mental health services to students. In addition, providers were asked to complete telephone interviews about their experiences and service provision.

EQ-4. What were the costs of providing Enhanced (per student and per school)?

To answer this question, we used cost data provided by CMS to estimate start-up costs and ongoing costs. Start-up costs were incurred at the beginning of the study before any students were provided services and consisted of training, hiring activities, meetings, materials, and some other small activities. Ongoing costs consisted of the labor, materials, and space costs associated with delivering the intervention.

EQ-5. Did outcomes change as a function of psychologist, therapist, and counselor service levels?

To answer this question, we analyzed the relationship between the outcomes (from EQ-1) and service levels (from EQ-2). We analyzed that relationship both for outcomes in the school year in which services were provided and for outcomes the following school year.

EQ-6. How did costs compare to improvements in outcomes?

We used a cost-effectiveness analysis (CEA) to answer this question. We used ongoing cost data that CMS provided to estimate the incremental costs between treatment groups, and outcomes data to estimate the incremental effectiveness between treatment groups. We then took the ratio of incremental costs to incremental effectiveness, indicating the cost of a one-unit improvement in the effectiveness outcome. Two CEAs were performed for this study, one for the aggression outcome and one for the victimization outcome.

2.5 DATA SOURCES

This section describes the data sources for the outcome evaluation, the implementation evaluation, and the CEA.

2.5.1 Implementation Data

Quantitative implementation data were collected through mental health provider surveys and logged records of service. Mental health providers included school counselors, psychologists, therapists, and social workers, except in non-SBMH schools, which did not have therapists or social workers. Survey and log data were aggregated at the school level to assess school implementation. A smaller group of providers was selected for in-depth interviews to provide qualitative data about implementation frequency and fidelity, as discussed further below.

2.5.1.1 *Provider Surveys*

Survey data were collected from all providers in fall 2016 and spring 2017, 2018, and 2019. Respondents completed the same survey items, regardless of their specific roles of counselor, social worker, psychologist, or therapist. ***Exhibit 2.4*** summarizes responses by provider type with response rates in parentheses.

Exhibit 2.4   Response Counts and Rates by Provider Type

Time Point	Counselor	Psych-ologist	Social Worker	Therapist
Baseline fall 2016	32 (40%)	16 (47%)	7 (44%)	10 (38%)
Spring 2017	28 (35%)	18 (53%)	1 (6%)	8 (31%)
Spring 2018	25 (60%)	8 (42%)	4 (31%)	8 (50%)
Spring 2019	6 (12%)	6 (25%)	4 (24%)	0 (0%)

The provider survey measured clinician attitudes toward and perceptions of enhanced therapies and evidence-based practices (EBPs; see **Exhibit 2.5**). Attitudes toward adopting EBPs were assessed along four dimensions using the Evidence-Based Practice Attitude Scale (EBPAS) (Aarons et al., 2010): (1) intuitive appeal of EBPs, (2) the likelihood of adopting EBPs given that adoption is a requirement, (3) openness to new practices, and (4) perceived divergence of EBPs from the usual practice. To measure appeal and requirement, respondents were asked how likely they would be to adopt a given therapy, using a 5-point scale ranging from *Definitely would not* to *Definitely would.* To measure openness and perceived divergence, respondents were asked to indicate the degree to which the statement described their practice, using a 5-point scale ranging from *Not at all* to *A very great extent.* One example was, "I am willing to try new types of therapy/interventions even if I have to follow a treatment manual."

To measure clinicians' self-efficacy in suicide assessment and intervention, the survey included three of the four original subscales of the Counselor Suicide Assessment Efficacy Survey (Douglas & Morris, 2017). Ability to assess student personal characteristics, assess suicide history, and carry out suicide intervention were assessed using a 5-point scale ranging from *Not confident* to *Highly confident.* Perceptions of organizational readiness for new or adapted therapies were assessed using a widely recognized tool (Austin & Claassen, 2008) that measures four components of readiness: (1) organizational capacity, (2) organizational culture/climate, (3) staff capacity, and (4) implementation plan. Each component includes five questions about readiness on a 4-point scale from *Not even close* to *We're there.*

Similar to the staff survey (described below), the provider survey also measured school safety and influences on collaboration. In addition to the subscales assessed in the staff survey, providers were asked about the frequency of their interactions with students, families, and school staff concerning consultations; completing assessments; plan development; interventions; and participation in teams.

Exhibit 2.5   Provider Survey Construct Summary

Construct	Source	Subscales
Attitudes	Evidence-Based Practice Attitudes Scale (Aarons et al., 2010)	Appeal Requirement Openness Divergence
Influences on collaboration	Expanded School Mental Health Collaboration Instrument— Community (Mellin et al., 2016)	Outreach and approach Interpersonal processes Outreach to communities and families Administrator support Interactions with students, families, and staff
Personal safety		Before school During school After school
Organizational readiness	Assessing Organizational Readiness for Implementation of Evidence-Based Practice (Austin & Claassen, 2008)	Organizational capacity Organizational culture Staff capacity Implementation plan

Demographic characteristics of the respondents are described in **Exhibit 2.6**. The majority of respondents were women at all four time points. School counselors accounted for the majority of respondents in the first three time points. In the last time point, counselors, and psychologists each accounted for one-third of the respondents, and there were no therapist respondents. The majority of respondents identified as white in all four time points, followed by Black or African American.

Exhibit 2.6   Provider Survey Sample Characteristics, by Time Point

	Survey Administration Period			
	Time 1 (n = 73)	Time 2 (n = 58)	Time 3 (n = 55)	Time 4 (n = 18)
Sex (% Male)	11.0	6.9	10.9	11.1
Counselor (%)	43.8	48.3	50.9	33.3
Psychologist (%)	21.9	31.0	14.5	33.3
Social Worker (%)	9.6	1.7	7.3	22.2
Therapist (%)	13.7	13.8	16.4	0.0
Other/Undisclosed (%)	11	5.2	10.9	11.1
Ethnicity (% Hispanic or Latino)	2.7	1.7	0.0	5.6
White (%)	61.6	70.7	54.5	55.6
Black or African American (%)	38.4	27.6	45.5	50.0
Asian (%)	2.7	3.4	1.8	0.0
American Indian or Alaska Native (%)	1.4	3.4	0.0	0.0
Native Hawaiian or Pacific Islander (%)	0.0	0.0	0.0	0.0

Notes: Provider type percentages may not sum to 100% due to rounding. Respondents were able to identify as more than one race; percentages therefore do not sum to 100%.

2.5.1.2   Provider Logs

Clinician logs were the primary metric of how often providers used evidence-based therapies and with whom. Clinicians entered data into the logs throughout the school year, and the study team collected the logs near the end of the school year, in spring 2017, 2018, and 2019. *Exhibit 2.7* details the number of logs collected for each provider type in each treatment arm.

Exhibit 2.7   Number of Schools Providing Service Logs

School Year	Treatment Group	Therapists	Social Workers	Psychologists	School Counselors
2016–2017	Non-SBMH	—	—	7	1
	TAU	8	3	9	7
	Expanded	8	6	8	6
	Enhanced	8	4	7	6
	Changed groups	1	—	2	—
2017–2018	Non-SBMH	—	—	7	3
	TAU	8	4	9	9
	Expanded	8	6	8	5
	Enhanced	8	4	8	7
	Changed groups	2	—	2	2
2018–2019	Non-SBMH	—	—	7	—
	TAU	9	6	9	—
	Expanded	6	4	6	—
	Enhanced	7	6	7	—
	Changed groups	1	—	1	—

Note: EN, Enhanced; EX, Expanded; SBMH, school-based mental health; TAU, Treatment as usual.

Therapists reported the number of sessions and the length of each session for every student they worked with during the year. Social workers submitted only the total number of students they saw during the year. Non-SBMH schools did not have therapists or social workers. Counselor data included the number of students seen and the number of sessions. Counselor data were not collected in the final year because the schools no longer required submission of detailed implementation data. Psychologists provided the most detailed level of intervention data with specific information about number of students seen and time spent with these students.

2.5.1.3   *Provider Interviews*

Beginning in late 2017 and continuing through early 2019, two sets of individual telephone interviews were completed with school counselors, psychologists, social workers, and therapists in CMS. The primary objective of this effort was to identify the respondents' perceptions of implementation and the impact of the implementation of SBMH practices in the schools where they worked. Included in this objective was the identification of

any differences in perception of the relative effectiveness of Expanded services and Enhanced services regarding student behavior and school safety.

A second objective was to identify the perceptions of student services staff regarding facilitators of and barriers to the implementation of SBMH in general and SPARCS and DBT in particular.

To do so, we examined information from all SBMH schools, with questions about DBT/SPARCS in Enhanced schools.

All counselors, psychologists, therapists, and social workers in SBMH-ET, Expanded, TAU, and non-SBMH schools were invited to participate in potentially two rounds of interviews. The interview questions for TAU and non-SBMH respondents did not include questions regarding the implementation of Enhanced or Expanded. In addition, because non-SBMH schools did not have therapists, interviews of these schools did not include therapists. Respondents were recruited via direct email to participate in confidential telephone interviews by BYC Consulting. Non-SBMH and TAU school student services staff participated in interviews that paralleled the interviews for the schools with SBMH programs without addressing experimental treatment conditions. Non-SBMH schools were included in the interviews to gain an understanding of how student mental health issues are typically addressed in settings where additional SBMH programming is not supported with additional funds. Respondents in all categories were identified by CMS student services administrators. Respondent contacts were made, and all interviews were done, by the project's local liaison. Approximately 150 individuals were available across the four categories of student service employment in each interview round.

The interviews varied in length from 20 to 45 minutes. These conversations were audio recorded and transcribed to facilitate analysis. A flexible interview protocol (***Appendix A***) was used to complete the conversations. In total, 22 respondents from 14 schools (12 TAU, Expanded, and SBMH-ET schools and two non-SBMH schools) contributed interviews.

2.5.2 Outcome Data

The three sources of quantitative outcome data were administrative data provided by CMS and surveys of students and staff (***Exhibit 2.8***).

Exhibit 2.8   Outcome Data Sources

Instrument	Respondents per School (34 Schools)	Mode	2016–17 School Year		2017–18 School Year		2018–19 School Year	
			Fall (T1)	Spring (T2)	Fall	Spring (T3)	Fall	Spring (T4)
Staff survey	40 randomly selected instructional staff and 20 noninstructional staff	Web-based survey lasting ~20 minutes	X	X		X		X[a]
Student survey	~120 students from randomly selected classes in 6th–8th grades	Paper-and-pencil survey; classroom setting; 1-hour session	X	X		X		X
Administrative data	All students	School records		X		X		X

[a]For Time 4, all staff at surveyed schools were recruited rather than a random selection of staff in order to conserve funds.

2.5.2.1 Survey Data

We surveyed students and staff at four points in time across 3 school years. ***Exhibit 2.9*** broadly summarizes the constructs measured in each survey and discusses the measures in more detail. We then describe sampling procedures, timeline, and administration for each survey.

The student and staff surveys combine subscales from various sources to measure school climate and school safety. ***Exhibits 2.10*** and ***2.11*** provide information about the measures used in each survey.

S*chool climate* was assessed using the Inventory of School Climate for staff (Brand et al., 2008) and students (Brand et al., 2003). The staff survey comprised six subscales measuring peer sensitivity (5 items), disruptiveness (5 items), teacher-pupil interactions (5 items), achievement orientation (5 items), support for cultural pluralism (5 items), and safety problems (4 items). The student survey comprised six subscales measuring teacher support (6 items), consistency and clarity of rules and

expectations (5 items), negative peer interactions (5 items), positive peer interactions (5 items), disciplinary harshness (5 items), and support for cultural pluralism (4 items). For a series of incidents, respondents were asked to indicate how often each incident occurred in their school using a 5-point scale ranging from *Never* to *Always*.

Exhibit 2.9   Constructs Measured in Student and Staff Surveys

Construct	Summary of Measures	Reporting Time Frame	Student Survey	Staff Survey
Personal safety	How safe does respondent feel before, during, and after school in each of 3 locations?	Past 30 days	Y	Y[a]
School Safety Problems Scale	How much of a problem in the school is each of 10 possible problems?	None specified	Y	-
	Have you observed the following behaviors?	Past 30 days	-	Y
Inventory of School Climate	Each survey had 6 subscales: 3 were common to both student and staff surveys and 3 were unique to the respective respondent groups	None specified	Y	Y
Aggression Scale	1 scale with 5 items	Past 7 days	Y	—
Victimization Scale	1 scale with 4 items	Past 7 days	Y	—
Psychological Sense of School Membership	1 scale with 18 items	None specified	Y	—

[a]Includes provider survey responses.

Exhibit 2.10   Student Survey Measures

Construct	Source	Subscales
School climate	Inventory of School Climate—Student (Brand et al., 2003)	- Teacher support - Consistency and clarity of rules and expectations - Negative peer interactions - Positive peer interactions - Disciplinary harshness - Support for cultural pluralism
Personal safety		- Before school - During school - After school
School safety	Schools and Staffing Survey (National Center for Education Statistics [NCES] & U.S. Department of Education, 2000)	Perceived threat to safety at school
Aggression and victimization	Victimization, Aggression, and Social Skills scales (Orpinas, 2009)	- Victimization scale - Aggression scale
School connectedness	Psychological Sense of School Membership (Goodenow, 1993)	Perceived sense of belonging in the school

Exhibit 2.11   Staff Survey Measures

Construct	Source	Subscales
School climate	Inventory of School Climate—Teacher (Brand et al., 2008)	- Peer sensitivity - Disruptiveness - Teacher-pupil interactions - Achievement orientation - Support for cultural pluralism - Safety problems scale
Personal safety	RTI International	- Before school - During school - After school
Influences on collaboration	Expanded School Mental Health Collaboration Instrument (Mellin et al., 2016)	- Outreach and approach - Interpersonal processes - Outreach to communities and families - Administrator support

Safety was assessed using the 10-item School Safety Problems Scale, which was developed for the Multisite Violence

Prevention Project (2004). In addition, to gauge personal safety, three sets of questions asked participants how often they felt unsafe before, during, or after school in various locations on school property. Respondents answered using a 3-point scale ranging from *Never, Occasionally,* or *Most or all of the time.* The student survey included an additional measure of school safety problems composed of 10 items that evaluated perceptions of threats to their safety at school. Items varied in the seriousness of the threat, from "Disrespect for teachers by students" to "Students carrying weapons." Three of the items were derived from the Department of Education School and Staffing Survey (now the National Teacher and Principal Survey). The remaining items were developed for the Multisite Violence Prevention Project. Responses were on a 4-point scale ranging from *Not a problem* to *Serious problem.*

The staff survey measured *influences on collaboration* using a subset of questions from the Expanded School Mental Health Collaboration Instrument (Mellin et al., 2014). Thirty-four items are grouped into four factors. The first factor, "Outreach and approach by mental health professionals from collaborating agencies," assessed how mental health professionals working together from multiple agencies approached work in schools. The second factor, "Interpersonal processes," assessed the relationship between school professionals and mental health professionals. The third factor, "School outreach to communities and families," assessed how the school supported community and family involvement. The final factor, "School administrator support," assessed the extent to which school administrators believed that expanded school mental health was important. Responses were measured on a 4-point scale ranging from *Strongly disagree* to *Strongly agree*, with an additional option for participants to indicate that an item did not apply at their school.

The student survey assessed bullying behavior using self-reported measures of *aggression* and *victimization*. Students were asked to report the number of times within the previous 7 days that they had experienced a specific event; the scale ranged from 0 to 6 or more times. Examples of events include "I teased students to make them angry" (aggression) and "A student teased me to make me angry" (victimization). The scales were developed at the University of Georgia and have been translated into many languages and used worldwide

(Orpinas, 2009) to measure aggression and victimization. The student survey also evaluated *school connectedness* using the Psychological Sense of School Membership scale, which measured students' perceived sense of belonging in their school (Goodenow, 1993). Students were asked to indicate how true each statement was using a 5-point scale ranging from *Not at all true* to *Completely true*. Examples include "I feel like a part of my school" and "People at my school notice when I am good at something."

Baseline data collection occurred in fall 2016. Subsequent follow-up data collection occurred in spring 2017, 2018, and 2019.

To obtain a representative sample of a school's total student population in grades 6, 7, and 8 at each time point, we randomly selected approximately six classrooms at each school to survey 120 students from grades 6, 7, and 8 (one to three classes from each grade). To ensure that each student had an equal chance to be selected, we drew student survey classes from academically diverse classes in subject areas that all students were required to take. Once classrooms were selected for the sample, teachers distributed "opt-out" forms to all students to ensure that parents were notified of the upcoming data collection. Parents had an opportunity to return a form to the school stating they did not want their child to participate, though few parents did this.

Student surveys were administered in individual classrooms or other appropriate locations at the school during a 1-hour administration period. The 30-minute student survey was administered using paper-and-pencil bubble sheets. Trained RTI staff and field staff members administered student surveys during a 1- to 2-day student survey administration at each school, accommodating individual school schedules.

The 20-minute web-based staff survey was conducted over the course of 4 weeks at each survey administration. All staff employed at the school, both instructional and noninstructional, were eligible to participate. For the first three data collection time points, 40 instructional staff and 20 noninstructional staff in each school were randomly selected for the survey. For the last staff survey, all staff were asked to participate for practical reasons (e.g., budgetary concerns, feasibility).

Student and staff survey data were collected and managed using REDCap electronic data capture tools hosted at RTI International (Harris et al., 2019; Harris et al., 2009). REDCap (Research Electronic Data Capture) is a secure, web-based software platform designed to support data capture for research studies, providing 1) an intuitive interface for validated data capture; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for data integration and interoperability with external sources.

Both the staff survey and student surveys were voluntary and anonymous. The number of students responding at each time point was 4,025 at Time 1 (baseline), 3,588 at Time 2 (Spring 2017), 2,600 at Time 3 (Spring 2018), and 2,621 at Time 4 (Spring 2019). During the course of the study, we were able to reduce the number of students surveyed per school by examining intraclass correlations and correcting our power analysis. High response rates were obtained for all four staff survey time points: 57.8% (1,066 of 1,845) at Time 1, 62.3% (1,160 of 1,863) at Time 2, 57.2% (1,081 of 1,889) at Time 3, and 41.8% (1,117 of 2,670) at Time 4.

The student and staff samples comprised serial cross-sectional samples each year. A cross-sectional sample collects data from a new random sample of students and staff at each time point. This approach is appropriate and commonly used when the focus is on whether the program effects change in a population of youth in intervention schools relative to a population of youth in control schools.

2.5.2.2   *Administrative Data*

CMS provided the evaluation team with a wealth of administrative data that included information on a comprehensive list of infractions, tardiness and attendance, and in-school and out-of-school suspensions. For the purposes of this project, we focused on infractions as a measure of school safety.

CMS provided a data set for each study school year that listed all reported infractions for any student charged with one or more infractions. There are 81 potential infractions, ranging from cell phone use to assault involving the use of a weapon.

2-17

Exhibit 2.12 presents the total number of infractions reported for each school year in each of the study schools, the number of students with one or more infractions, and the range of the number of reported infractions per student.

Exhibit 2.12   Summary Information on Reported Infractions in Study Schools

School Year	Total Number of Infractions Reported	Students With One or More Infractions	Range of Infractions per Student
2015–16	17,291	5,728	1–50
2016–17	15,543	5,479	1–40
2017–18	16,548	5,464	1–49
2018–19	17,079	5,889	1–47

Note: Information in this table is based on the raw data on infractions, not on data from our manipulations described below.

There was substantial variation between schools and across years in the total number of infractions each school reported, ranging from 13 to 1,675, but the average number of infractions was most commonly between 750 and 850 per school per school year. This variation is not fully explained by differences in enrollment: the number of infractions *per student enrolled* also varied substantially, ranging from 0.02 to 2.75. In addition, schools varied in the number of reported individual infractions—for example, aggressive behavior. Such variation between schools in reported infractions (total or individual infractions) could be due to differences in *student behavior* (i.e., the actual frequency of behaviors meant to be identified as infractions), differences in *reporting practices* (e.g., the threshold for reporting a given behavior as an infraction), or *both*.

Not all the infractions were pertinent to the study outcomes of interest. Therefore, for our analytic data set, we filtered out 45 infraction types such as cell phone use, excessive displays of affection, and falsification of information.

Using the remaining 36 infraction types, we worked with CMS to create six categories such that infractions were most similar to other infractions within a category and distinct from infractions in the other categories. The six categories are (1) harassment, threats, and bullying; (2) insubordination and disrespect; (3) disruption and misbehavior; (4) aggressive behavior; (5) fighting and assault; and (6) weapons (including

threats of weapons use) and gangs. ***Exhibit 2.13*** shows the infractions in each category and the total counts of each infraction in the study schools over the 4 study years.

Exhibit 2.13   Infraction Categories Created for the Evaluation

Category & Infractions	4-Year Count
1. Harassment, Threats, & Bullying	
032 Inappropriate language/disrespect	**5,678**
052 Bullying	787
019 Communicating threats	761
038 Harassment—sexual	685
025 Harassment—verbal	610
094 Cyber-bullying	141
107 Threat of physical attack without a weapon	141
101 Harassment—racial	26
109 Harassment—sexual orientation	15
102 Harassment—disability	5
110 Harassment—religious affiliation	5
023 Extortion	4
Category Total	8,858
2. Insubordination & Disrespect	
033 Insubordination	**9,044**
061 Disrespect of faculty/staff	2,899
Category Total	11,943
3. Disruption & Misbehavior	
042 Disruptive behavior	**13,615**
037 Bus misbehavior	2,961
114 Inappropriate behavior	1,214
022 Disorderly conduct (G.S. 14-288.4(a)(6))	408
Category Total	17,790
4. Aggressive Behavior	
027 Aggressive behavior	**10,891**
5. Fighting & Assault	
024 Fighting	**9,053**
044 Assault on student	865
003 Assault on school personnel, not serious	219
072 Assault on student w/o weapon	85
045 Assault—other	47
014 Sexual assault, not rape or sexual offense	17
093 Robbery without a weapon	12
090 Violent assault not resulting in serious injury	8
021 Affray	8
071 Assault on nonstudent w/o weapon	2
001 Assault resulting in a serious injury	1
Category Total	10,317

(continued)

Exhibit 2.13 Infraction Categories Created for the Evaluation (continued)

Category & Infractions	4-Year Count
6. Weapons (Including Threats) & Gangs	
009 Possession firearm or powerful explosive	376
105 Threat of physical attack with a firearm	42
079 Gang activity	30
106 Threat of physical attack with a weapon	28
008 Possession of a weapon (excluding firearm, explosives)	12
002 Assault involving the use of a weapon	6
043 Bomb threat	1
Category Total	495

Note: Boldface indicates the most frequent infraction in each category.

CMS linked all infractions to disciplinary *incidents*, specific events for which students received a formal reprimand. A student could receive multiple infractions within an incident, and each incident could involve multiple students receiving infractions. To avoid overcounting similar infractions (e.g., classroom disruption and insubordination) in our six infraction categories, we created binary variables to indicate whether an incident included one or more infractions in a category. For example, an incident including infraction 033 Insubordination and 042 Disruptive behavior would be coded as "1" in infraction categories 2 and 3, respectively. Summing the count of these binary variables within a school year for each student yielded the total number of incidents per school year in which a student received at least one infraction in a given category, which were the data we used in our analysis.

Analyzing infractions in this way provided two benefits. First, it reduced the number of variables from 36 infractions to six categories, which made statistical analyses much more manageable; this was especially important given our multicomponent study design that included both school randomization for SBMH schools and a quasi-experiment comparing infractions in SBMH intervention schools with those in matched non-SBMH schools. Second, it is likely that using the categories helped to reduce variation due to differences among reporters in applying the infraction types. For example, a given behavior may be seen as "fighting" by one reporter and as "assault on student" by another. Including both infraction types in a single category avoids that potential problem. Moreover, we think that most reporters would be likely to

report any given behavior using an infraction type that we would place into the same category rather than infraction types that we would place into different categories. For example, it would be relatively less likely that one reporter would report an infraction in the Harassment, Threats, and Bullying category and another reporter would report the same behavior in the Fighting and Assault category. In short, examining categories rather than individual infractions likely helps to smooth out some artifactual differences in reporting.

In addition to the category-indicator counts, we also analyzed the counts of the most frequent infraction in each category (shown in bold in *Exhibit 2.13*). Compared to the category-based data, the data on these most frequent infractions were less complete because they excluded the less-frequent infractions. However, focusing on the most frequent infractions offered the advantage of knowing which specific infraction was under consideration; in contrast, the category totals included multiple infractions. (Aggressive behavior was a category unto itself, and therefore we are not including it in the analysis of infraction models.)

2.5.3 Cost Data

The cost data structure comes from the principles of activity-based costing. We first categorized costs by a limited set of comprehensive and exclusive activities (see *Exhibit 2.14*). These are the basic throughputs by which the interventions can improve mental health and reduce violence. We estimated the cost of each activity by assessing the quantity and price of each type of resource used to produce the activity (e.g., the FTE and salary labor).

Estimating economic effects entails four sequential steps:

(1) Identify the activities needed to implement the program.

(2) Identify the resources used to execute each activity.

(3) Determine the quantity of each resource used.

(4) Assess the unit cost for each resource.

Exhibit 2.14 Activity Categories and Resource Types to Estimate Start-Up and Ongoing Costs
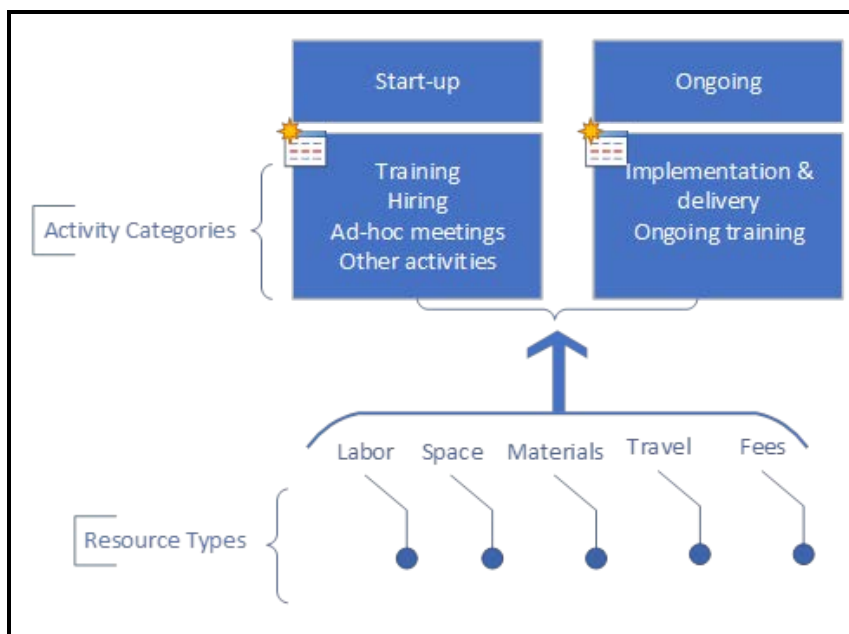
Exhibit 2.15 describes the general categories for each of the activities and resources and the quantity measure and data source for each. **Exhibit 2.16** similarly lists the activities and resources, along with their unit cost measures and data sources. Each exhibit provides the quantity and unit cost data sources for both start-up and ongoing costs. To best measure real-world implementation costs, we did not include costs that are research-only elements of the study, such as staff time for completing grant paperwork.

We identified the activities and the resources through semistructured telephone interviews with two CMS staff (the SBMH specialist and the grant coordinator). We conducted a first interview to identify the activities and resources for starting up SBMH at the study schools and then a series of ad hoc phone calls to resolve questions about the data being reported. On the basis of these discussions, we assembled a start-up cost questionnaire to collect the quantity and unit costs of the resources and asked the CMS staff to complete the guide.

Exhibit 2.15   Measures and Data Sources for Resource Quantities

Activity	Resource	Quantity Measure
Start-Up Costs (via Start-Up Cost Questionnaire)		
Trainings	Trainer fee	One per training
	CMS staff time	Hours
	Provider time	Hours
	Materials	Count of materials
	Space	Square footage
	Travel	Travel cost
Hiring activities	CMS staff time	Hours
Ad hoc meetings	CMS staff time	Hours
Other activities[a]	CMS staff time	Hours
	Materials	Count of materials
Ongoing Costs (via Ongoing Cost Questionnaire)		
Implementation	Counselors	FTE[a]
	Social workers	FTE
	Psychologists	FTE
	Providers	Billed amount
	SBMH specialist	FTE
	Grant coordinator	FTE
	Space	Square footage
	Interns	FTE
	Materials	Count of materials
Ongoing trainings	See above list	Billed amount

[a]Coordination of training, legal, and IT staff.

Exhibit 2.16   Measures and Data Sources for Unit Prices

Activity	Resource	Unit Cost Measure	Data Source
Start-Up Costs			
Trainings	Trainer fee	Cost per training	SUQ
	CMS staff time	Wage rates	SUQ, expense reports, BLS
	Provider time	Billed rate	BLS
	Materials	Cost per material	SUQ
	Space	Cost per sq ft	Loopnet.com
	Travel	Travel cost	SUQ
Hiring activities	CMS staff time	Wage rates	SUQ, expense reports, BLS
Ad hoc meetings	CMS staff time	Wage rates	SUQ, expense reports, BLS
Other materials	Materials	Cost per item	SUQ
Other activities[a]	CMS staff time	Wage rates	SUQ, expense reports, BLS
	Materials	Cost per item	SUQ
Ongoing Costs			
Implementation	Counselor	Salary	OGQ
	Social worker	Salary	OGQ
	Psychologist	Salary	OGQ
	Provider	Billed amount	OGQ
	SBMH specialist	Salary	Expense report
	Grant coordinator	Salary	Expense report
	Space	Cost per sq ft	Loopnet.com
	Interns	Salary	OGQ
	Materials	Cost per item	OGQ
	Ongoing trainings	Billed amount	OGQ

Note: BLS, Bureau of Labor Statistics, U.S. Department of Labor; OGQ, Ongoing Cost Questionnaire; SUQ, Start-Up Questionnaire.

[a]Coordination of training, legal, and IT staff.

We followed a similar process to collect ongoing costs in each of the 3 implementation years. In the first year of implementation, we interviewed CMS staff to identify all ongoing activities and resources at the study schools and then followed up with an Ongoing Cost Questionnaire to collect the quantity and unit cost data. We followed the same approach in the next 2 years of implementation. We used CMS administrative data on annual school enrollment to apportion district-wide costs to schools and to generate cost-per-student estimates. The district-level resource types were time for district staff, any trainings they attended, and any materials purchased at the district level. Because the ongoing cost guide asked CMS for the average annual salary district-wide for school counselors, social

workers, and psychologists, any within-district wage difference was not accounted for.

When needed, we supplemented unit cost data from other sources, including project expense reports and Occupational Employment Statistics from the Bureau of Labor Statistics in the U.S. Department of Labor. Additionally, CMS was not able to provide us with a cost per square foot of school space; we instead used the average cost of office space in Mecklenburg County from Loopnet.com.

2.5.4 Data Collection Timeline

The calendar of data collection and sources (described in preceding sections) is summarized in *Exhibit 2.17,* Data Collection Calendar. The grant-funded school mental health services supported by the study were delivered in the 2016–17, 2017–18, and 2018–19 school years. Throughout each year of delivery, providers kept service logs, which the schools provided to RTI at the close of the school year. In addition, CMS provided RTI with administrative data on student infractions for each of those school years, in addition to data for the 2015–16 school year (prior to delivery of grant-funded supplemental mental health services). For each school year, RTI tallied services and infractions for analysis. Finally, staff, student, and provider surveys were administered in fall and spring of the 2016–17 school year (Time 1 and Time 2, respectively), and in the spring of the 2017–18 (Time 3) and the spring of the 2018–19 school year (Time 4). The administrative data from the 2015–16 school year and the survey Time 1 are considered baseline, capturing outcomes prior to service delivery (to a large extent). The administrative data for the following 3 school years and survey time points (Time 2, Time 3, and Time 4) measure outcomes that follow service delivery during each of those years. The provider service logs and administrative infractions data capture events throughout the school year, whereas each survey time point captured events and perceptions at the time of the survey or within the preceding 30 days. Details on how we used these data sources in analyses are provided in later discussions of analysis plans and results.

Exhibit 2.17   Data Collection Calendar

	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
2015–16 School Year											
Administrative Data						2015–16					
2016–17 School Year											
Provider Service Logs						2016–17					
Administrative Data						2016–17					
Survey Reporting Period		Time 1							Time 2		
2017–18 School Year											
Provider Service Logs						2017–18					
Administrative Data						2017–18					
Survey Reporting Period									Time 3		
2018–19 School Year											
Provider Service Logs						2018–19					
Administrative Data						2018–19					
Survey Reporting Period									Time 4		

2.6   ANALYTIC APPROACH

Both the RCT component and the quasi-experimental component were analyzed under an intention-to-treat framework to assess intervention effects under real-world effectiveness conditions for implementation.

Power analyses were conducted for the cluster-randomized design of the study using Optimal Design version 3.0 to assess the minimum detectable effect size at which .80 power would be achieved across multiple parameters: (1) differences in school-level intraclass correlations (.05, .10), (2) differences in the number of schools in the non-SBMH (16 for pairwise intervention condition comparisons, 32 for comparison between all SBMH schools and control schools), and (3) average within-school sample sizes (30 for students in SBMH, 100 for the student survey). Under all scenarios examined, the range of detectable effect sizes for .80 power is .24–.55, which is considered between small and medium.

To ensure that an equal proportion of the three randomized conditions were represented in each of the four mental health

2-26

provider agencies, randomization was conducted (using SAS Proc Plan) with the provider agencies as a stratification factor. Before propensity score weighting, comparisons between the combined SBMH schools and the non-SBMH schools showed baseline covariate effect sizes ranging from |.14| to |1.49|, with SBMH schools (not unexpectedly) showing lower school enrollment, a higher percentage of students receiving free or reduced-price lunch, higher suspension rates, and higher student-reported aggression and victimization. After propensity weighting, differences between SBMH and non-SBMH schools were reduced to statistically nonsignificant differences, effect sizes that were smaller than |.10| or both; importantly, the effect sizes for the baseline levels of aggression and victimization were each below |.10| after weighting. Of the other key covariates, short-term suspension rates were nonsignificant (p = .09) but still had an effect size difference at baseline of |.61|. As a result, suspensions were included as a covariate in outcomes analysis (along with the propensity weights) to assess whether there were any differences in outcomes results (e.g., statistical significance, effect sizes). As there were no differences when suspensions were included as a covariate along with propensity weights, suspensions were removed from outcome models.

2.6.1 Implementation Analysis

We estimated the provision of SBMH and other student support services by examining (1) the number of students that providers served, and (2) the amount of time providers spent with students. The number of students seen and minutes spent with students were collected from the provider logs and used to create five implementation measures (***Exhibit 2.18***). Because the current study is testing the impacts of SBMH on the whole school, we converted each measure into a rate per student by dividing by the total number of students enrolled in the school, resulting in "minutes per student" and "percentage of students" served by each provider type – therapist, psychologist, and counselor. This allows us to interpret the effects of implementation across the entire student population accounting for the size of the student body and to compare results between schools with different enrollments. For the 11 K–8 schools, the providers reported the number of students served and minutes of service for the entire school, not specifically for students in grades 6–8. Most of the schools serve only grades

6–8 and the provider log data are specific to students in those grades. During the study, two schools changed treatment arms and were dropped from the final analyses for all time points.

Exhibit 2.18   Provider Log Data Sources

	Raw Measures	Rates
Psychologists	# of students per year	# of students seen per year, divided by school enrollment
	# of minutes per year	# of minutes per year, divided by school enrollment
Therapists	# of students per year	# of students seen per year, divided by school enrollment
	# of minutes per year (based on 50 min session)	# of minutes per year, divided by school enrollment
Counselors	# of students per month	# of students seen per month, divided by school enrollment

We analyzed implementation rates between treatment groups within each of the 3 school years, 2016–17, 2017–18, and 2018–19. We conducted pairwise comparisons of each service-level measure between each pair of treatment groups (e.g. Enhanced compared with TAU) in each school year.

2.6.2   Outcome Analysis

The staff and student survey outcomes were analyzed using multilevel models in which respondents were nested within schools. Neither staff nor students were tracked longitudinally and consequently there are no repeated measures across the four assessments nested within respondents. Although individual respondents were not tracked across assessments, it is possible that individual staff members or students responded at more than one data collection wave.

The administrative data on disciplinary infractions were longitudinal at the student level, i.e. data for a given student could be linked over the years. However, students changed schools, including moving between schools in different treatment groups – which made longitudinal analysis problematic. We explored several approaches to

accommodating that movement between schools: a) weighting each student's by the number of days he or she was in each school each year; b) assigning each student to the school that they attended the most days each school year; or c) dropping students who changed schools from the analysis. For technical reasons, the first two solutions encountered analytic challenges that made them not viable or appropriate. The third solution, dropping from analyses students who changed schools, was not desirable because those students were more likely (than students who did not change schools) to receive one or more disciplinary infraction reports and to account for a disproportionate number of infractions. For those reasons, it seemed important to include them in the analyses. Our ultimate solution was to not make use of the longitudinal nature of the infractions data and instead conduct serial cross-sectional analysis by assigning each student to the school he or she attended the most days each school year, but not tracking individual students across time, which avoided many of the problems associated with students changing schools across school years in the longitudinal analyses. Under the serial cross-sectional framework, the analysis that we applied to the infractions data was very similar to the analysis of the staff and student survey data, which helped to standardize interpretation of results across the different data sources.

All models used the propensity score weights described in the beginning of Section 2.6 and included short-term suspension rate as a school-level covariate.

Treatment group models. The treatment group multilevel models included group and time as categorical variables. Preliminary models that treated time as a measure of linear change across the 3 follow-up years (i.e., a longitudinal growth model approach) suggested that schools did not follow a well-defined linear functional form. An alternative modeling approach was used that estimated differences by treatment condition in changes from baseline (Time 1) to each follow-up in turn. Estimates from this model were then used to calculate and evaluate pairwise comparisons of each group within each follow-up assessment. School was included as a random intercept in all models. Models were estimated in SAS PROC GLIMMIX using a binary or continuous response distribution as appropriate for the outcome being modeled.

2-29

Service-level models. Five school-level implementation measures were used as predictors in models that ignored treatment group and focused on the amount of services rendered by mental health professionals in the schools. Service levels were measured throughout the school years, with the student and staff surveys administered late in those school years. Change over time in implementation was not examined due to the low school-level sample size and variability in implementation across the time points. Instead, the multilevel model used each implementation measure as a time-varying predictor of each staff or student outcome. Separate models were used for each implementation item and all models were weighted and used short-term suspensions and baseline outcome values as covariates. Two variants of the models were used. The first used concurrent measures of implementation and outcomes (i.e., service levels and survey outcomes from the same school year) and the second estimated lagged effects by predicting the outcome with the previous year's service levels. Models were again estimated in SAS PROC GLIMMIX using a binary or continuous response distribution as appropriate for the outcome being modeled.

2.6.3   Cost-Effectiveness Analysis

The analytic perspective for the economic analysis was that of CMS, meaning that the analysis included costs accruing to the district: school staff time, trainer fees, materials and supplies paid for by the district, and school space. Costs accruing outside the district, such as the time costs for students and their families, were not included in this perspective. We chose the district perspective because it seems most likely that future SBMH programs will most likely be considered and initiated at the school district level.

Students at the study schools received mental health services from several staff types. Most schools had a mix of counselors, psychologists, social workers, and community-based therapists, and schools in the Enhanced and Expanded conditions had student services facilitators. To adequately compare the costs of SBMH across schools and its relationship to school safety outcomes, this economic evaluation considered SBMH to be any mental health services provided at school.

We collected cost data at the school level. SBMH is not a specific intervention but rather a variety of services employed

by multiple specialists within a school. Although an economic analysis at the student level would have allowed for student-level variation in the cost data, collecting these data would have been cost prohibitive for the study and highly burdensome for school staff.

2.6.6.1   Start-Up Costs

To estimate start-up costs, we first estimated the cost of each activity in Exhibit E by combining the unit cost (or price) and quantity of each resource (Exhibits E and F); we then summed over all activities. The mathematical expression for start-up costs SC_{ar} is

$$SC_{ar} = \sum_{a=1}^{l} \sum_{r=1}^{k} P_{ar} Q_{ar} \ ,$$

where P_{ar} and Q_{ar} are price and quantity for activity a_{ar} and resource r_{ar}.

2.6.6.2   Ongoing Costs

To estimate the average ongoing cost per student in each study year, we used an approach similar to that for start-up costs. The annual school-level cost per student $\overline{OC}_{sy}$ is the sum of the product of the unit costs (i.e., prices) for each of the resources listed in Exhibits E and F, divided by the school enrollment:

$$\overline{OC}_{sy} = \sum_{a=1}^{l} \sum_{r=1}^{k} \frac{P_{syar} Q_{syar}}{E_{sy}} \ .$$

In addition to the terms defined above for start-up costs, costs are at the level of the school, s, and study year, y; E_{sy} is school enrollment.

Because two resources of the SBMH intervention—community provider time and trainings— affected only students in grades 6–8, we adjusted the denominator E_{sy} appropriately. For schools that served only grades 6–8, the denominator E_{sy} was the same across all resources. For schools that served grades K–8, we used K–8 enrollment as the denominator for community provider time and training resources and the full K–8 enrollment for the remainder of resources.

2.6.6.3   Cost-Effectiveness Analysis

We computed the relative cost-effectiveness for each treatment arm by combining cost estimates with predicted values of scores on the student-reported outcomes of victimization at the school level. We used the random effects model to predict at

2-31

the level of the treatment arm and derived from that school-level marginal effects for each year. We did not compute confidence intervals because of the limited number of schools per treatment arm.

2.6.4 Qualitative Analysis

Qualitative analysis of provider interviews was carried out through an inductive review across response categories from each participant. This "grounded theory" (Glaser & Strauss, 1967) technique provides for a three-step approach in analyzing the interview data:

- Interviews are transcribed and read multiple times.

- "Responses" within and across respondents are open coded (Corbin & Strauss, 1990). This was carried out by examining and coding common themes in each conversation and then refining the results using a constant comparison (Padgett, 2008) approach.

- Respondent "constructs" and subsequent themes are constructed to group together response types that reflect related beliefs and experiences.

The constructs and related responses were also entered in a database along with demographic (student services employment category, school) and research variables (treatment category, interview round) to facilitate analysis and create descriptive statistics. The responses to the interviews were analyzed in two different ways. First, we describe the 27 response categories and eight constructs we identified. Second, we describe the overall nature of the responses to the questions contained in the interview protocol and identify differences between or similarities in responses between treatment conditions. To achieve validity of the constructs used in the analysis, we familiarized three independent individuals not associated with the project but familiar with social science research with the construct descriptions and had them assign a random selection of 20 participant responses to the appropriate construct. Interrater agreement in this exercise was calculated at 86.4%.

2.6.5 Mixed Methods

Our process evaluation used a sequential explanatory mixed-methods design, in which quantitative data inform qualitative data collection and analysis. We used quantitative data to identify high and low implementers (dosage) of SBMH and

student support providers. We used qualitative interviews to explain those quantitative data by exploring barriers to and supports for implementation. The qualitative interviews provide context and clarification of how differences in levels of dosage and implementation may have occurred. The mixed-methods model is shown in ***Exhibit 2.19***.

Exhibit 2.19   Mixed-Methods Process Evaluation Model
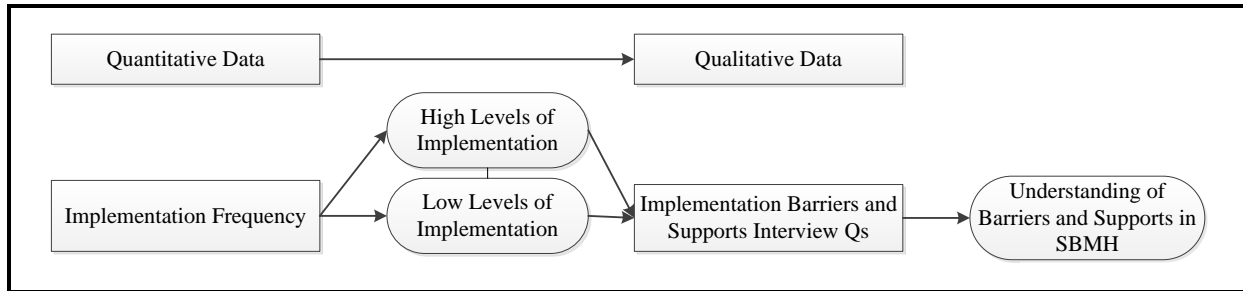
3 **Results**

In this chapter we present the study findings. First, we describe the level of mental health services provided in the study schools. Next, we present school safety and student behavior outcomes for the treatment groups, as reflected first in the staff survey, then in the student survey, and finally in administrative data on infractions. We also examine the associations between the outcomes and the levels of mental health services provided in the schools. We then present findings presenting results from the economic analysis.

To shed light on the quantitative results, we offer themes identified in qualitative analysis of interviews with school counselors, psychologists, social workers, and therapists. Finally, in a mixed-methods analysis, we explore how themes identified in the qualitative analysis cast light on the levels of services provided.

3.1 SERVICE-LEVEL RESULTS

The project provided funds for additional mental health support services, with increased services from therapists in TAU, Expanded, and Enhanced groups, increased psychologist services in Expanded and Enhanced groups, and increased time for counselors to provide direct services, rather than do administrative work, in Expanded and Enhanced groups. Accordingly, we expected higher rates of implementation by psychologists in Expanded and Enhanced groups compared to TAU and non-SBMH schools, and higher rates of implementation by counselors in SBMH schools (TAU, Expanded and Enhanced) compared to non-SBMH schools.

We analyzed implementation rates between treatment groups within each of the 3 school years, 2016–17, 2017–18, and 2018–19. We conducted pairwise comparisons of each service-level measure between treatment groups in each school year. This analysis yielded 66 separate comparisons using T-tests (3 years * 6 pairwise comparisons for 2 psychologist measures = 36, and 3 years * 3 pairwise comparisons for 2 therapist measures = 18, and 2 years * 6 pairwise comparisons for 1 counselor measure = 12). None of these comparisons were statistically significant at the 0.05 level, and only two approached significance. During the 2018–19 school year, the percentage of students seen by a psychologist was marginally higher in TAU schools compared to Expanded (p = .07) and in Enhanced compared to Expanded schools (p = .09).

However, because the service-level data were aggregated to the school, the unit of analysis was the school and each pairwise comparison was based on a relatively small number of schools in each treatment group. The comparisons had limited statistical power and the lack of significant differences should be considered in that light. To further explore service levels within schools in the treatment groups, we graphed each school's service level during each implementation year; see **Exhibits 3.1** through **3.5.** Schools varied widely for each service-level measure, with substantial variation *within* treatment groups. Therefore, in our outcome analysis results discussed in **Section 3.2,** in addition to examining differences in outcomes between treatment groups we also examine differences in outcomes as a function of service level, irrespective of treatment group.

Exhibit 3.1   Percentage of Students Served by Psychologist, by Treatment Group
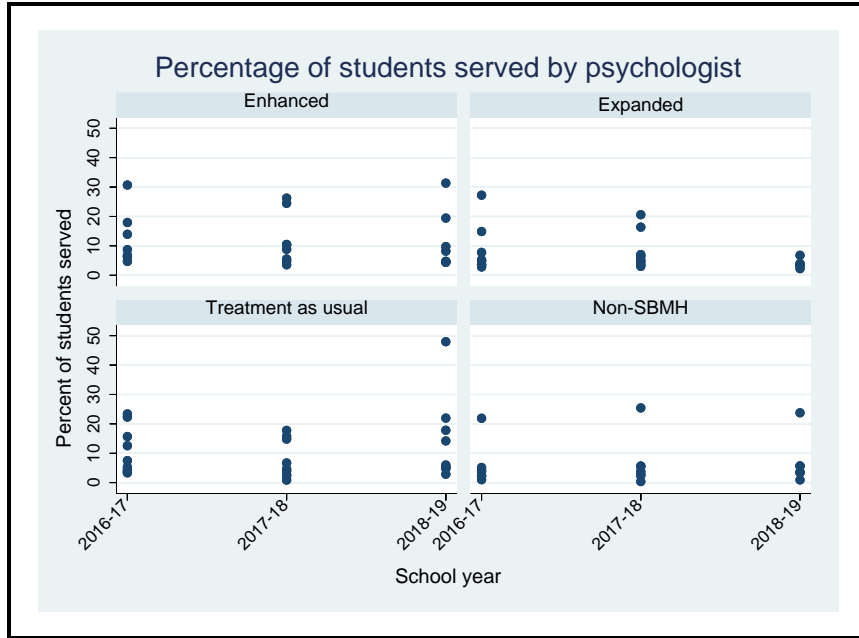
Exhibit 3.2   Minutes of Psychologist Services per Student, by Treatment Group
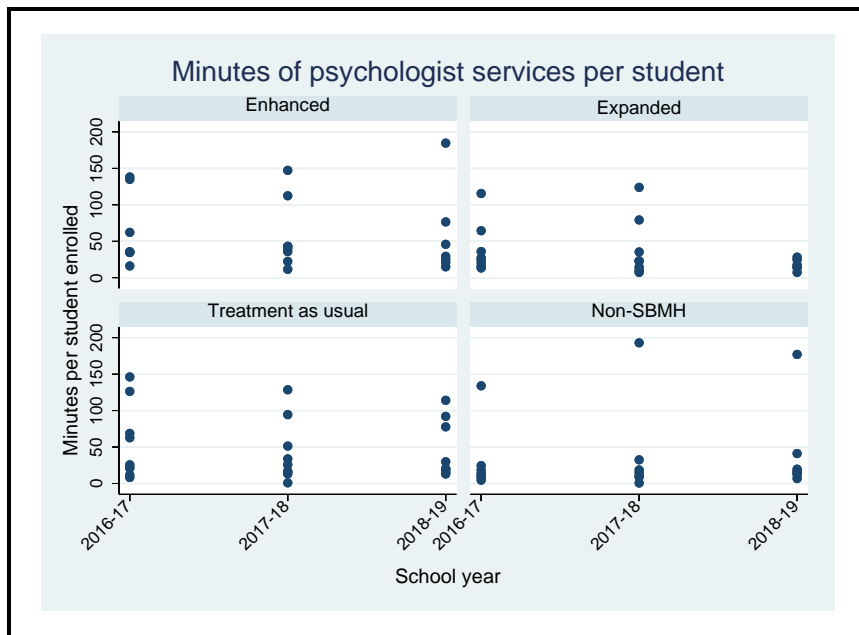
Exhibit 3.3 Percentage of Students Served by Therapist, by Treatment Group
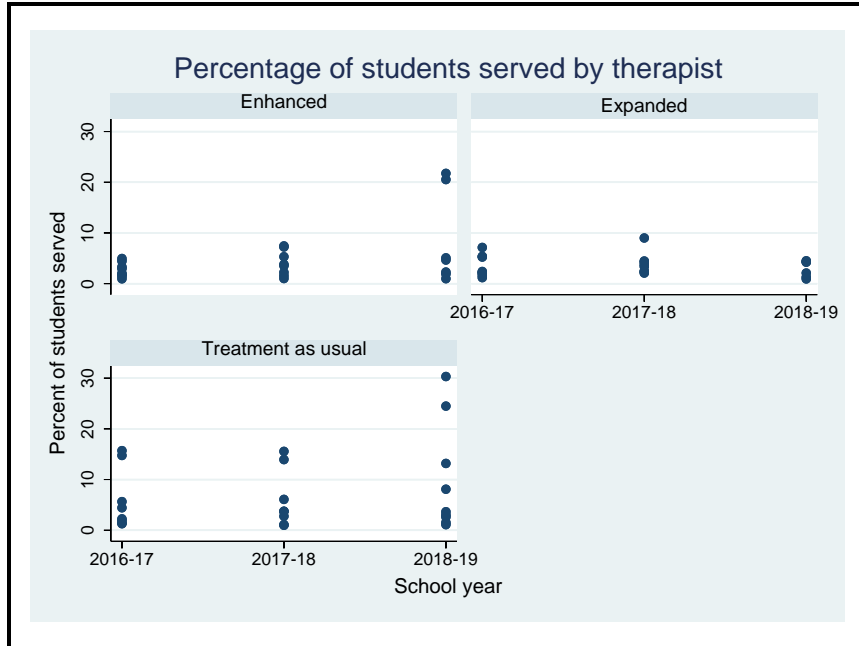
Exhibit 3.4 Minutes of Therapist Services per Student, by Treatment Group
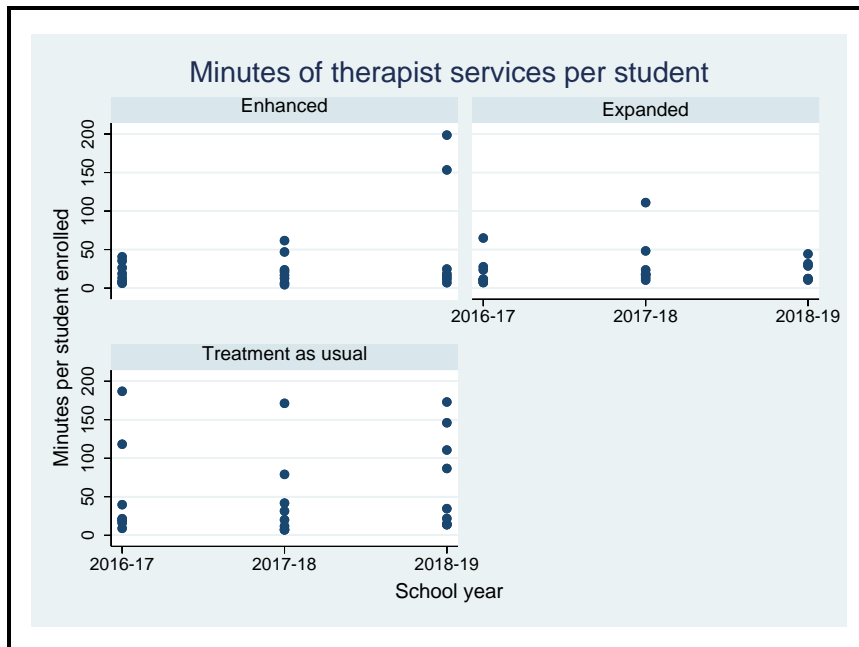
Exhibit 3.5   Percentage of Students Served by Counselor, by Treatment Group

Percentage of students served by counselor
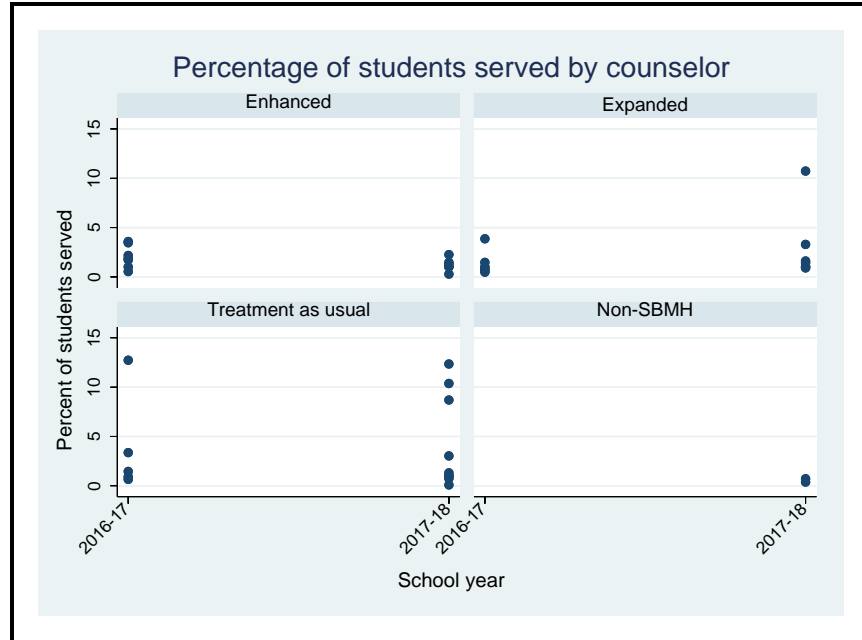
3.2   STAFF SURVEY RESULTS

Staff survey respondents were asked about various aspects of school safety and student behavior, including their own feelings of safety and personal experiences at school and their observations and perceptions of student behavior. They were also asked about their school's and principal's support of community mental health professionals in the school, topics pertinent to SBMH.

3.2.1   Effect of Treatment Group on Staff Survey Outcomes

We first examined differences in outcomes by school treatment group. Using the multilevel model with group by time interactions described in Section 2.6.2X, we did pairwise group comparisons (e.g., Enhanced group vs. Expanded group) of the outcomes reported by staff in each school treatment group.

Staff Feeling Unsafe. Staff were asked how often during the past 30 days they felt unsafe before, during, and after school hours in each of three locations: "in their classroom or office"; "in other areas of the school building (i.e., hallways, stairwells, restrooms, and cafeteria)"; and "in the parking lot, sports fields, or other areas of the school grounds". Responses of *Most or all of the time* were rare and therefore were combined with responses of *Occasionally* to simplify analyses. We assigned

3-5

staff who gave either response a single indicator, *Ever*. For all nine combinations of time frame and location across four survey time points and four treatment groups (144 combinations total), most staff felt safe. The percentage of staff who felt unsafe ranged from 5% (in the classroom or office after school, measured at Time 3 for Expanded group staff) to 29% (in the classroom or office during school, measured at Time 4 for Enhanced group staff). For 49 of the 144 combinations (34%), less than 10% of staff reported feeling unsafe; 10% to 20% of staff reported feeling unsafe in 86 of the combinations (60%).

Responses did not differ significantly by location, so we assigned a single binary indicator for each time frame if staff said they felt unsafe in any location. **Exhibit 3.6** shows the model-estimated percentage of staff who reported feeling unsafe before, during, or after school. (Model-estimated results are the results — here, percentages — after adjusting the raw survey findings by applying the propensity weights and suspensions covariate described in Chapter 2).

We found few statistically significant differences in the percentage of staff in each group who reported feeling unsafe at each survey time point for each time frame. For the *before-school* time frame, at Time 4, staff in Enhanced schools reported feeling less safe than those in Expanded schools ($p = .040$). For the *during-school* time frame, at Time 3, staff in Enhanced and non-SBMH schools reported feeling less safe than staff in Expanded schools ($p = .019$ and $p = .041$, respectively); this difference dissipated at Time 4 because of an uptick in the percentage of Expanded school staff who felt unsafe. For the *after-school* time frame, there were no statistically significant differences between groups at any survey time point; at Time 4, the difference between staff in Enhanced and Expanded schools approached significance ($p = .085$).

Exhibit 3.6   Percentage of Staff Who Felt Unsafe Before, During, and After School Hours
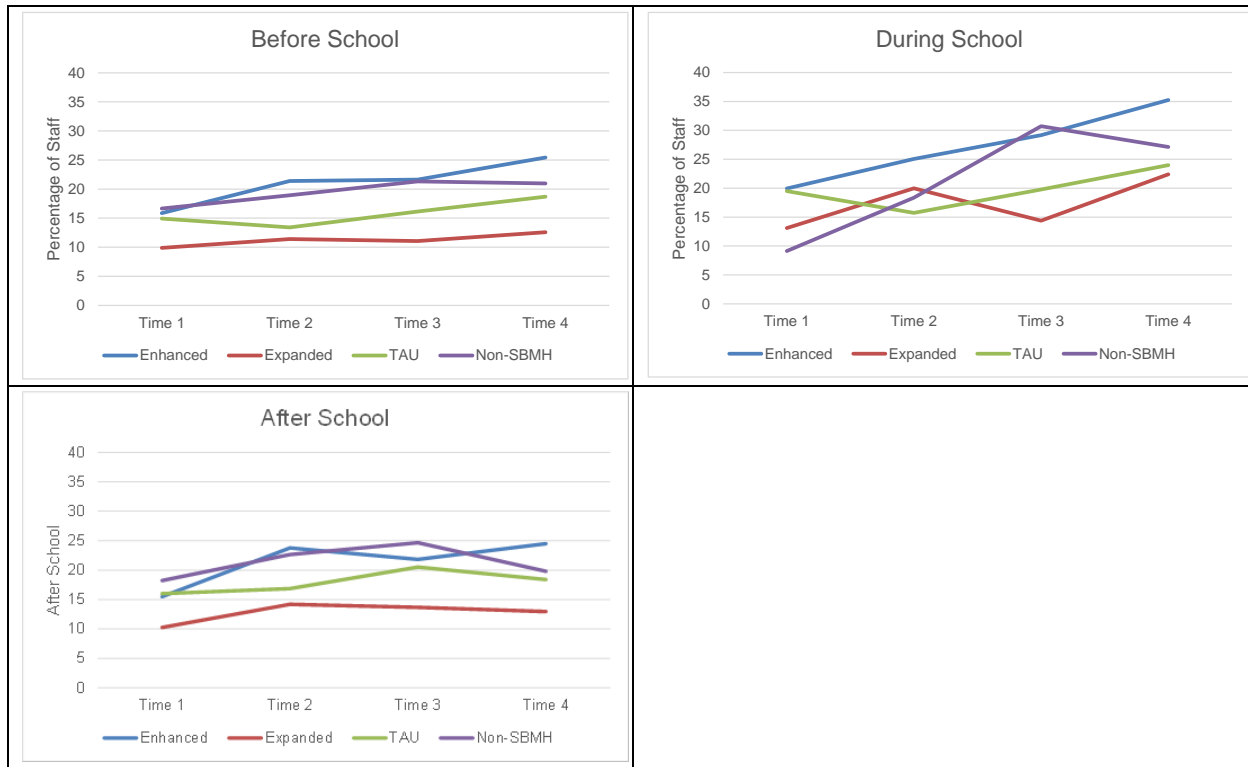
Staff Personal Experience. The Safety Problems Scale of the Inventory of School Climate includes items asking staff how often each of the following happened:

- Student attempted to or actually hit or assaulted you when you were at school.

- You have been afraid that a student will hurt you at school.

- Student at school threatened to hurt you if you didn't give them your money or something else that belonged to you.

- You bring something to school to protect yourself.

Likert scale response options ranged from *Never* (scored 1) to *Always* (scored 5). The most frequent response was *Never,* and responses at the other end of the range were infrequent (e.g., at Time 1, 89% of staff reported never having been threatened if they didn't give students money or belongings, and less than 2% reported this happened most of the time or always); thus, we again created a binary indicator of *Never* versus *Ever* (which combined *Hardly ever*, *Sometimes*, *Most of the time*, and

3-7

Always). We then made a binary indicator of the scale: staff who responded *Ever* on any item received the *Ever* indicator for the scale. The model-estimated percentages of staff with the *Ever* indicator are presented in **Exhibit 3.7**.

Exhibit 3.7   Percentage of Respondents with the Safety Problems Indicator

Safety Problems Indicator
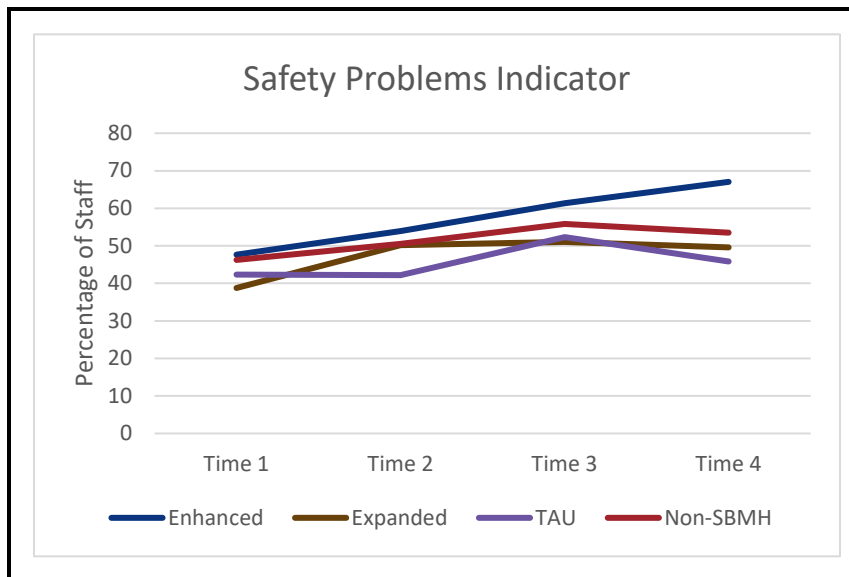
At each of the first three time points, pairwise comparisons between groups found no statistically significant differences. At Time 4, however, the percentage of staff with the Safety Problems indicator continued to increase in Enhanced schools while decreasing slightly for the other three groups; differences were significant between the Enhanced group and the Expanded (p = .009) and TAU (p = .001) groups. (The difference between the Enhanced and non-SBMH groups approached significance, p = .068.)

Staff Observations of Student Behavior. Staff survey staff were asked whether they had personally observed a number of student behaviors in their designated work area (i.e., classroom, cafeteria, or entire school) in the past 30 days. Most germane to this report are staff reports of student bullying, fighting, robbery, and possession of a weapon.

Student Bullying. At Time 1, a significantly higher model-estimated percentage of staff in the Enhanced group reported student bullying than staff in the Expanded group (p = .046) or the TAU group (p = .026; **Exhibit 3.8**). The percentage of staff in Enhanced group who reported student bullying remained

fairly constant over the subsequent survey time points, whereas the percentage in the other groups increased from Time 1 such that the differences between groups were not significant. Relatively speaking, this pattern indicates a better outcome for the Enhanced group compared with the other groups.

Exhibit 3.8   Percentage of Staff Who Observed Student Bullying, Past 30 Days
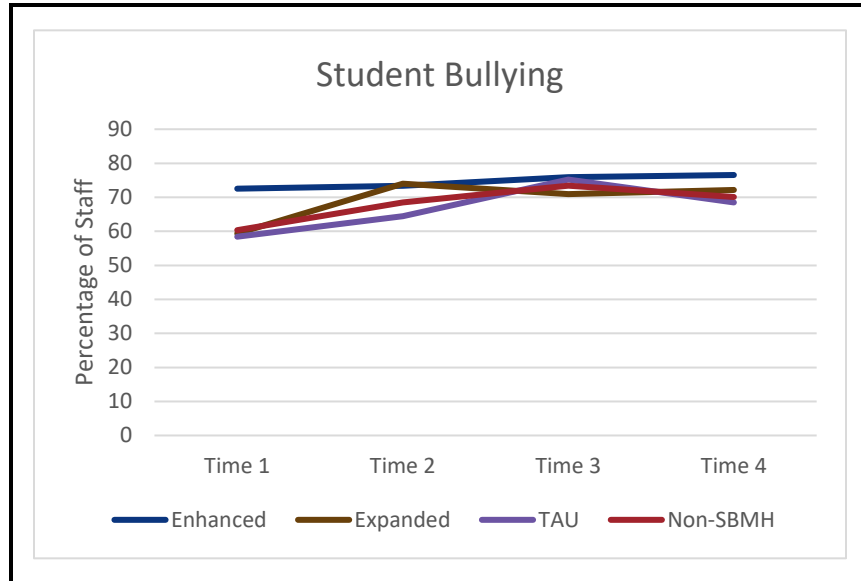
Student Fighting. The percentage of staff who reported student fighting at Time 1 was significantly higher in Enhanced schools than in TAU (p = .011) and non-SBMH (p = .005) schools (***Exhibit 3.9***). At Times 2 and 3, the percentage in Enhanced schools changed very little, whereas it increased in the other three groups; the pairwise comparisons for those two time points found no significant differences between groups. Through Time 3, this pattern indicated a better outcome for the Enhanced group relative to the other groups. At Time 4, however, the percentage increased in Enhanced schools but stayed relatively constant in the other groups. At Time 4, the percentage in Enhanced schools significantly exceeded that in TAU schools (p = .041). (The differences with Expanded schools and Non-SBMH schools approached but did not reach significance, p = .069 and .096, respectively).

Exhibit 3.9   Percentage of Staff Who Observed Student Fighting, Past 30 Days

Student Fighting
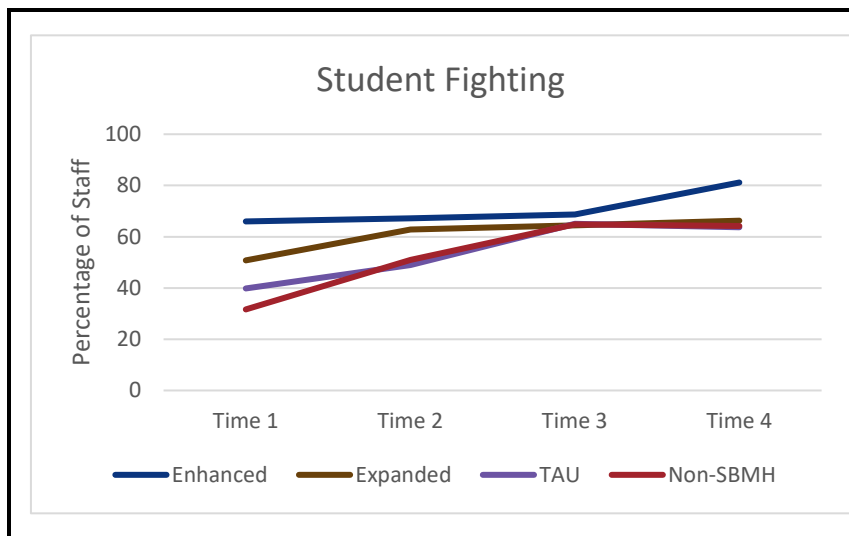
Student Robbery. Regarding staff who reported observing student robbery in the past 30 days (***Exhibit 3.10***), *t*here were no statistically significant differences between treatment groups at any of the four staff survey time points.

Exhibit 3.10   Percentage of Staff Who Observed Student Robbery, Past 30 Days

Student Robbery
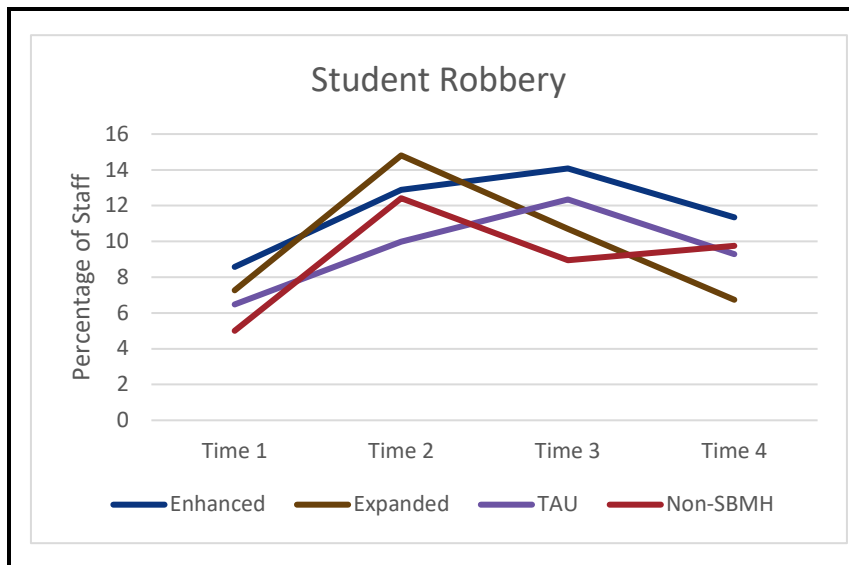
Student Possession of Weapons. The percentage of staff who reported observing student possession of weapons was much lower than the percentage who reported student bullying or fighting. The percentage who observed weapon possession ranged from .04 to .18 (**Exhibit 3.11**), whereas the percentage who observed bullying ranged from .58 to .77 and the percentage who observed fighting ranged from .32 to .81.

At Times 1, 2, and 3, we found no significant differences among the treatment groups. At Time 4, however, the percentage of staff in the Enhanced group who observed student weapon possession increased dramatically, and the difference between the Enhanced and TAU groups was significant (p = .044).

Exhibit 3.11   Percentage of Staff Who Observed Student Weapon Possession, Past 30 Days
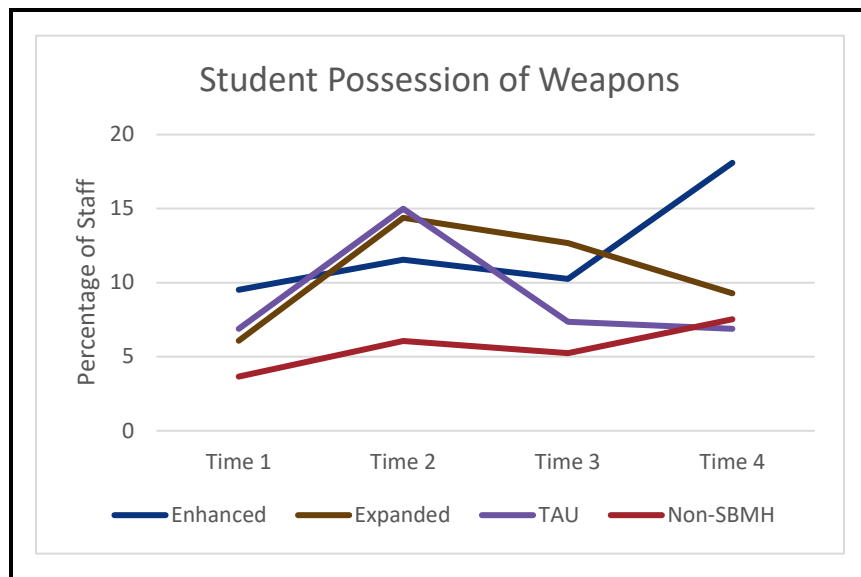
Student Disruptive Behavior. The Disruptiveness Scale of the Inventory of School Climate included items asking staff how often students at their school disrupt what others are doing, are inattentive, bicker and quarrel with each other, are restless, or call out answers out of turn. Likert scale response options ranged from *Never* (scored 1) to *Always* (scored 5), where lower ratings (showing less-frequent disruptive behaviors) are more favorable. Mean ratings are presented in **Exhibit 3.12.**

**Exhibit 3.12   Mean
Staff Ratings of
Frequency of Disruptive
Behaviors by Students**
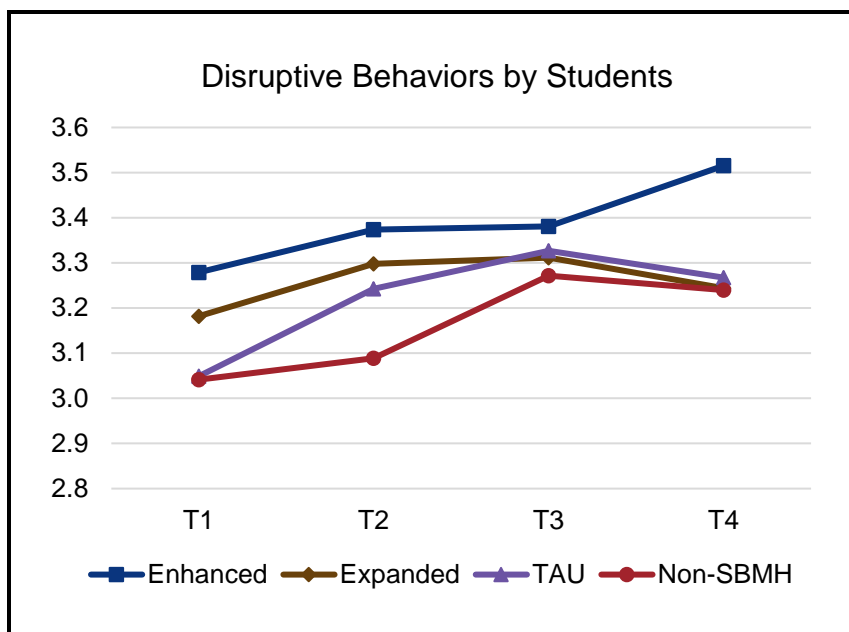
At Time 1, the mean rating on the Disruptiveness Scale was significantly higher for staff in the Enhanced group than in the TAU group (p = .009) and the non-SBMH group (p = .018). At Time 2, the Enhanced group mean remained significantly higher than the non-SBMH group mean (p = .004); the Expanded group mean was also significantly higher than the non-SBMH group mean (p = .045). At Time 3, the group means largely converged, and differences were no longer significant, suggesting that the Enhanced group had improved relative to the other groups. At Time 4, differences between the non-SBMH, TAU, and Expanded group means were still not significant. However, the Enhanced group mean increased sharply and was significantly higher than the Expanded (p = .004), TAU (p = .005), and non-SBMH (p = .005) group means—reversing the Enhanced group's relative improvements at Time 3.

Principal Support of Community Mental Health Professionals. The School Administrator Support Scale in the Expanded School Mental Health Collaboration Instrument asked staff the extent to which they agreed that the principal did each of the following:

- Supports community mental health professionals.
- Communicates with community mental health professionals.
- Advocates on behalf of community mental health professionals working in schools.
- Arranges teacher/community mental health professional meetings.
- Communicates the importance of community mental health professionals to teachers.
- Includes community mental health professionals in school leadership activities.
- Addresses students' mental health needs in meetings with teachers.

Response options ranged from *Strongly disagree* (scored 1) to *Strongly agree* (scored 4). We computed the average rating across the seven items to form a scale score. Mean scale scores are presented in **Exhibit 3.13.**

Exhibit 3.13   Mean Staff Ratings of Principal Support for Community Mental Health Professionals
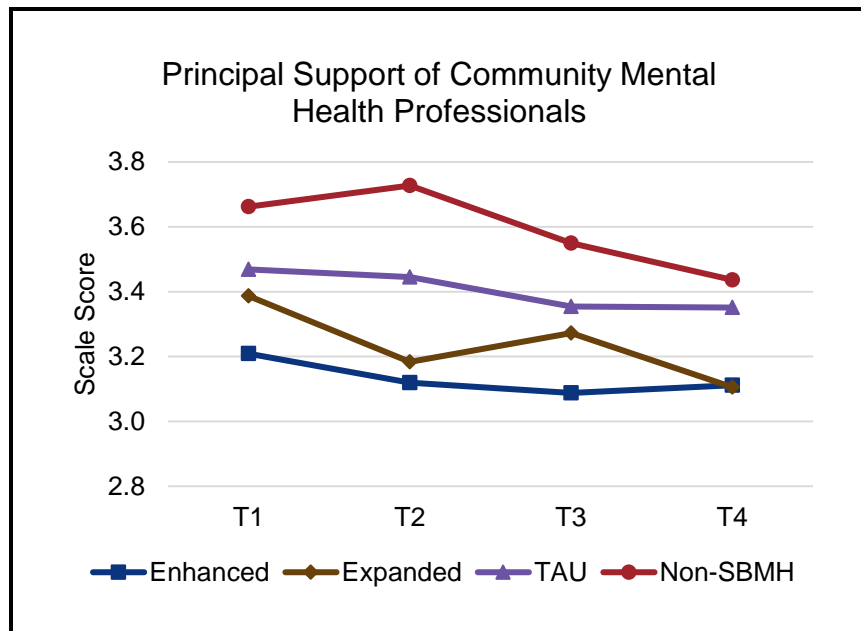
Mean ratings were significantly lower for the Enhanced group than for the non-SBMH group at Time 1 (p = .010), Time 2 (p = .0006), and Time 3 (p = .011). At Time 4, the difference in ratings approached significance (p = .061). In addition, the Enhanced group mean was lower than the TAU group mean at Time 2 (p = .039).

The Expanded group mean was lower than the non-SBMH group mean at Time 2 (p = .004). At Time 4, the difference approached significance (p = .077).

3.2.2 Association Between Staff Survey Outcomes and Mental Health Services

As discussed in **Section 3.1**, the levels of student services provided by psychologists, therapists, and counselors in each school were not always aligned with the expected increases in services for a school's treatment group based on intended implementation of SBMH (see Exhibit 2.2). Therefore, to examine the effect of services on outcomes, we assessed the association between the key staff survey outcomes and five school-level implementation measures (i.e., the amount of services rendered by mental health professionals in the schools).

The implementation measures reflect information gathered from psychologists, therapists, and counselors in each school over the course of the 2016–2017(Time 2), 2017–2018 (Time 3), and 2018–2019 school years (Time 4). Each analysis examined the association between student service levels and outcomes in the same school year (which we refer to as *concurrent* effects) and student service levels in one school year and outcomes in the following school year (which we refer to as *lagged* effects). The service data were recorded throughout the school year, and the outcomes were measured near the end of the school year. Thus, even in the concurrent models, most of the reported services preceded outcome measurement—which helps us to interpret the temporal relationships.

Statistically significant associations are summarized in **Exhibit 3.14**. (For clarity, nonsignificant results are not shown.) Parameters for binary 0/1 outcomes reflect differences (in logit scale from multilevel logistic regression models) in the likelihood of a staff member reporting a "1" on the outcome for a one-unit increase in the implementation predictor. For example, for models based on student support service implementation minutes, the values shown represent the decrease (for negative relationships) in the likelihood of the categorical outcome for staff per minute of service in that school. For the Disruptive Behaviors Scale score, the value shown is the change in the scale score reported by staff

associated with a 1-minute increase in service. For the models based on the percentage of students served, the estimates are interpreted similarly but are larger in magnitude due to the difference in scaling between rates of students served per school versus the average number of minutes of treatment per student per school. The size of the coefficients may be less meaningful than the direction of the association and the fact that these statistically significant associations are unlikely to be due to chance ($p < .05$).

Exhibit 3.14   Associations Between Implementation Levels and Staff Survey Outcomes

Outcome	Psychologist				Therapist				Counselor	
	Minutes		Students		Minutes		Students		Students	
	Con-current	Lagged	Con-current	Lagged	Con-current	Lagged	Con-current	Lagged	Con-current	Lagged
Feeling Unsafe (Binary)										
Before School										
During School					−0.0038				−19.404	
After School	0.00565		2.5741							
Personal Experience (Binary)							−2.4158		−7.0409	−11.7748
Observed Student Behaviors (Binary)										
Bullying			−2.1797							
Fighting			−3.054		−0.00367	−0.00803	−2.929			−15.0378
Robbery	−0.00669		−3.9788				−5.2853		−27.471	
Weapons Possession					−0.01042		−9.947			
Disruptive Behaviors (Score)					−0.00081		−0.6539		−2.3590	−2.9524

Psychologist Service Levels. All significant effects related to psychologist services were in the concurrent models; no significant effects were seen in the lagged models. The percentage of students who received psychologist services was negatively associated with staff reports of student bullying, fighting, and robbery. That is, in schools in which a higher percentage of students received psychologist services, staff were less likely to report observing student bullying, fighting, and robbery than in schools in which a lower percentage of

students received services. Minutes of psychologist service provided were negatively associated with staff reports of student robberies. Less favorably, the percentage of students who received psychologist services and the minutes of psychologist services provided were positively associated with staff being more likely to report feeling unsafe after school.

Therapist Service Levels. All significant effects related to therapist service levels except one were in the concurrent models. The percentage of students who received therapist services was negatively associated with staff reports of student fighting, robbery, and weapon possession; of negative personal experiences with students (e.g., being hit or assaulted); and of student disruptive behaviors.

The minutes of therapist services provided were negatively associated with staff reports of student fighting and weapon possession, feeling unsafe during school hours, and frequency of student disruptive behaviors.

The lagged effect models found one significant effect involving therapist services. Minutes of therapist services provided during one school year were negatively associated with staff reports of student fighting in the following school year.

Counselor Service Levels. Counselor services were measured in terms of the number of students served; data on minutes of service were not available. In the concurrent models, the percentage of students who received counselor services was negatively associated with staff reports of student robbery, negative personal experiences with students, feeling unsafe during school, and frequency of student disruptive behaviors.

The lagged effect models found several significant effects involving counselor services. The percentage of students who received counselor services during one school year was negatively associated with staff reports the following year of student fighting, having negative experiences with students, or student disruptive behaviors.

3.3   STUDENT SURVEY RESULTS

The student survey asked students about their experiences with violence and aggression involving other students, whether they felt unsafe at different times of day in different locations at school (similar to the staff survey items discussed in

Section 3.2), and their perceptions of their school in terms of safety problems and negative peer interactions. In this section, we present findings on those outcomes as a function of the school's treatment group and the levels of mental health services in the school.

Exhibit 3.15 shows demographic information for students who responded to the student survey. At each administration, most students identified as Black or African American, and all samples were well-balanced with respect to age, sex, and Hispanic ethnicity. Grade levels were balanced due to the student survey sampling strategy.

Exhibit 3.15   Student Survey Sample Characteristics, by Time Point

	Survey Administration Period			
Characteristic	**Time 1 (n = 3,997)**	**Time 2 (n = 3,612)**	**Time 3 (n = 2,585)**	**Time 4 (n = 2,448)**
Sex (% Male)	50.1	48.7	49.0	47.3
Age, mean (SD)	12.3 (1.0)	12.7 (1.0)	12.7 (1.0)	12.7 (1.0)
Grade 6 (%)	34.5	38.0	34.8	36.4
Grade 7 (%)	30.5	29.7	34.3	31.7
Grade 8 (%)	34.9	32.3	31.0	31.9
Ethnicity (% Hispanic or Latino)	35.5	37.3	37.5	43.0
White (%)	31.6	30.9	30.9	35.3
Black or African American (%)	61.8	62.3	62.7	57.0
Asian (%)	8.4	8.3	9.2	9.5
American Indian or Alaska Native (%)	8.8	7.7	8.6	8.3
Native Hawaiian or Pacific Islander (%)	3.1	4.3	3.1	4.0

Note: Grade-level percentages may not sum to 100% due to rounding. Respondents were able to identify as more than one race; therefore, percentages do not sum to 100%.

3.3.1   Effects of School Treatment Group on Student Survey Outcomes

As in the staff survey analysis, we first examined the student survey data for differences in outcomes among the school treatment groups. Using the multilevel model with group by time interactions described in Chapter 2, we did pairwise group comparisons (e.g., Enhanced group vs. Expanded group) of the outcomes reported by students in each school treatment group.

Student Aggression and Victimization. Students were asked about their personal experience with violence and aggression, both as perpetrator (which we refer to as aggression) and as victim. On the Aggression Scale, students were asked how often in the past 7 days they were involved with each of the following behaviors. Response options were *0 times*, *1 time*, *2 times*, *3 times*, *4 times*, *5 times*, or *more times* (scored as 6). The victimization items were based on the same behaviors (e.g., "A student teased me to make me angry") and used the same response options.

1. I teased students to make them angry.
2. I got angry very easily with someone.
3. I fought back when someone hit me first.
4. I said things about other kids to make other students laugh.
5. I encouraged other students to fight.
6. I pushed or shoved other students.
7. I was angry most of the day.
8. I got into a physical fight because I was angry.
9. I slapped or kicked someone.
10. I called other students bad names.
11. I threatened to hurt or to hit someone.

The model-estimated group means on the Aggression Scale are presented in *Exhibit 3.16*. At Times 1, 2, and 3, the non-SBMH group mean was significantly lower than the other group means, and the latter three groups were not significantly different from each other. At Time 4, the non-SBMH group mean remained significantly lower than the Enhanced and TAU group means; the difference between the non-SBMH group and the Expanded group was approached statistically significant ($p = .052$). None of the pairwise comparisons between the Enhanced, Expanded, and TAU groups were significant.

Exhibit 3.16   Mean Scores on Self-Reported Aggression, by Condition
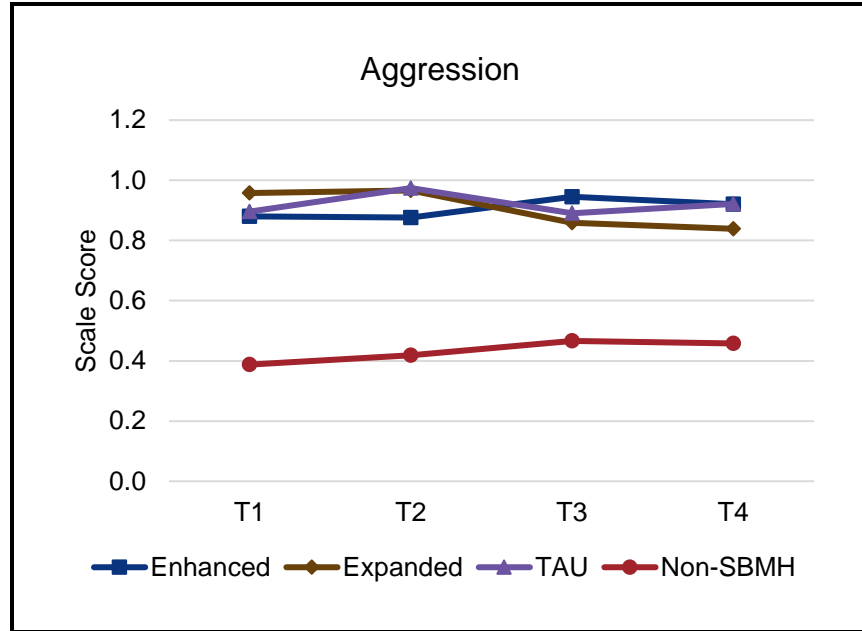
The model-estimated group means on the Victimization Scale are presented in ***Exhibit 3.17***. At Times 1 and 2, the non-SBMH group mean was significantly lower than the Expanded group mean. At Time 3, the Expanded group mean decreased and was no longer significantly different from the non-SBMH group mean. That difference was also not significant at Time 4, reflecting substantial improvement in the Expanded group relative to the non-SBMH group. The TAU group mean was significantly higher than the non-SBMH group mean at Time 1 but not at Times 2 or 3, suggesting relative improvement in the TAU group. However, the TAU group mean increased at Time 4 and was again significantly higher than the non-SBMH group mean. The Enhanced group mean was not significantly different from the non-SBMH group mean at Times 1, 2, or 3; however, at Time 4, the Enhanced group mean increased, and the non-SBMH group mean decreased, resulting in a significantly higher mean in the Enhanced group than in the non-SBMH group. Again, no pairwise comparisons between the Enhanced, Expanded, and TAU groups were significant.

3-19

Exhibit 3.17   Mean Scores on Self-Reported Victimization, by Condition
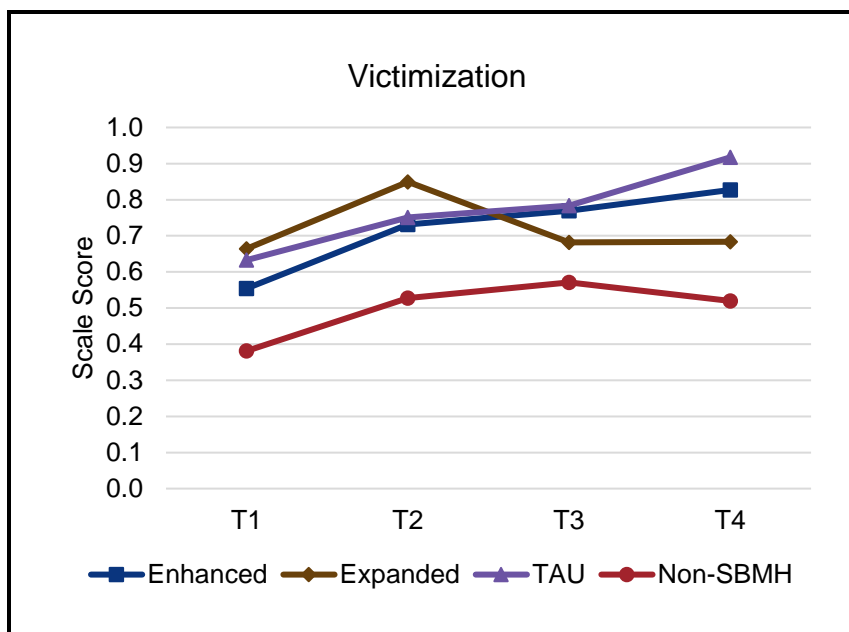
Students Feeling Unsafe. As in the staff survey, students were asked how often during the past 30 days they felt unsafe before, during, and after school hours in each of three locations: in their classroom or office; in other areas of the school building (i.e., hallways, stairwells, restrooms, cafeteria); and in the parking lot, sports fields, or other areas of the school grounds. For each time frame, we formed a single binary indicator if a student reported feeling unsafe in any location.

The model-estimated percentage of students who reported feeling unsafe before school is shown in *Exhibit 3.18*. At Times 1 and 2, the percentage in non-SBMH schools was lower than in Enhanced, Expanded, and TAU schools. At Time 3, the percentage increased substantially in non-SBMH schools, and the difference between those schools and Enhanced and Expanded schools was no longer significant; the difference between non-SBMH and TAU schools remained significant. At Time 4, however, the percentage in non-SBMH schools decreased somewhat, and the difference was again significant between those schools and Enhanced, Expanded, and TAU schools. At all four time points, no pairwise comparisons between the Enhanced, Expanded, or TAU groups were significant.

**Exhibit 3.18   Proportion
of Students Who
Reported Feeling Unsafe
Before School, by
Condition**

Before School
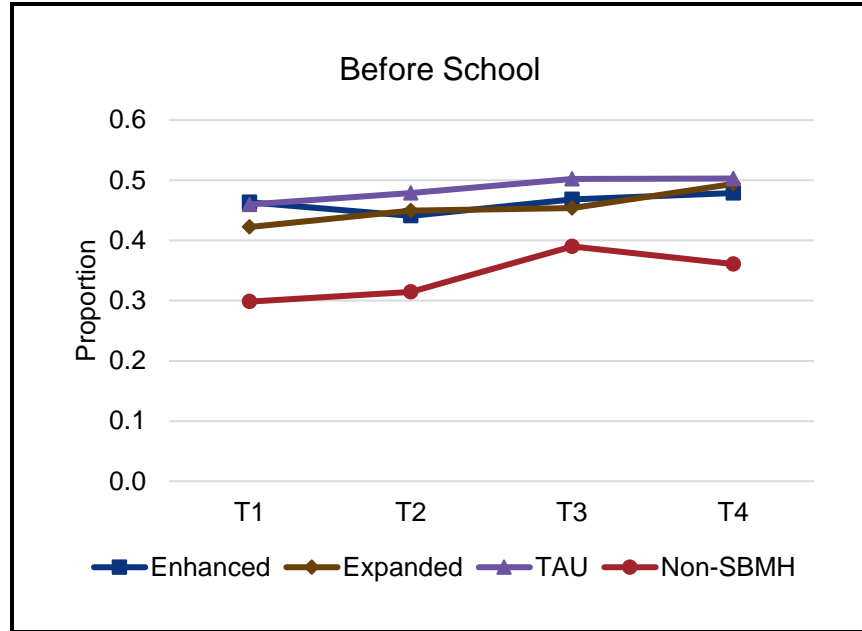
The model-estimated percentage of students who reported feeling unsafe during school is shown in ***Exhibit 3.19***. At Time 1, the percentage in non-SBMH schools was lower than in Enhanced and TAU schools. At Time 2, the percentage in Enhanced schools decreased substantially, and the difference with non-SBMH schools was no longer significant. The difference between non-SBMH schools and TAU schools remained significant, and the difference between non-SBMH schools and Expanded schools became significant. At Time 3, the percentage in non-SBMH schools increased, and the difference between that group and the other groups was no longer significant—showing relative improvements in Enhanced, Expanded, and TAU schools compared with non-SBMH schools. At all four time points, no pairwise comparisons between the Enhanced, Expanded, or TAU groups were significant.

Exhibit 3.19   Proportion of Students Who Reported Feeling Unsafe During School, by Condition
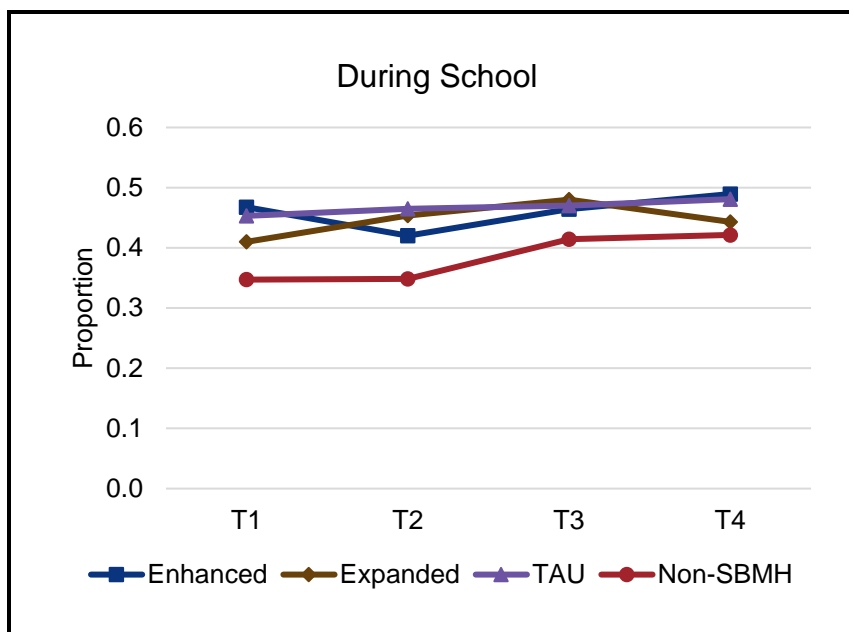
The model-estimated percentage of students who reported feeling unsafe after school is shown in ***Exhibit 3.20***. At Time 1, the percentage in non-SBMH schools was lower than in Enhanced and TAU schools. At Time 2, those differences were again significant, as was the difference between non-SBMH and Expanded schools. At Time 3, the percentage in non-SBMH schools increased somewhat, and the differences were no longer significant. None of the differences at Time 4 were significant, showing relative improvements in Enhanced, Expanded, and TAU schools, compared with non-SBMH schools. At all four time points, no pairwise comparisons between the Enhanced, Expanded, or TAU groups were significant.

Exhibit 3.20 Proportion of Students Who Reported Feeling Unsafe After School, by Condition
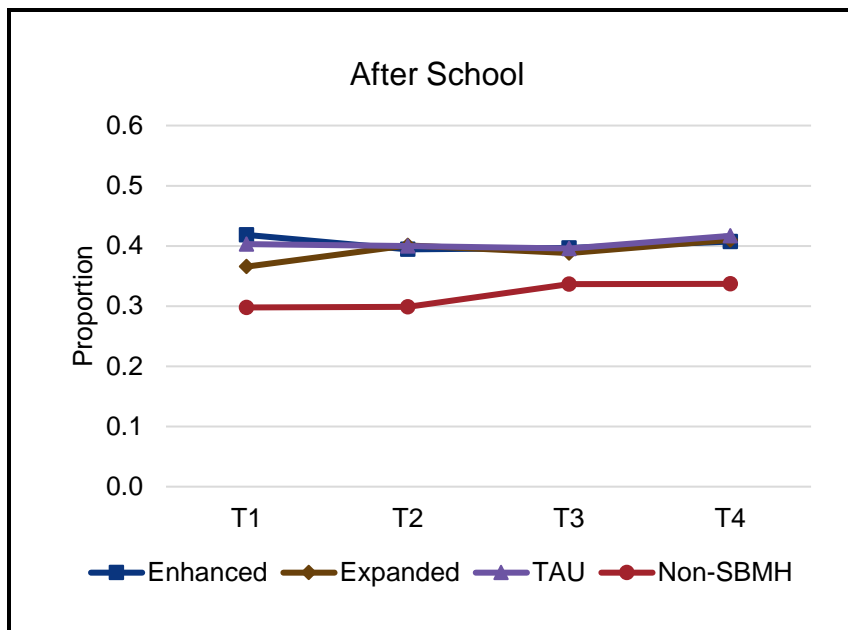
Negative Peer Interactions. To measure negative peer interactions, students were asked how often the following things happen at their school. Response options were *Never* (scored 1), *Hardly ever* (scored 2), *Sometimes* (scored 3), *Most of the time* (scored 4), and *Always* (scored 5). A scale score was computed for each respondent by averaging scores across the items.

1. Students in this school have trouble getting along with each other.

2. Students in this school are mean to each other.

3. In classes, students find it hard to get along with each other.

4. There are students in this school who pick on other students.

5. Students in this school feel students are too mean to them.

The model-estimated group means on the Negative Peer Interaction Scale are shown in *Exhibit 3.21.* At Time 1, the non-SBMH group mean was significantly lower than the Enhanced, Expanded, and TAU group means. At Time 2, the differences between non-SBMH schools and Enhanced and Expanded schools remained significant, but the difference

3-23

between non-SBMH and TAU schools fell slightly short of significance (p = .052). At Time 3, only the difference between Expanded schools and non-SBMH schools remained significant; the differences between the latter schools and the Enhanced and TAU schools were slightly short of significance (p = .07 and .06, respectively). At Time 4, the non-SBMH group mean decreased and was significantly lower than the Enhanced and TAU group means. The Expanded group mean also decreased sharply, and the difference with the non-SBMH group mean fell slightly short of significance (p = .065). Again, no pairwise comparisons between the Enhanced, Expanded, and TAU groups were significant.

Exhibit 3.21 Mean Scores on Self-Reported Negative Peer Interactions, by Condition
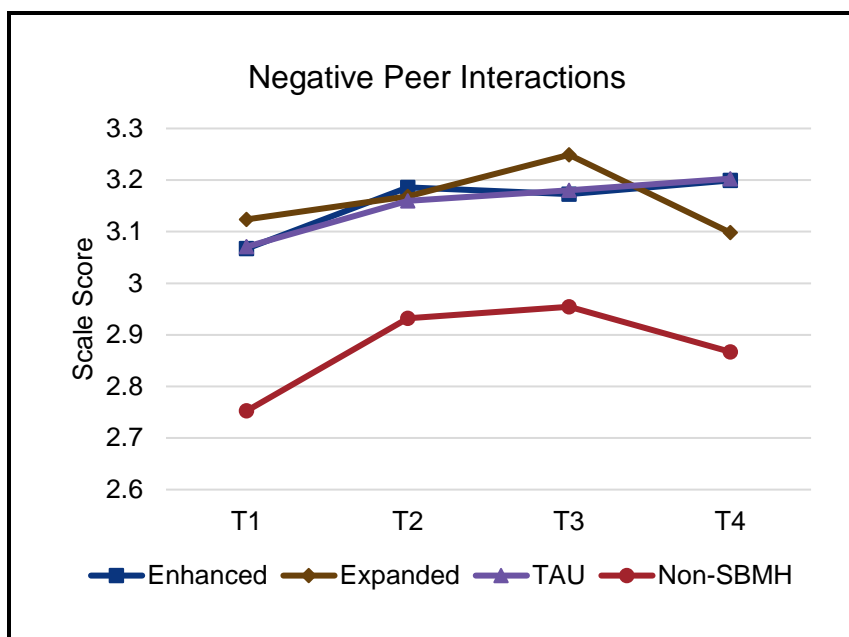
School Safety Problems. Students were asked the extent to which each of the following was a problem in their school. Response options were *Not a problem* (scored 1), *Minor problem* (scored 2), *Moderate problem* (scored 3), and *Serious problem* (scored 4). A scale score was computed for each respondent by averaging scores across the items.

1. Fighting (hitting and kicking) among students
2. Students wrecking school property
3. Students carrying weapons
4. Student disrespect for teachers

5. Racial tension or racism

6. Gangs

7. Unsafe areas in the school

8. Teachers ignore it when students threaten other students

9. Teachers ignore it when students tease other students

10. Teachers not knowing what students are up to

The model-estimated group means on the School Safety Problems Scale are presented in ***Exhibit 3.22.*** At all four time points, the non-SBMH group mean was significantly lower than each of the other group means, and the latter three groups were not significantly different from each other.

Exhibit 3.22   Mean Scores on Perceptions of School Safety Problems, by Condition
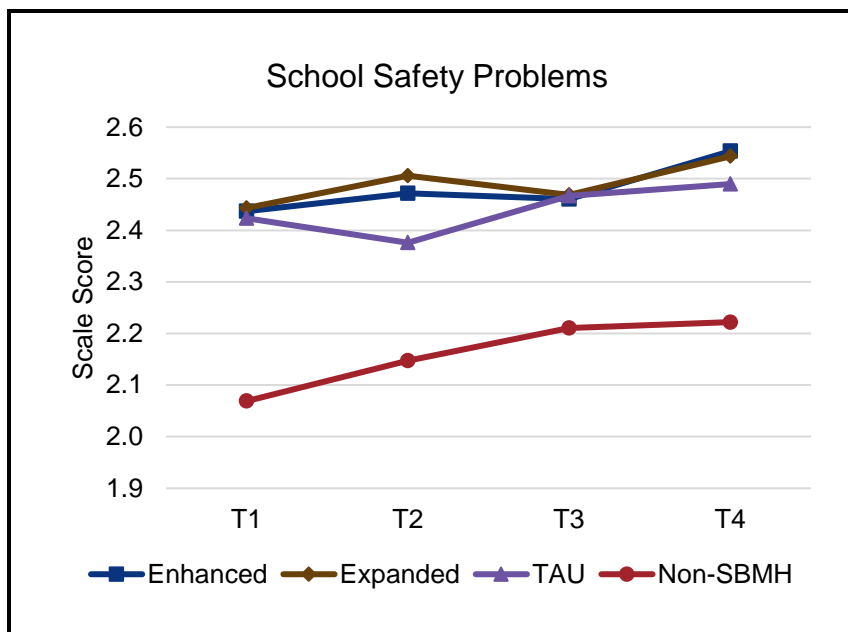
3.3.2   Association Between Student Survey Outcomes and Mental Health Services

Mirroring the staff survey data analysis, we examined the effect of school mental health services on student outcomes by analyzing the association between the key student survey outcomes and five school-level implementation measures (i.e., the amount of services rendered by mental health professionals in the schools). We again examined the association between service levels and outcomes in the same school year (*concurrent* effects) and service levels in one school year and outcomes in the following school year (*lagged* effects).

3-25

Statistically significant associations are summarized in
Exhibit 3.23.

Exhibit 3.23   Associations Between Service Levels and Student Survey Outcomes

Student Outcome	Psychologist				Therapist				Counselor	
	Minutes		Students		Minutes		Students		Students	
	Con-current	Lagged	Con-current	Lagged	Con-current	Lagged	Con-current	Lagged	Con-current	Lagged
Aggression (score)							0.837 (.047)			
Victimization (score)			.5867 (.012)		.000925 (.034)		1.2929 (.0004)			
Feeling Unsafe (binary)										
Before School										
During School										
After School										
Negative Peer Interaction (score)	−0.0066 (.041)									
School Safety Problems (score)	−.0019 (.0002)		−0.5224 (.001)		−0.00120 (.049)					

Psychologist Service Levels. All significant effects related to psychologist services were in the concurrent models; no significant effects were seen in the lagged models. The percentage of students in a school who received psychologist services was negatively associated with student ratings of school safety problems in that school. That is, in schools in which a higher percentage of students received psychologist services, students rated school safety as less of a problem, compared with schools in which a lower percentage of students received psychologist services. Also, minutes of psychologist service provided were negatively associated with student ratings of school safety problems and frequency of negative peer interactions. Less favorably, the percentage of students who received psychologist services was positively associated with students reporting more victimization (i.e., higher scores on the Victimization Scale).

Therapist Service Levels. Similar to the unfavorable finding regarding psychologist services, the percentage of students who received therapist services was positively associated with

students reporting more victimization as well as more aggression. Minutes of therapist services received was also positively associated with students reporting higher levels of victimization. More favorably, minutes of therapist services provided during one school year were negatively associated with student ratings of school safety problems the following school year.

Counselor Service Levels. The percentage of students who received counselor services was not significantly associated with any of the student outcomes examined. Two concurrent associations approached significance: (1) the percentage of students who received counselor services was negatively associated with student self-reported aggression ($p = .07$) and ratings of school safety problems ($p = .08$), and (2) the percentage of students who received counselor services one year was negatively associated with student ratings of school safety problems the following year ($p = .07$).

3.4   STUDENT INFRACTIONS RESULTS

As described in *Section 2*, we used student disciplinary infractions data provided by CMS to examine differences among the treatment groups during each of 4 school years. Note that in the following discussion, the 2015/16 school year is considered a baseline period, prior to implementation of study-supported services that started during the 2016/17 school year. For analysis purposes, this 2015/16 school year infractions baseline is functionally similar to the staff and student baseline surveys conducted in the fall of the 2016/17 school year, immediately prior to the start of study-supported services.

These analyses focused on the infraction categories described earlier: disruption; insubordination or disrespect; harassment, threats, or bullying; aggressive behavior; fighting or assault; and weapons or gangs. For each category, we examined two types of measures:

For each school and in each study year, the percentage of students in grades 6 through 8 who received one or more disciplinary infraction reports; we refer to this measure as "binary" because each student either did or did not receive one or more infraction reports.

3-27

For students who received one or more infraction reports in a category, the number of reports in that category that each student received; we refer to this measure as "count."

For each category and data type, we conducted two complementary sets of analyses (similar to our analysis of the staff and student survey data): 1) pairwise comparisons between each of the treatment groups for each time point, and 2) the association between each outcome and each of the mental health service measures, i.e. the percentage of students who received the service and the minutes of service provided per student.

Also, we explored conducting similar analyses using outcomes based on the single most frequent infraction in each category, rather than the broader category. Some of these statistical models would not converge. In all cases where the models converged, there were no differences in inference compared with the models using the infraction categories that we discuss below. Therefore, we do not present results from these supplemental analyses.

3.4.1 Effects of School Treatment Group on Student Infractions Outcomes

For both the binary and the count measures, there were numerous significant differences involving the Non-SBMH group. For each of the binary outcomes, more than half of the 12 pairwise contrasts involving the Non-SBMH group (contrasts between the Non-SBMH group and each of the other three groups at each of four time points) showed statistically significant differences (see ***Exhibit 3.24***).

Exhibit 3.24   Significant Differences involving the Non-SBMH Group

	Significant Differences in Binary Measures (of 12 Contrasts Examined)		Significant Differences in Count Measures (of 12 Contrasts Examined)	
	Number	**Percentage**	**Number**	**Percentage**
Insubordination or Disrespect	11	92%	5	42%
Fighting or Assault	8	67%	3	25%
Harassment, Threats, or Bullying	7	58%	3	25%
Aggressive Behavior	9	75%	2	17%
Disruption	8	67%	3	25%
Weapons or Gangs	7	58%	0	0%

As will be seen in the graphs presented below, the percentages of students with one or more infractions in a category was almost always lower in the Non-SBMH group than the other three treatment groups. This is not surprising: SBMH services are typically provided in schools with more students with the greatest need, and that need is often related to many student outcomes, including disciplinary infractions.

Fewer significant differences between Non-SBMH and SBMH groups were found for the count measures (though significant differences were still more common than would be expected by chance, 5%). This finding may be due in part to these analyses including far fewer students than the analysis of binary outcomes: fewer than 10% of students had one or more infractions in each category, with fewer than 1% for the Weapons or Gangs category. The finding may also reflect the likelihood that the number of infractions per student does not differ as greatly between different types of schools as does the percentage of students with one or more infractions.

Given the large pre-existing and ongoing differences between the Non-SBMH group and the other treatment groups on the binary outcomes, we focus the remainder of this discussion on differences among the Enhanced, Expanded, and TAU groups. For completeness, we include the Non-SBMH group in the graphs below showing the binary and count outcomes. Our discussion, however, focuses on differences among the other three treatment groups, which had been randomly assigned to condition.

Insubordination or Disrespect. On the binary measure (**Exhibit 3.25**), in 2015/16 the percentage of students with at least one infraction in the Insubordination or Disrespect category was significantly lower in the TAU group than in the Enhanced group ($p < .0001$) or the Expanded group ($p = .001$). Both of these differences remained significant in 2016/17 ($p < .0001$ and $p = .002$, respectively) but in 2017/18 and 2018/19 they were no longer significant, due mainly to the TAU group percentage increasing and, to a lesser extent, the Enhanced group percentage decreasing.

On the count measure of infractions per student (**Exhibit 3.26**), in 2015/16 the TAU group count was lower than the Expanded count ($p = .004$). This difference remained

significant in 2016/17 2 (p = .002) and the TAU group count was also lower than the Enhanced group (p = .022). In 2017/18 and 2018/19 these differences were no longer significant, due mainly to the TAU group count increasing and, to a lesser extent, the Enhanced group count decreasing.
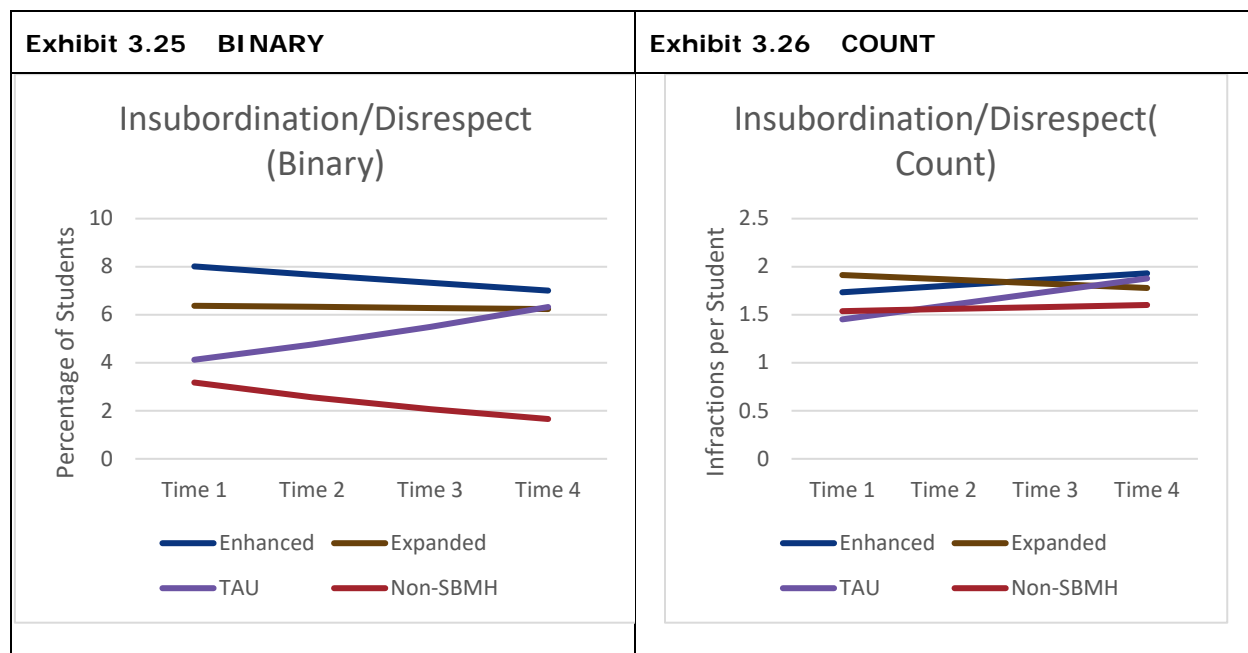
Exhibit 3.25 BINARY	Exhibit 3.26 COUNT

For both the binary and count measures, the pattern of trends suggests better outcomes in the Enhanced and Expanded groups relative to the TAU group.

Fighting or Assault. On the binary measure of the percentage of students with one or more infractions in the Fighting or Assault category (***Exhibit 3.27***), in 2015/16 the Enhanced group percentage was significantly higher than the percentage for the Expanded group (p = .003) or the TAU group (p = .007). In 2016/17, both differences remained significant (p = .003 and p < .0001, respectively). In 2017/18 the Enhanced group percentage remained higher than the TAU group percentage (p < .0001) but did not significantly differ from the Expanded group percentage. In 2018/19, the percentages did not significantly differ in any of the pairwise contrasts between these three groups. This pattern of trends suggests better outcomes in the Enhanced group relative to the Expanded and TAU groups.

On the Fighting or Assault category count measure (***Exhibit 3.28***), in 2015/16 and 2016/17, there were no significant differences among the Enhanced, Expanded, or TAU groups. Due to decreases in counts for the TAU group and slighter increases for the Enhanced group, the former group count was significantly lower than the latter group count in 2017/18 (p = .024) and 2018/19 (p = .035), suggesting better outcomes in the TAU group relative to the Enhanced group.

Exhibit 3.27 Fighting BINARY	Exhibit 3.28 Fighting COUNT
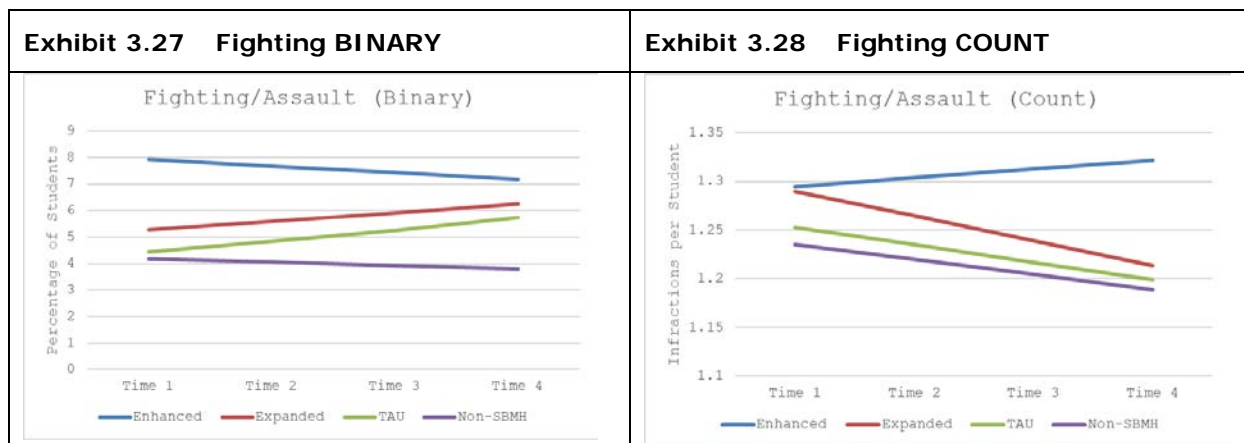
For the Fighting or Assault category, trend patterns differed between the binary and count outcomes. On the binary measure, the Enhanced group improved relative to the Expanded group or the TAU group in 2018/19. However, on the count measure, the Enhanced group worsened relative to the TAU group in 2017/18 and 2018/19. Relatively speaking, in the Enhanced group fewer students received at least one Fighting or Assault category infractions, but those who did receive at least one received slightly more, on average. It is worth noting that for these three groups the range in Fighting or Assault category counts is quite small, 1.20 to 1.32.

Harassment, Threats, or Bullying. For the Harassment, Threats, or Bullying category, on the binary measure there were no significant differences among the Enhanced, Expanded, or TAU groups at any time point (***Exhibit 3.29***).

On the count measure (***Exhibit 3.30***), the Enhanced group count was higher than the TAU group count in 2015/16 (p = .001), 2016/17 (p = .003), and 2017/18 (p = .023). The Enhanced group count was also higher than the Expanded group count in 2016/17 (p = .023), 2017/18 (p = .005) and

3-31

2018/19 (p = .007). With some slight deviations, the Enhanced group count was consistently higher than the counts for the Expanded group or the TAU group.

Exhibit 3.29 Harass. BINARY	Exhibit 3.30 Harass. COUNT
Harassment/Threats/Bullying (Binary)	Harassment/Threats/Bullying (Count)
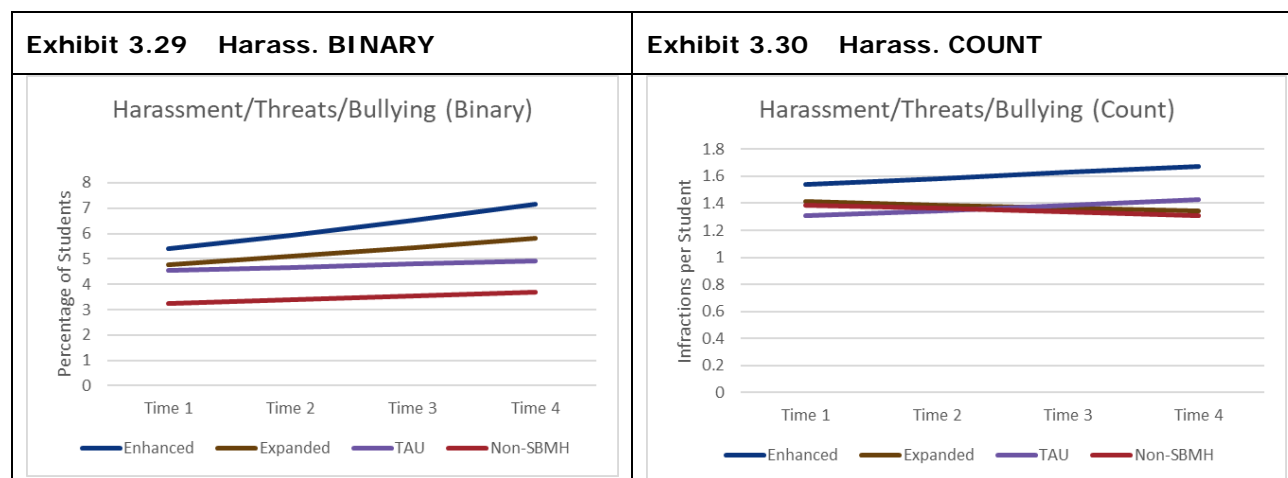
Aggressive Behavior. For the Aggressive Behavior category, on the binary measure there were no significant differences among the Enhanced, Expanded, or TAU groups at any time point (***Exhibit 3.31***).

On the Aggressive Behavior count measure (***Exhibit 3.32***), in 2016/17 the Enhanced group mean count was higher than the Expanded group mean count (p = .038) and the TAU group mean count (p = .033). In Exhibit Aggress. Count, the mean counts for the Enhanced group increased more than the TAU group in 2017/18 and 2018/19, which would seem to suggest that the Enhanced group mean count remained significantly higher. However, the differences were not significant in 2017/18 or 2018/9, likely due to high variability: the mean counts for a small number of schools within the groups became higher over time and drove much of the change from 2016/17 to 2018/19 but other schools within those groups had much smaller increases, stayed the same, or even decreased. This larger variability (standard error in statistical terms) led to the differences being statistically nonsignificant.

Exhibit 3.31 Aggress. BINARY	Exhibit 3.32 Aggress. COUNT
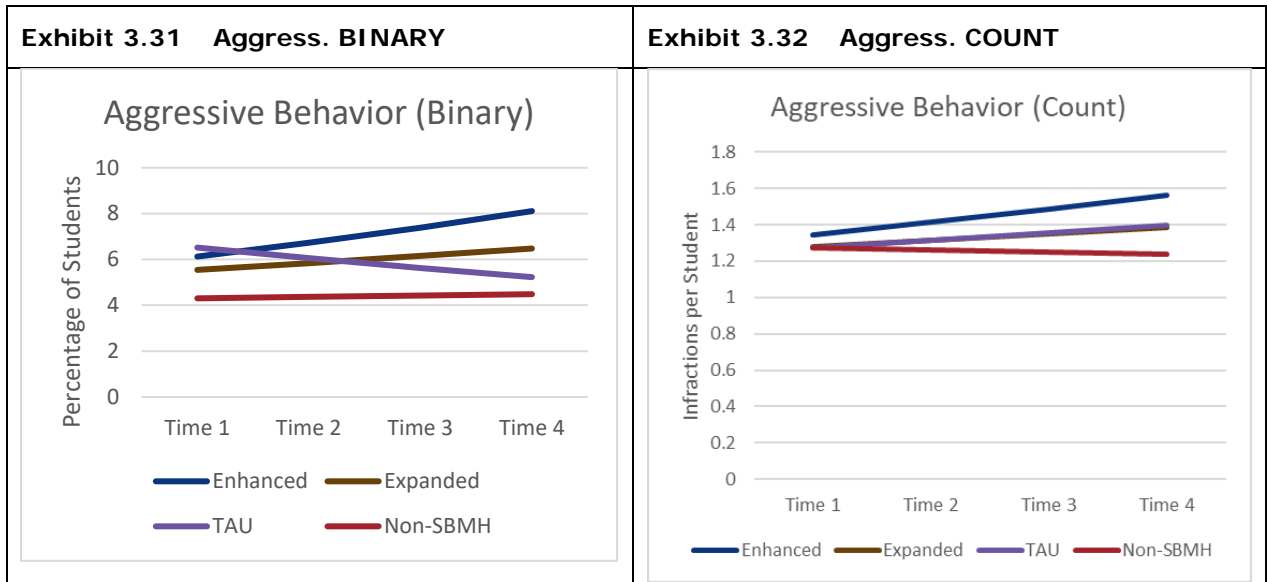
Disruption. For the Disruption category, there were no significant differences among the Enhanced, Expanded, or TAU groups at any time point on either the binary measure (***Exhibit 3.33***) or the count measure (***Exhibit 3.34***).

Exhibit 3.33 Disrupt. BINARY	Exhibit 3.34 Disrupt. COUNT
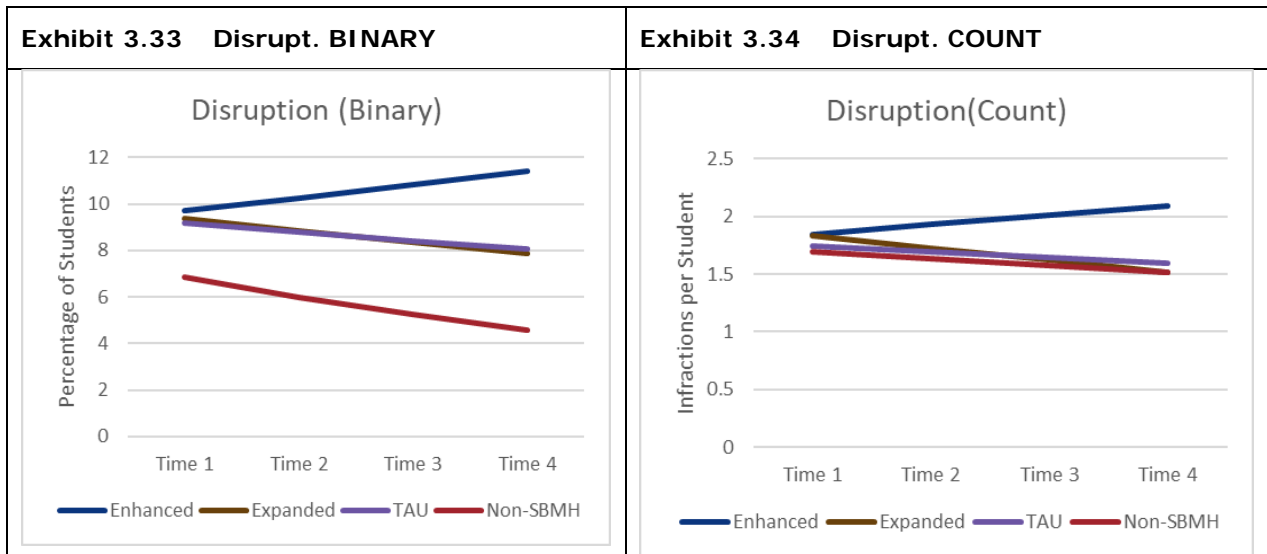
Weapons or Gangs. We attempted to analyze infractions involving weapons or gangs in the same manner as the other infractions discussed above. However, infractions involving weapons or gangs were very rare and the statistical models were extremely unstable, leading us to question whether the results were meaningful and reliable. Therefore, we are not presenting and interpreting these results.

3-33

3.4.2 Association Between Student Infractions Outcomes and Mental Health Services

Next, we present the results of analyses that examined the association between each infraction category outcome and each of the mental health service measures, i.e. the percentage of students who received the service and the minutes of service provided per student. Service levels were not often associated with infraction outcomes, for either the concurrent or lagged models.

The concurrent models found four significant associations, 7% of the 60 models tested, just slightly above chance. Three of the significant models found that higher service levels were associated with worse infractions outcomes:

- In schools in which more minutes of psychologist services were provided (compared with schools in which fewer minutes of psychologist services were provided), a higher percentage of students were charged Disruption infractions (p = .032).

- In schools in which more minutes of therapist services were provided (compared with schools in which fewer minutes of therapist services were provided), a higher percentage of students were charged with Harassment, Threats, or Bullying infractions (p = .048).

- In schools in which a higher percentage of students received psychologist services (compared with school in which a lower percentage of students received psychologist services), students who were charged with one or more Aggressive Behavior infractions were charged with more of those infractions, on average (p = .038).

Only one concurrent model found that higher service levels were associated with better infractions outcomes: in schools in which a higher percentage of students received counselor services (compared with school in which a lower percentage of students received counselor services), students who were charged with one or more Disruption infractions were charged with more of those infractions, on average (p = .037).

The lagged models also found four significant associations, which was again 7% of the 60 models tested, just slightly above chance. All four of the significant models found that

higher service levels were associated with worse infractions count outcomes the following year:

In schools in which a higher percentage of students received psychologist services (compared with school in which a lower percentage of students received psychologist services), the following year students who were charged with one or more Aggressive Behavior infractions were charged with more of those infractions, on average (p = .046).

In schools in which a higher percentage of students received therapist services (compared with school in which a lower percentage of students received therapist services), the following year students who were charged with one or more Insubordination or Disrespect infractions were charged with more of those infractions, on average (p = .026).

In schools in which more minutes of therapist services were provided (compared with schools in which fewer minutes of therapist services were provided), the following year students who were charged with one or more Insubordination or Disrespect infractions were charged with more of those infractions, on average (p = .001).

In schools in which more minutes of psychologist services were provided (compared with schools in which fewer minutes of psychologist services were provided), the following year students who were charged with one or more Disruption infractions were charged with more of those infractions, on average (p = .031).

3.5   ECONOMIC ANALYSIS RESULTS

In this section we first discuss the results for the start-up costs of the Expanded and Enhanced approaches, which are the value of resources, such as training, which are used before any student is engaged with the approaches. We then describe the ongoing costs and show how they are used in the cost-effectiveness analysis (CEA). The section concludes with a sensitivity analysis of the findings to see how robust the conclusions are to changing a key cost estimation assumption.

3.5.1 Start-up Costs

Exhibit 3.35 presents the start-up costs per school for the Expanded and Enhanced conditions. The primary start-up cost

driver is training for SPARCS and DBT, totaling $43,852 per school. These costs are only incurred by Enhanced schools and represent the majority of all start-up costs. Costs are presented by school to help administrators anticipate costs for their districts. We estimated per school costs by dividing total start-up costs by the number of schools; this approach is the reason for equal costs across conditions in some of the categories. TAU schools did not expand or change their services in any way and therefore incurred no start-up costs.

Exhibit 3.35   Start-up Costs per School (2020 dollars)

Activity	Expanded (N = 8)	Enhanced (N = 8)
Training and related activities	$0	$43,852
Hiring activities	$1,800	$1,800
Meetings with school staff and administrators	$489	$489
Materials	$59	$2,610
Other (e.g., legal, information technology, staff selection)	$56	$56
Total	*$2,403*	*$48,807*

Note: EX: Expanded, EN: Enhanced; SBMH, school-based mental health.

3.5.2 Ongoing Costs

Exhibit 3.36 presents ongoing costs in a box plot. Each school year and study condition has a box, and that box comprises a median (middle bar), and the 25th and 75th percentiles (lower and upper bars, which form the end of the box.) The plot also shows the mean as a short dash and the cost-per-student estimate for each school (circles). The potential outlier values are the circles beyond the boxes.

The Expanded and Enhanced groups have a relatively tight spread of cost estimates in each year, while the TAU group has a wider spread. The same two schools in the TAU have outlier high cost-per-student estimates from year to year. Compared to other schools in the treatment arm, these two schools have fewer students but the same total costs; the small number of students therefore results in a higher cost per student estimate. Given the small number of participating schools in the study, we are unable to perform further statistical adjusting to account for the imbalance in student counts across conditions.

Exhibit 3.36   Ongoing Costs per Student by Year (2020 Dollars)
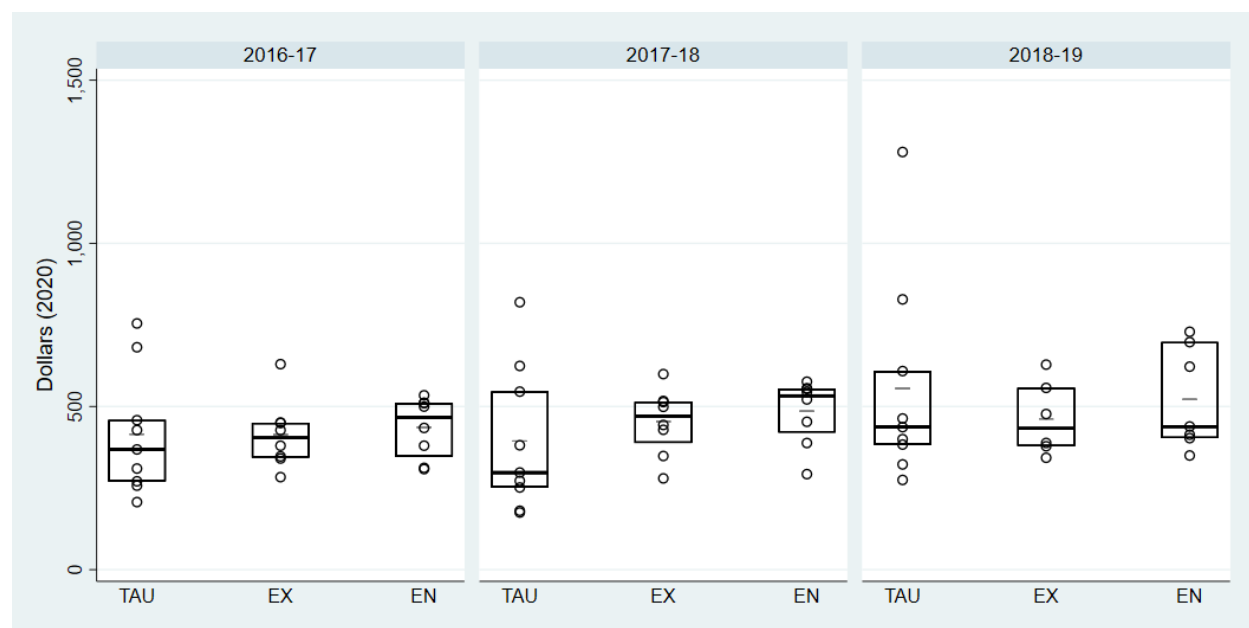
TAU: Treatment as usual, EX: Expanded, EN: Enhanced

We present the ongoing costs per student for each year and averaged across years in ***Exhibit 3.37***. We use medians to represent central tendency in the estimates because of the outlier costs of the two TAU schools. For 2016–17 and 2017–18 the costs align with the intervention levels, where TAU has the lowest average cost, the Expanded group is higher, and the Enhanced group is the highest. In 2018–19 median costs are similar across all arms. The last row of the column displays average annual costs.

Exhibit 3.37   Ongoing Costs per Student, Average Annual and by Year (2020 dollars)

	2016–17			2017–18			2018–19			Avg Ann
	N	Median	Min	Max	N	Median	Min	Max	N	Median	Min	Max	
TAU	9	$368	$207	$755	9	$298	$175	$820	9	$437	$275	$1,280	$368
Expanded	8	$404	$284	$630	8	$471	$280	$599	6	$433	$343	$628	$436
Enhanced	8	$467	$308	$534	8	$533	$293	$576	7	$439	$350	$729	$480

TAU: Treatment as usual; SD: Standard deviation; Avg Ann: Average Annual

Exhibit 3.38 presents median ongoing costs over time. Years 2016–17 and 2017–18 show costs as we might expect: Enhanced is higher than Expanded and both increase slightly, while TAU is lower than both and falls in the second year. Year

2018–19 shows all three converge to the same cost per student.

Exhibit 3.38   Median Ongoing Costs per Student by Year (2020 dollars)
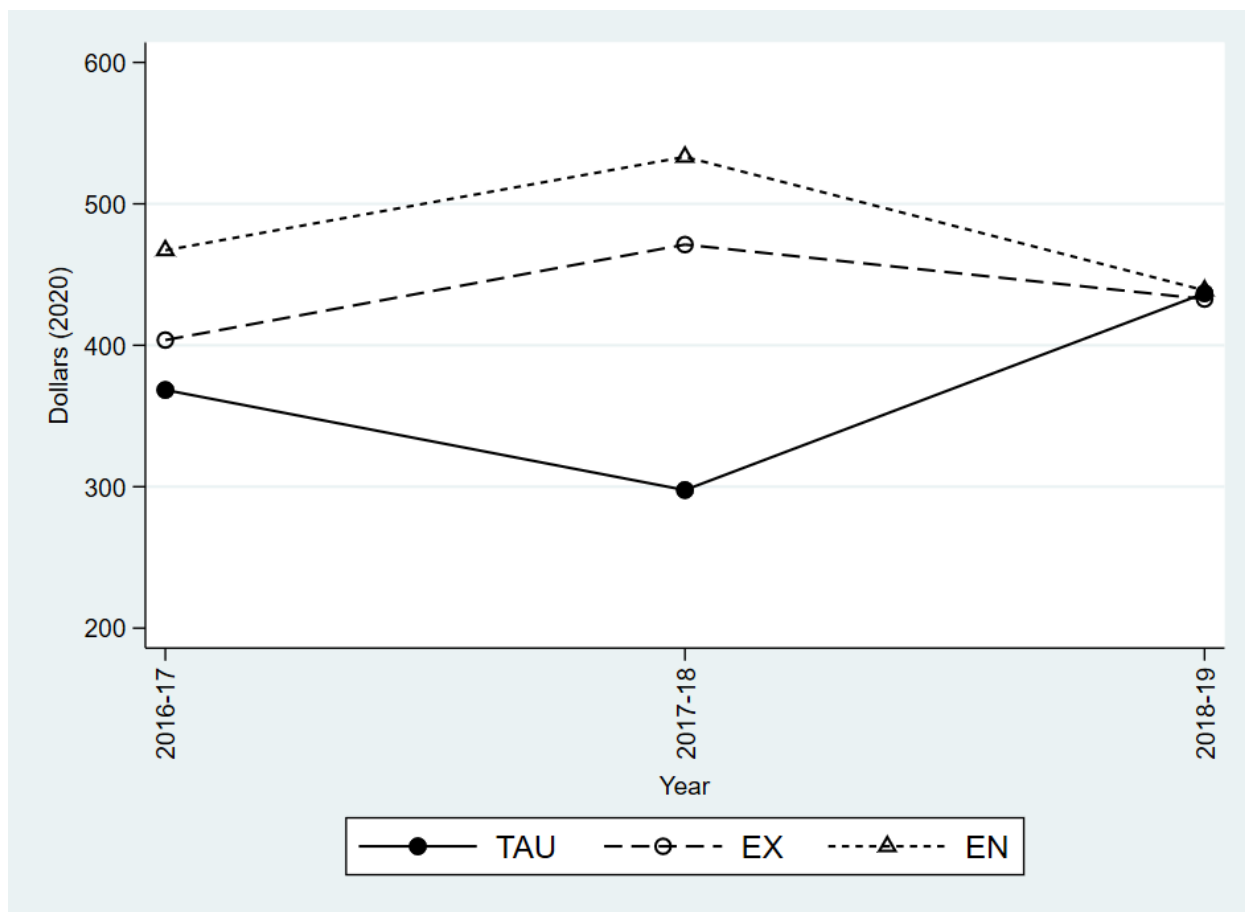
TAU: Treatment as usual, EX: Expanded, EN: Enhanced

3.5.3 Cost-effectiveness Analysis

The CEA combines the cost estimates presented in
Exhibit 3.37 with predicted values of the victimization scale
from the outcomes study, shown in ***Exhibit 3.39***. The
outcomes are counts of victimization events collected from the
student survey and then adjusted for covariates in the
outcomes study. A lower value represents a better outcome,
and the estimates indicate that both the Expanded and
Enhanced conditions reduce victimization compared to TAU.

Exhibit 3.39 Predicted Victimization Scale, by Year and Average Annual

Victimization				
	2016–17	2017–18	2018–19	Avg Ann
TAU	0.731	0.805	0.879	0.805
Expanded	0.775	0.724	0.674	0.724
Enhanced	0.724	0.748	0.772	0.748

TAU: Treatment as usual; Avg Ann: Average Annual

To compute the CEA, we combine the estimates for costs and victimization averaged over three years. *Exhibit 3.40* presents the steps of the CEA. The first panel lists the per student cost and victimization for each treatment group, ranked in ascending order of cost.

The second panel then shows the pairwise differences in costs and victimization between study condition. The final column of that panel shows how these pairwise differences translate to an incremental cost-effectiveness ratio (ICER), which divides the difference in cost by the difference in victimization. In any two comparisons of treatment groups, a lower ICER is preferred because it means that it costs less to achieve a given improvement in victimization.

Exhibit 3.40 Costs and Effectiveness of Treatment Groups Ranked in Order of Cost

Group	Cost	Victimization	
TAU	$368	0.805	
EX	$436	0.724	
EN	$480	0.748	

	Difference in cost	Difference in victimization	Incremental cost-effectiveness ratio (ICER)
EX-TAU	$68	-0.081	-$847
EN-TAU	$112	-0.057	-$1,971
EN-EX	$44	0.024	$1,855

TAU: Treatment as usual, EX: Expanded, EN: Enhanced

To interpret the results in the second panel of **Exhibit 3.40**, consider the following example. The first row indicates that relative to TAU the Expanded treatment group both costs $68 more per student per year and is associated with a decrease in the victimization scale of 0.081 victimization incidents per student per week. The ICER indicates that it costs $847 per year to achieve a one-unit improvement in this victimization

3-39

scale when using the Expanded treatment approach rather than TAU.

The second row shows that relative to treatment as usual, the Enhanced treatment group costs $112 more per student per year and is associated with a decrease in the victimization scale of 0.057 victimization incidents per student per week. The ICER indicates a rate of $1,971 per year to achieve a one-unit decrease in the victimization scale. Finally, the last row shows that the Enhanced group costs more but is less effective than the Expanded group. For this reason, a decision maker seeking to make the optimal cost-effective choice would not choose the Enhanced approach.

The CEA estimates have significant practical value to school administrators. The results suggest that at a hypothetical school of 1,000 students, using a $6,800 grant on the Expanded approach would result in nearly 300 fewer victimization events at the school that year.

How, then, to interpret the results in a way that means something to school administrators? Consider the following hypothetical example. In a school of 1,000 students, an administrator has a grant of $6,800 on school-based mental health to reduce victimization and wants to optimally spend that grant. The CEA findings indicate that relative to what the school is doing now, for every $68 spent on a student using the Expanded approach, weekly victimization is reduced by 0.081 units – or events – per student per week. In this hypothetical school, the administrator would spend $6.80 per student (6,800/1,000 = 6.8), and in turn, that would translate to 0.0081 events per student per week [(6.8/68) * 0.081 = 0.0081]. Given that there are 37 weeks in the school year and 1,000 students in the school, that effectiveness impact means that the grant would be expected to result in nearly 300 fewer victimization events (1,000 * 0.0081 * 37 = 299.7) in the school that year.

Although, the CEA estimates do not suggest the Enhanced approach is optimal — the Expanded is less expensive and more effective — the Enhanced approach is still more effective than the TAU. Administrators could still choose to use the Enhanced approach and improve victimization. Continuing the example of the same hypothetical school, spending the $6,800 grant on the Enhanced approach would result in approximately

128 fewer victimizations than TAU at the school of 1,000 students (see ***Exhibit 3.41***).

Exhibit 3.41 Illustrative Example: 1,000 Student School with a $6,800 Grant

Use the Grant for the Expanded Approach	300 fewer victimization events than TAU
Use the Grant for the Enhanced Approach	128 fewer victimization events than TAU

TAU: Treatment as usual

3.5.4 Sensitivity Analysis

In the base case analysis, we followed standard practice and excluded start-up costs from the CEA. The reason for this is that CEA is a marginal analysis, where the costs and effectiveness for the next student are relevant, whereas start-up costs are the same regardless of the number of students to which the approach is applied. Nevertheless, Table 3.35 above shows that the start-up costs for the Enhanced condition are substantial. We performed a sensitivity analysis to determine the degree to which including the costs would change our conclusions.

We instead amortized start-up costs by condition over the 3 study years and allocating start-up costs across schools and then reassessed the CEA. The results in ***Exhibit 3.42*** show that the sensitivity analysis does not change the conclusions from the base case. The ranking by cost of treatment group does not change, nor does eliminating the Enhanced approach in the interests of an optimal choice. Because additional costs are applied to the estimates, the ICERs are slightly larger than in the base case.

Exhibit 3.42 Cost-effectiveness Analysis

Group	Cost	Victimization
TAU	$368	0.805
EX	$437	0.724
EN	$503	0.748

	Difference in cost	Difference in victimization	Incremental cost-effectiveness ratio (ICER)

EX-TAU	$70	-0.081	-$863
EN-TAU	$136	-0.057	-$2,386
EN-EX	$66	0.024	$2,799

TAU: Treatment as usual, EX: Expanded, EN: Enhanced

The above example of the school with 1,000 students and a grant of $6,800 can be extended to the sensitivity analysis. In this case, the estimates indicate that spending the grant on the Expanded approach results in 291 fewer victimization events.

3.6   PROVIDER SURVEY RESULTS

As described in **Section 2**, the provider survey included measures of school safety that were also included in the staff survey. For analysis, we combined the provider survey data for those measures with the staff survey data; the results were presented in **Section 3.2**.

This section presents provider survey results on three constructs measured only in the provider survey: attitudes toward EBPs, self-efficacy regarding suicidality, and perceptions of organizational capacity. The number of respondents at each time point was less for the provider survey than for the staff survey or the student survey: 63 providers completed the survey at Time 1, 55 at Time 2, 45 at Time 3, and 16 at Time 4. Given these relatively small samples, we analyzed the provider survey results for all providers together, irrespective of their school's treatment group. Thus, we present provider survey results for descriptive purposes, to convey provider perspectives on constructs potentially pertinent to implementation of school mental health services and student supports. We do not use the provider survey results to help account for either aggregate service levels or safety outcomes.

We used the EBPAS (Aarons et al., 2010) to measure provider attitudes in four areas: (1) intuitive appeal of EBPs; (2) the likelihood of adopting EBPs if adoption is a requirement, (3) openness to new practices, and (4) perceived divergence of EBPs from the usual practice. **Exhibit 3-44** summarizes the results of each EBPAS subscale using box plots. For each subscale, the middle horizontal line of the box plot displays the mean score. The part from the top horizontal line (the maximum) to the top of the box is the highest (fourth) quartile.

The values from the top of the box to the middle line (mean score) are in the third quartile. The values from the middle line to the bottom of the box represent the second quartile, and the values from the bottom of the box to the bottom horizontal line (the minimum) reflect the first quartile. Dots reflect outliers in the data.

Exhibit 3.40   Subscale Scores From the Evidence-Based Practice Attitude Scale
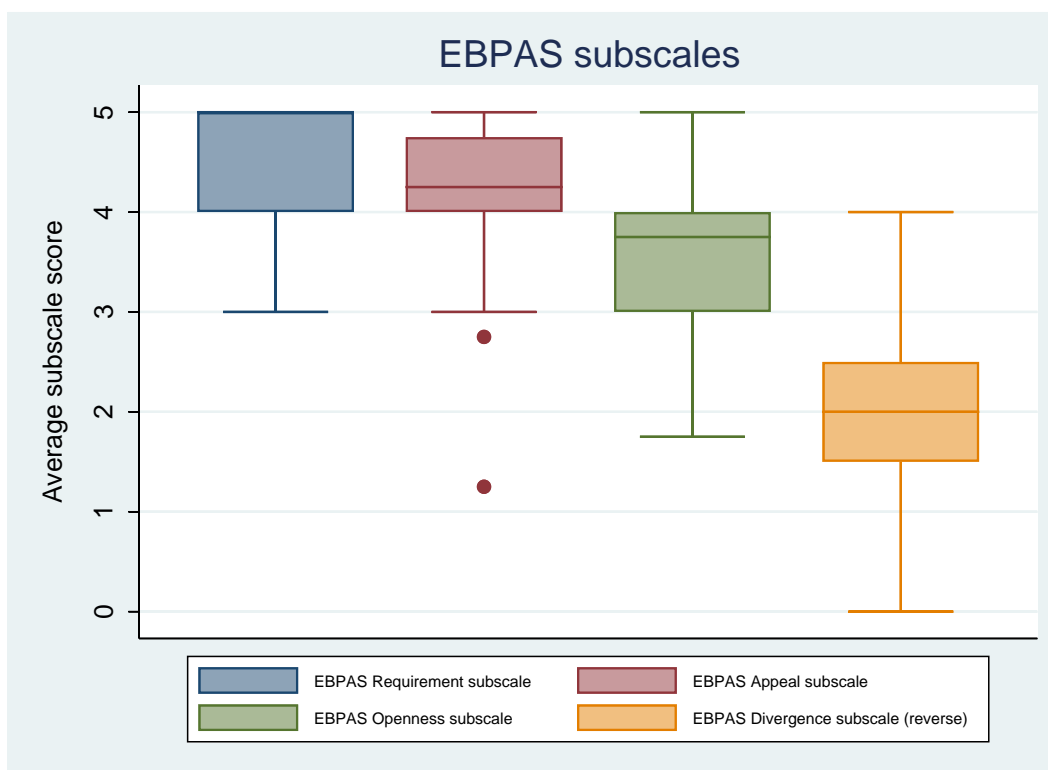
Providers scored highest on the Requirement subscale, indicating that they would likely adopt EBPs if they were required. They also scored fairly high on the Appeal subscale, indicating that they would likely adopt EBPs if they were personally intuitive and interesting. In contrast, providers reported hesitance to adopt EBPs if they are very different from typical therapies and interventions.

To measure clinicians' self-efficacy in suicide assessment and intervention, the survey included two of the four original subscales of the Counselor Suicide Assessment Efficacy Survey (Douglas & Morris, 2017). As shown in *Exhibit 3-45*, on average providers felt confident in their abilities to effectively

address issues of abuse and thoughts of self-harm and to take appropriate and immediate action in the event of imminent risk for suicide.

Exhibit 3.41   Subscale Scores on Suicide Clinical Self-Efficacy
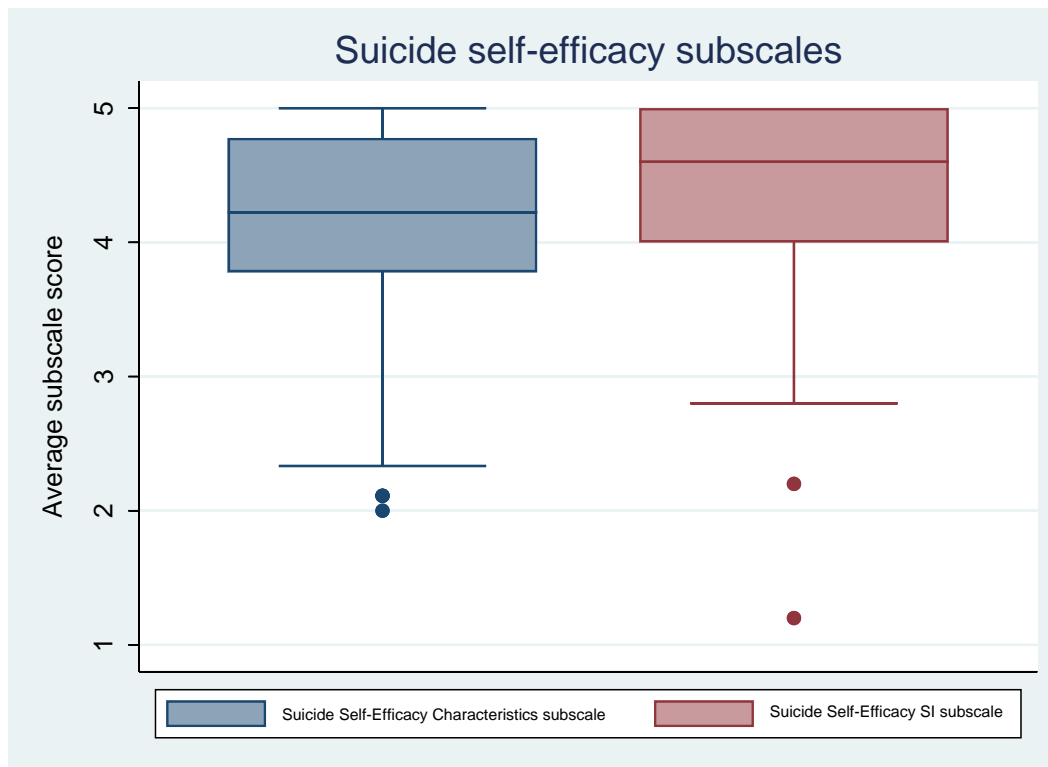
Provider perceptions of their schools' organizational readiness for new or adapted therapies were assessed in four areas: (1) organizational capacity, (2) organizational culture/climate, (3) staff capacity, and (4) implementation plan. As shown in ***Exhibit 3-46***, scores for the organizational subscales were in the middle range of the scales, suggesting that providers viewed their schools as not quite ready to take on or sustain EBPs.

Exhibit 3.42   Subscale Scores on Organizational Readiness
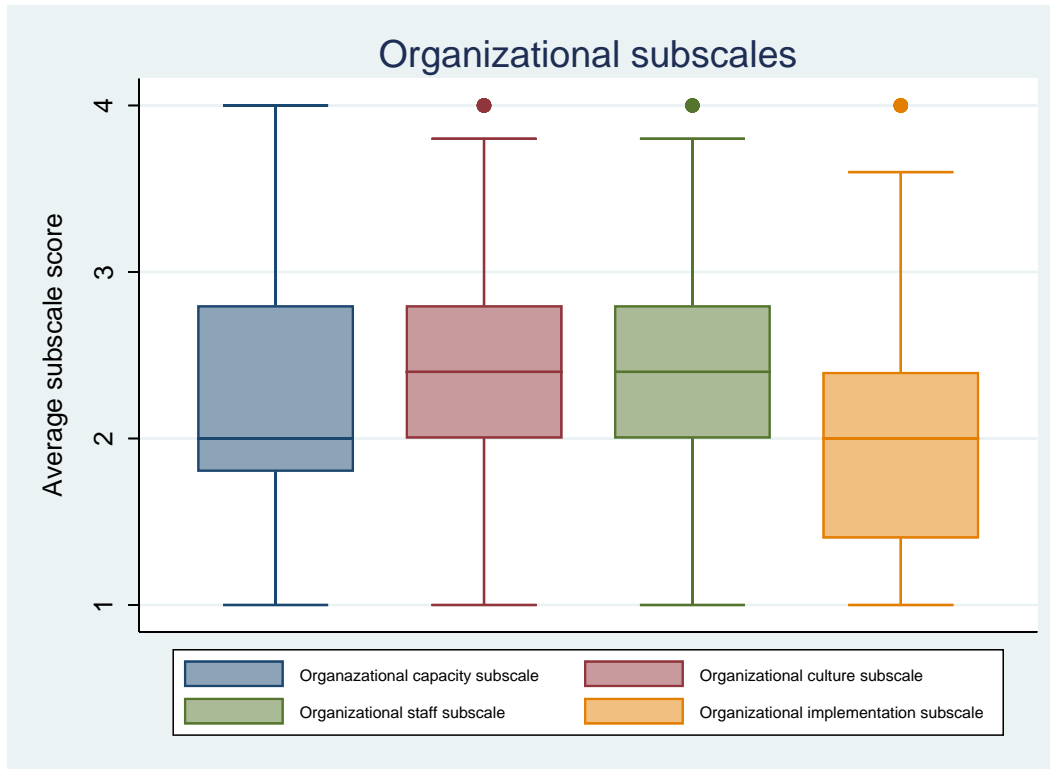
We examined the pairwise correlations between all subscales. Unsurprisingly, subscales within the same construct were generally correlated with each other. Moreover, the EBPAS Openness subscale was moderately positively correlated with all Organizational subscales (correlation coefficients from 0.27 to 0.40), suggesting that providers in more organized environments were more open to innovative practices. In addition, the Suicide Self-Efficacy Characteristic subscale was also modestly positively correlated with three Organizational subscales (Culture, r = 0.26; Staff, r = 0.24; Capacity, r = 0.15), suggesting that providers in more organized environments felt more confident in their ability to address suicidality. The complete correlation matrix is provided in ***Appendix B***.

3.7   THEMES FROM PROVIDER INTERVIEWS

In this section we present themes identified through qualitative analysis of interviews with 22 providers from 14 schools (12 TAU, Expanded, and SBMH-ET schools and two non-SBMH

schools). The interviews addressed provider perceptions of implementation, the impact of the implementation, and facilitators of and barriers to implementation of SBMH in general and SPARCS and DBT in particular. We discuss themes within the broad domains of collaboration, capacity, provider burden, student need, impact, and barriers and facilitators. As discussed in a subsequent section, we also used these provider interview themes in mixed-methods analysis of service levels.

> *Absolutely, a huge increase in capacity… now we have a therapist here five days a week, but we have two days a week where we have two therapists. So that's a huge change. It's a great sign that the kids are getting— the kids are being identified for the services they need.—Counselor, Expanded school*

3.7.1 Collaboration

Theme 1. Provider collaboration increased and relationships improved at TAU and Expanded schools since the beginning of the grant, as increased SBMH services appeared to facilitate collaboration between school providers and community therapists.

Theme 2. A counselor from a non-SBMH school stated that the psychologist on staff was only at the school "one day a week, and I never see her."

> *One real benefit of doing the school-based is that a lot of times I can work with the school counselor and it's a more—it sort of we can coordinate care or exchange observations or let each other know if there's anything the other needs to know. That can really work out well for helping the students and the families is that relationship. So that's a positive.— Therapist, TAU school*

Theme 3. A psychologist from a TAU school noted that collaboration can depend on the individual provider, citing that the original therapist was very collaborative while the new therapist was not. A respondent from one Enhanced school stated that collaboration became more difficult in the later years of the grant once psychologists were consolidated to work across multiple schools.

3.7.2 Capacity

Theme 4. Respondents from several Expanded and Enhanced schools reported an increase in their schools' capacity to

provide mental health services to students since the grant began. Having additional providers helped ease the workload, and respondents appreciated being able to use their professional knowledge and training in a purposeful way to provide services rather than spending a lot of time doing paperwork.

Theme 5. Similarly, respondents from Expanded and Enhanced schools described the positive role that student services facilitators and additional staffing played, particularly in making it possible for them to serve more students. One of the Expanded schools planned to have full-time providers in the future rather than sharing providers across multiple schools. Respondents from non-SBMH and TAU schools expressed a need for increasing the scope and availability of all student support services through facilitators and additional staffing.

Theme 6. Respondents from a few Enhanced schools and a TAU school expressed concern, however, as to whether students were getting the help they needed from individual therapists who had less expertise, were less committed, and/or did not spend adequate time with students (e.g., some providers were just "checking in" quickly with the student, and others were not in the building enough to really get to know the students and the culture or be a part of the school). A respondent from an Enhanced school also wondered whether SBMH students received the same amount of thorough counseling as they would in an outpatient setting.

3.7.3 Provider Burden

Theme 7. Despite increases in capacity in Expanded and Enhanced schools, respondents from nearly all schools (across all study groups and both provider interview time points), cited that student support staff were overburdened and only able to provide limited, reactive mental health

> *It can't be a layer on top of 12 other layers. You know, and I think what ends up happening is the excitement that we might feel can easily get dampened when we realize, "Wow, this is amazing, and they've not taken anything off my plate. It's just another thing I have to do," and that's very unfortunate… and [the district] tried to address that, but again, you know, as much as we want it not to be another layer on top of multiple other layers, that's kind of what it is, because structurally, things are going to have to change for it to be different.—Social Worker, Enhanced school*

services for students. Referred students were on waitlists in Expanded schools because providers were overwhelmed. Providers from Enhanced schools said they were only able to provide one-time or temporary support for students experiencing behavioral/mental health problems and that more regular follow-up with students was needed. Providers from Enhanced schools also said implementing DBT and SPARCS placed an additional time burden on staff. Across most interviews, respondents perceived that student support staff were spread thin and unable to address their service needs fully with the available staffing.

3.7.4   Student Need

Theme 8. Provider burden was likely exacerbated by high levels of student need for mental health services cited by several respondents from Enhanced, Expanded, and TAU schools, which respondents said increased since the grant began. With 80% to 85% of the student population in need of mental health services at some schools, respondents said their schools did not have the capacity needed to address the needs of all students. Both respondents from non-SBMH schools also reported high levels of student need that increased throughout of the grant.

> *The problem in schools like [ours] is that that triangle is inverted…We should not have 80 or 85% of our kids at the base of that triangle, but the fact is we do. So then it becomes impossible to be proactive and intentional about how you're meeting the needs of kids—it's just overwhelming, and the way that things are structured, the way that systems are put in place to address the needs of kids, can't accommodate the reality on the ground.—Social Worker, Enhanced school*

Theme 9. Some respondents believed the higher levels of student need were due to the schools' increased focus on mental health issues, rather than increased prevalence—that is, increases in awareness and follow-up by staff led to increases in the number of student assessments, combined with people now viewing minor attention-getting behaviors as reportable.

3.7.5   Impact

Theme 10. Respondents from nearly all Enhanced schools and two TAU schools reported seeing a positive impact of SBMH on students. Students who participated in Enhanced services demonstrated enhanced interpersonal skills, improved well-being according to pre/post surveys, movement toward goals, reduced impulsivity, increased emotional control and mindfulness, improved academic outcomes, better coping skills, and reduced behavioral/counseling service needs. Providers from TAU schools also mentioned seeing students make progress toward their goals and fewer students needing mental health services overall.

Theme 11. Providers at all six Enhanced schools were particularly pleased with SPARCS because, as one respondent noted, SPARCS was less intense and less difficult to implement

than DBT. Another respondent said the skills groups worked particularly well for students who were less likely to open up to a therapist in DBT. Respondents said SPARCS was effective and that students found the activities engaging and looked forward to the sessions.

Theme 12. Respondents from the other four Enhanced schools and one Expanded school indicated limited impact on students, citing challenges getting DBT groups started or indicating that students with higher needs do not benefit as much because they need more outside counseling or other services. One respondent expressed concern about the small percentage of the total school population impacted by SPARCS/DBT relative to the need in the school: "I just worry that researchers are going to look at this and say, 'We only made a small improvement, so it wasn't really worth it.' Yes, it does make a difference for that population of kids, but also we need more help and I don't want it to not be effective and then they say, 'We're going to stop doing it.'"

Theme 13. Respondents from some SBMH schools (Enhanced, Expanded, and TAU) expressed concerns about whether SBMH services were preparing students for the real world (e.g., ignoring a bully will not work in a community where standing up for yourself is critical to ensuring one's safety; what might be acceptable behavior at school may not be acceptable at home or in the community; the peer environment does not support positive relationship-building skills).

3.7.6 Barriers and Facilitators

Theme 14. Respondents in several schools (Enhanced, TAU, and non-SBMH) cited challenges getting buy-in from stakeholders; five of six Enhanced schools faced these challenges at multiple levels (from the district, principals, providers, and parents). Three Enhanced schools noted a lack of support from administrators, citing a general lack of support from principals, a lack of needed supplies from the district, and a concern that services would be discontinued after the grant ends. Respondents described both

> *A lot of kids I have, the parents might be working, and they go home and then they're on—they're on technology, and it's really hard to help them change those behaviors when there aren't adults around to help them.—Therapist, TAU school*

cooperation and resistance by families regarding their children's participation in SBMH. On one hand, families were eager to receive needed services for their child; on the other hand, pervasive family economic/social challenges limited their participation. Enhanced schools also faced challenges with mental health providers. In general, some participants expressed subtle or outspoken resistance to implementation. Some implementers believed that the Enhanced therapies were inappropriate (based on their theoretical/philosophical underpinning—i.e., "behavior modification," whether its content is applicable to middle school and younger students, that "therapy" should not be provided by school counselors, or Enhanced therapies were redundant to therapy techniques already available in schools). The reluctance of school and external agency providers to support DBT/SPARCS implementation ("lack of buy-in") was identified as a cause for incomplete implementation. In contrast, the only two Expanded schools that mentioned buy-in described positive support from administration.

Theme 15. SBMH schools (all six Enhanced schools, one Expanded school, and one TAU school) described how schools prioritize academics over mental health services; for several schools, this meant that scheduling a time to meet with students around academic instructional time during the school day was a challenge. Several Enhanced schools found ways to overcome this challenge, such as holding groups during electives instead of core classes, overlapping two blocks instead of one so students only miss class twice per month instead of four times per month, and rotating the schedule so that students do not miss the same class each week. Respondents from a few Enhanced schools asserted that these challenges are precisely why SBMH services are needed—to reduce the amount of class time students miss and to help students move forward academically; as one respondent said, "Those are some of the basic areas that are in front of student learning."

Theme 16. Some providers described an unclear boundary between imposing consequences for inappropriate behavior and meeting students' needs for emotional and behavioral counseling. Respondents from Enhanced, TAU, and non-SBMH schools said that schools often employ reactive, punitive disciplinary actions, whereas mental health providers support the use of proactive, positive reinforcement, incentives, and

3-51

morale-boosting techniques. On the other hand, respondents from a TAU and a non-SBMH school described students being "enabled" by staff via support, with discipline not always being harsh enough to hold students accountable for their actions.

Theme 17. Providers who discussed record keeping (from Enhanced, TAU, and non-SBMH schools) described maintaining limited documentation of the students they saw, such as calendar notations of how often and how long they met with students and what they were working on or diagnoses. Providers were expected to figure out their own documentation method, which could take years of trying different things. Data on student progress consisted of anecdotal notes, academic achievement, and pre-/post-intervention surveys/ psychological rating scales (but no one mentioned relying on all three of these).

Theme 18. Respondents from several Enhanced schools and one Expanded and TAU school mentioned the need for a whole school component to, for example, back up providers' messages to students and educate teachers on mental health issues—perhaps through turning SPARCS into a schoolwide guidance curriculum .

Theme 19. Several respondents from Enhanced and TAU schools said that not all schools had adequate, dedicated space where providers could see students. At times, this created potential lapses in confidentiality and conditions where students might be hesitant to share personal stories.

Theme 20. Many Enhanced schools and one Expanded school described ways in which the SBMH process can be burdensome for providers and teachers (e.g., approvals to provide services can take a long time to obtain, which leaves students at risk; it places an additional burden on already overwhelmed teachers, scheduling large SBMH team meetings can be difficult).

Theme 21. The level of implementation of Enhanced and Expanded was frequently described as limited in extent (time of initiation, number of participants, and duration/frequency of sessions) and not strictly adhering to the parameters defined in training.

3.8 MIXED METHODS

We used the mixed-methods process model (Exhibit 2.18) to assist in answering EQ-3. What were the barriers and supports to implementation of different SBMH services? The sequential explanatory mixed-methods design used quantitative provider log data to establish levels of implementation at each school. Qualitative data from provider interviews were matched with their school's level of implementation - high, moderate, or low implementation.

3.8.1 High and Low Levels of Implementation Between Schools

To identify high, moderate, and low levels of implementation, each measure was combined across all 3 school years (2016–17, 2017–18, and 2018–19) and divided into thirds using benchmark values, as shown in *Exhibit 3.47*. Schools were assigned an implementation level and score for each measure based on the average time rate and student rate over the study period. High implementation levels received a score of 3, moderate levels a score of 2, and low levels a score of 1. A final score was created by averaging scores for psychologist and therapist measures. Counselor rates were excluded because of high missingness. For non-SBMH schools, which did not have therapists, we averaged psychologist rates only. Across all schools, 26% were high implementers, 33% were moderate implementers, and 41% were low implementers. Among SBMH schools, 30% were high implementers and 33% were low implementers.

Exhibit 3.43   Scoring Implementation Measures

Implementation Measure	High	Moderate	Low
Psychologist, percentage of students	15% or more	5%–14%	Less than 5%
Psychologist, minutes per student	50 or more	20–49	Less than 20
Therapist, percentage of students	5% or more	3%–4%	Less than 3%
Therapist, minutes per student	30 or more	14–29	Less than 14
Score	3	2	1

To better understand the context and fidelity of SBMH implementation, we examined themes from the provider interviews in combination with school-level program

implementation as categorized in Exhibit 3.47. In ***Exhibit 3.48,*** we display themes form provider interview in the rows, with each column representing a different respondent. Across the 12 respondents from SBMH schools, five were from schools with high implementation, two were from schools with moderate implementation, and five were from schools with low implementation. In Exhibit 3.48, plus signs (+) denote that a theme was mentioned as a facilitator to implementation, whereas minus signs (–) denote that the theme was described as a challenge to implementation by the respondent.

This analysis revealed that providers from schools with high and moderate implementation faced just as many challenges as those at low implementation schools, and low implementation schools noted just as many facilitators as high and moderate implementation schools. More high than low implementation schools' respondents mentioned increases in capacity to provide mental health services, while low implementation schools described more challenges with space issues. Otherwise, differences in facilitators and challenges faced were minimal across implementation levels.

Exhibit 3.44   Joint Display of Provider Experiences Themes by School SBMH Implementation Level

Theme Category	Theme Code	Theme	High					Moderate		Low				
			EN School 4	EN School 31	EX School 33	TAU School 3	TAU School 19	EN School 5	EN School 20	EN School 14	EN School 26	EX School 16	TAU School 21	TAU School 30
Collaboration	1	Provider collaboration increased			+							+	+	+
	2	Provider collaboration (lack of)								−				
	3	Provider collaboration (mixed)							−/+				−/+	
Capacity	4	Increase in capacity to provide mental health services	+	+	+			+	+			+		
	5	Additional support staff provided through the grant are needed/helpful	+	+	+	+				+		+	+	
	6	Concern about whether students are getting the help they need	−			−				−	−			
Provider Burden	7	Providers are overburdened and provide limited, reactive services. Comments noted the time burden required for DBT/SPARCS training on top of existing responsibilities.	−	−	−	−		−	−	−	−	−	−	
Student Need	8	High level of student need	−	−	−	−		−					−	−
	9	Artificially high prevalence		−					−				−	

(continued)

Exhibit 3.48   Joint Display of Provider Experiences Themes by School SBMH Implementation Level (continued)

Theme Category	Theme Code	Theme	High					Moderate		Low				
			EN School 4	EN School 31	EX School 33	TAU School 3	TAU School 19	EN School 5	EN School 20	EN School 14	EN School 26	EX School 16	TAU School 21	TAU School 30
Impact	10	Positive impact of SBMH in students	+	+					+	+	+		+	+
	11	Positive impact of SPARCS/DBT	+	+				+	+	+	+			
	12	Lack of impact	−	−	−				−	−	−			
	13	Concern over real-world application of SBMH	−		−	−		−						
Other Barriers and Supports	14	Buy-in (challenges or support)	−	−	+			+	−	−	−	+	+	−
	15	Academics are prioritized over mental health services	−	−	−			−	−	−	−			−
	16	Schools need to use more positive discipline				−		−	−	−				
	17	Documentation of services is limited	−			−	−						−	
	18	Need more resources	−		−			−	−	−				−
	19	Space issues	−	+		+				−			−	−
	20	SBMH process can be burdensome	−	−				−				−		
	21	Fidelity, adaptations made/are needed	−	−	−			−		−	−	−		

Note: +, −, −/+ indicate theme was described as a facilitator (+ positive), challenge (− negative), or mixed facilitator/challenge (−/+). Blank cells indicate the theme was not mentioned by the respondent. DBT, Dialectical Behavioral Therapy; EN, Enhanced; EX, Expanded, SBMH, school-based mental health; SPARCS, Structured Psychotherapy for Adolescents Responding to Chronic Stress; TAU, Treatment as usual.

4 Discussion

This study was ambitious in its goals and scope, with mixed success. The study was grounded in the premise that improving mental health services to students whose needs were the highest and were not previously being adequately addressed would help them to improve their behavior and—to the extent that those students contributed to a disproportionate share of behavioral infractions or violence in their schools—such improvements in their behavior would, we posited, result in substantial schoolwide improvements in safety and climate.

This premise is supported by research supporting both pillars of the premise: students with mental health challenges have been shown to contribute to a disproportionate share of behavioral infractions or violence (Sullivan et al., 2013), and evidence-based mental health services have been shown to improve student behaviors and academic performance (Powers et al., 2014; Sanchez et al., 2018). Together, those two bodies of research provide rationale for believing that improved mental health services provided to a carefully targeted group of students could improve school safety and climate for all students. However, the premise has not been widely subject to rigorous empirical test. One exception is a study by Ballard and colleagues (2014) that found that targeting selective interventions to students who already display aggressive behavior or mental health needs can reduce suspensions and expulsions for the entire student body. Our study used a similar model but expanded the scope by comparing multiple approaches' effect on multiple outcomes including overall school climate. In contrast, most studies have focused on individual outcomes of treated students or descriptive studies of

effects on school climate. If empirically supported by this study, our premise and approach would hold great promise in helping to improve school safety and climate—and students' school experiences.

The study was ambitious in scope in several regards, as follows.

Approach: The study included extensive support to improve mental health services in selected schools as well as rigorous implementation, outcome, and cost evaluation. This support was made possible through a grant from the National Institute of Justice's Comprehensive School Safety Initiative (CSSI) and a steadfast commitment from the Charlotte-Mecklenburg school district and participating schools. Without such grant support and partner commitment, the study's ambitious approach would not have been possible.

Design: The study comprised both a school-randomized controlled that included 25 middle schools, and a quasi-experimental supplemental design in which an additional nine schools that did not have school-based mental health services (and thus could not be included in the RCT) were selected that best matched the RCT schools, with propensity weighting being used to statistically adjust for differences between the non-SBMH schools and the SBMH schools. It is important to note that throughout the results, non-SBMH schools differed from SBMH schools at baseline and, sometimes, at follow-up, which is due to the fact that they were inherently lower-need schools given that their need had not risen to the level of receiving SBMH services at the outset of the study. In keeping with this attention to school needs, CMS introduced SBMH into two schools that had not had them, because of their emergent needs for services. For the study, this meant that two non-SBMH moved to the TAU condition; we largely excluded these two schools from the outcome analyses. Through this innovative and rigorous design, the study accommodated the existing situation regarding school mental health services in CMS middle schools and generated a wealth of rigorous information.

Data sources: The study included multiple data sources measuring outcomes (staff and student surveys, administrative data on infractions), implementation (provider log data on students served and quantity of service; qualitative interviews

and surveys with providers) and cost (start-up and ongoing), as well as qualitative interviews with service providers. These complementary data sources allowed the study to a) describe implementation of the interventions as well as the associated costs, b) measure changes in school safety and climate over time and from multiple perspectives, c) use the implementation measures to explain and account for changes in outcomes, d) estimate costs relative to improvements in key outcomes, and e) use qualitative findings and themes to better understand and interpret the quantitative results.

Duration: The four-year CSSI grant with an additional no-cost extension year provided support over a sufficient length of time to support a) a partial year of preparation (Spring 2016in which arrangements were put in place with study schools, and data collection instruments and procedures were developed; b) 3 years of data collection, including four waves of staff and student surveys (Times 1–4), service logs kept by providers over 3 school years (2016–17 through 2018–19), and 4 years of administrative data on infractions during the year prior to service delivery and the 3 years of service delivery (2015–16 through 2018–19).

4.1 SUMMARY AND DISCUSSION OF FINDINGS

4.1.1 Service Levels

As described in the Results section, service levels varied significantly based on school and time point, but comparisons between treatment conditions indicated that they did not appear to systematically vary between conditions. Our interventions were designed to increase services differently in each treatment condition across each of the three providers— therapists, psychologists, and counselors (see Exhibit 2.2). For therapists, additional SBMH payments for therapists' services were added to each of the three SBMH treatment conditions. Despite this, therapist logs indicate that between 10 and 50 students were seen per year for therapy depending on the school. This is not substantially different from the rate of 20 to 40 students who were seen for therapy in each school prior to the study.

Additional psychologists were hired to increase the amount of time psychologists spent at Enhanced and Expanded schools. Analyses showed that, for the most part, there were no

statistically significant differences between treatment conditions in the percentage of students seen by psychologists or the number of minutes psychologists spent serving students. One exception was a finding that approached significance in the last year of the study (2018–2019) that showed Enhanced and TAU schools had marginally higher percentages of students seen by a psychologist than Expanded schools. The reasons for the relative dip percentage of students seen by a psychologist in that year in the Expanded treatment condition are unclear, but qualitative data discussed later suggests service provision levels may be driven by school need, rather than by assigned treatment condition.

A student services facilitator was assigned in each Expanded and Enhanced school to complete some of the counselor's administrative duties and allow that counselor more time to provide direct services to students. No statistically significant differences were found between counselors' percentage of students serviced at Expanded, Enhanced, or TAU schools at either of the two time points that data were collected from counselors.

With a lack of systematic differences in service levels between treatment groups, we explored service levels at the school level to examine within treatment group variability. In reviewing plots in Exhibits 3.1–3.5, we see that schools within treatment groups did vary in their levels of service within each time point and across time. This suggests that service level differences were more pronounced between schools within each treatment condition than between treatment conditions.

The service level variability within treatment condition might be explained by many factors related to how providers work within different schools. For example, some data from the provider survey pointed to the relationship between organizational readiness for EBP and providers' use of EBPs. This relationship may affect service levels if certain schools' leadership and overall climate supports higher service levels, while others do not support increased service levels. We also know that changing circumstances from year to year within a school may significantly affect service levels, but these circumstances are hard to measure. Pertinent circumstances subject to unmeasured changes across years may include staff turnover, student cohort differences, new administrators, and changing

community contexts, each of which may affect service needs and the practices of service providers within a school over time.

To further understand the effect of different levels and types of SBMH, we first tested treatment group differences in outcomes each year and assessed patterns in significant differences across the years. Then, in acknowledgement that service levels did not align with treatment group as we had expected, we tested how service levels were associated with outcomes, irrespective of treatment condition.

4.1.2 Changes in Outcomes as a Function of Treatment Group

The study examined school safety outcomes as reflected in staff and student personal experiences, feelings of safety, and perceptions of the school. We also examined administrative data on student infractions, cognizant that those data likely reflect not only actual student behaviors but also differences in reporting (e.g. which behaviors are seen as rising to the level of an infraction that should be reported). As discussed below, many school safety indicators were worse in SBMH schools than non-SBMH schools, which is not surprising given that SBMH services are purposively placed in schools in which more students have higher needs, which are often associated with higher levels of misbehavior. Of greater interest and pertinence to this study, some patterns of outcomes showed relative improvement in the SBMH groups compared with the non-SBMH groups and, to a lesser degree, in the Enhanced group compared with the other SBMH groups. In both sets of comparisons, the patterns were seen in only some outcomes and at some time points—that is, they were not pervasive or consistent.

Staff Survey Outcomes. One important measure of school safety is the extent to which staff and students feel unsafe. The staff survey included a set of items that asked how often during the past 30 days the respondent felt unsafe before, during, and after school hours in each of three locations: in their classroom or office, in other areas of the school building, and on school grounds. For all nine combinations of time frame and location across four survey time points and four treatment groups (144 combinations total), most respondents felt safe. For 49 of the 144 time/location combinations (34%), less than 10% of staff reported ever feeling unsafe; 10% to 20% of staff reported ever feeling unsafe in 86 of the combinations (60%). Thus, for

4-5

135 of the 144 combinations (94%), 20% or fewer staff reported ever feeling unsafe. We found few statistically significant differences in the percentage of respondents in each treatment group who reported feeling unsafe at each survey time point for each time frame.

Similarly, staff were unlikely to report experiencing safety problems such as being hit or assaulted by students and there were very few differences between the treatment conditions.

Staff were more likely to report having observed student behaviors such as bullying or fighting in their designated work area (e.g., classroom) in the past 30 days. The percentage of staff who reported observing student bullying ranged from 58% to 77%; at most time points, more than half the staff in each treatment group reported student fighting. It is not surprising that staff were more likely to report observing student misbehavior than personally experiencing it, given that each staff member could observe numerous students and that students are more likely to misbehave toward each other than toward staff.

Trends (i.e. changes over the four survey time points) in the percentage of staff who reported student bullying and fighting were more favorable for the Enhanced group than for the Expanded or TAU groups, although the effect for fighting was no longer seen at Time 4. Similarly, Enhanced staff ratings of student disruptive behavior showed improvements relative to other groups between Time 1 and Time 3, but at Time 4 the Enhanced group was worse than the other groups. These patterns showing better outcome trends for the Enhanced group through Time 3 support the central premise of the study, but that effect did not last through Time 4. Perhaps this is not surprising, in light of the finding discussed earlier that service levels were not consistently in line with the levels that were expected based on treatment group. Further below, we discuss associations between service levels and outcomes, irrespective of treatment group.

Student Survey Outcomes. Students reported low rates of aggressive behavior perpetration or victimization: mean responses were equivalent to slightly less than one time in the past 7 days for students in schools in the three SBMH groups and between zero times and one time for students in the non-

SBMH schools. The non-SBMH group means for aggression perpetration were consistently lower than the means for the SBMH groups. This difference is not surprising because the SBMH and non-SBMH schools differed substantially prior to the study: SBMH services had been dedicated to schools based on need, which included the prevalence of student behavioral problems. The SBMH groups usually did not differ from each other in their trends in aggression or victimization scores. On victimization, there were some patterns in differences with the non-SBMH groups that indicated relative improvements for the Expanded and TAU groups, but the differences were inconsistent across the survey time points and did not suggest compelling or lasting improvements relative to the non-SBMH group.

Student ratings of the frequency of negative interactions with peers followed a pattern similar to that seen for victimization. Ratings of negative peer interactions were substantially better in non-SBMH schools at Time 1 than in the SBMH groups, but the non-SBMH ratings worsened at Time 2 and Time 3, suggesting some relative improvement in the SBMH groups, compared with the non-SBMH group. However, the patterns were inconsistent across groups and did not persist at Time 4— and therefore did not suggest compelling or lasting improvements in the SBMH groups relative to the non-SBMH group.

The percentage of students who reported feeling unsafe in any of the three locations ranged from 30% to 50%. In the SBMH groups, the percentages were quite similar for before school (42% to 50%) and during school (41% to 49%) and somewhat lower for after school (37% to 42%). Percentages were uniformly lower for the non-SBMH group than the SBMH groups. In the non-SBMH group, percentages were slightly higher for during school (35% to 42%) than before school (30% to 39%) and after school (30% to 34%). The mean percentages of students who felt unsafe remained fairly consistent over the survey time points for the SBMH groups, whereas the mean percentages for the non-SBMH group increased, suggesting relative improvement for SBMH groups. This pattern was seen for the "during school" and "after school" timeframes.

4-7

Students were asked the extent to which each of 10 types of problems (e.g. student fighting) was a problem in their school. At all four survey time points, the non-SBMH group mean was significantly lower than each of the other group means, and the latter three groups were not significantly different from each other. This pattern is similar to patterns seen in other outcomes: compared to the SBMH groups, the non-SBMH group had lower levels of problems at Time 1 but worsened over time, whereas the SBMH groups remained relatively consistent, suggesting improvement relative to the non-SBMH group. However, for this school safety problems outcome, the SBMH mean ratings remained significantly worse than the non-SBMH mean ratings.

Disciplinary Infractions Outcomes. We used student disciplinary data provided by CMS to examine differences among the treatment groups during each of 4 school years in the following infraction categories: disruption; insubordination or disrespect; harassment, threats, or bullying; aggressive behavior; fighting or assault; and weapons or gangs. For each category, we examined 1) the percentage of students in grades 6 through 8 who received one or more disciplinary infraction reports, labeled "binary"; and 2) for students who received one or more infraction reports in a category, the number of reports in that category that each student received, labeled "count."

For both the binary and the count measures, there were numerous significant differences involving the non-SBMH group. For each binary outcome, more than half of the 12 pairwise contrasts between the non-SBMH group and each of the other three groups at each of four time points showed statistically significant differences. On the binary measure, the percentage of students with one or more infractions in a category was almost always lower in the non-SBMH group than in the SBMH groups. Fewer significant differences between non-SBMH and SBMH groups were found for the count measures, though significant differences were still more common than would be expected by chance. There may have been fewer differences in part because these analyses included far fewer students than the analysis of binary outcomes: fewer than 10% of students had one or more infractions in each category and were thus included in analysis of the count measure. The finding of fewer differences in the infractions count data may also reflect the likelihood that the number of infractions per

student does not differ as greatly between different types of schools as does the percentage of students with one or more infractions.

Given the large pre-existing and ongoing differences between the non-SBMH group and the SBMH groups on the binary outcomes, we focused on differences among the Enhanced, Expanded, and TAU groups. Specifically, we examined the patterns of between-group differences over the 4 years of infraction data. For each infraction category, the main findings were:

- *Insubordination or Disrespect:* for both the binary and count measures, the pattern of trends suggests better outcomes in the Enhanced and Expanded groups relative to the TAU group.

- *Fighting*: for the binary measure, the pattern of trends suggests better outcomes in the Enhanced group relative to the Expanded and TAU groups. However, for the count measure there were better outcomes in the TAU group relative to the Enhanced group.

- *Harassment, Threats, or Bullying:* for the binary measure there were no significant differences among the SBMH at any time point. For the count measure, the Enhanced group mean was consistently higher than the means for the Expanded or TAU groups; the groups did not substantially differ in in the trend patterns.

- *Aggressive Behavior:* for the binary measure there were no significant differences among the SBMH groups at any time point. For the count measure, the trend patterns were inconsistent and did not suggest compelling differences among the groups.

- *Disruption:* there were no significant differences among the SBMH groups at any time point on either the binary measure or the count measure.

In short, there was some indication of better infractions outcomes for the Enhanced group compared with the Expanded or TAU groups, but those indications were limited and inconsistent across infraction categories.

4.1.3   Outcomes as a Function of Service Levels

The levels of student services provided by psychologists, therapists, and counselors in each school were not always aligned with the expected increases in services for a school's treatment group based on intended implementation of SBMH.

Therefore, we assessed the association between key outcomes and five school-level implementation measures (i.e., the amount of services rendered by mental health professionals in the schools). These analyses examined the association between student service levels and outcomes in the same school year (which we refer to as concurrent effects) and service levels in one school year and outcomes in the following school year (which we refer to as lagged effects).

In Chapter 3, we described each of the statistically significant associations between service levels and staff survey, student survey, and student infractions outcomes. To facilitate discussion, we now present a summary of those statistically significant associations (see **Exhibit 4-1**). Associations were considered *favorable* when higher service levels were associated with better outcomes (i.e. less frequent negative events) and *unfavorable* when higher service levels were

Exhibit 4.1   Percentage of Significant Favorable and Unfavorable Associations between Service Levels and Outcomes

Outcome	Psychologist				Therapist				Counselor	
	Minutes		Students		Minutes		Students		Students	
	Con-current	Lagged	Con-current	Lagged	Con-current	Lagged	Con-current	Lagged	Con-current	Lagged
Staff Survey (for 9 outcomes)										
Favorable	11%	0	33%	0	44%	11%	55%	0	44%	33%
Unfavorable	11%	0	11%	0	0	0	0	0	0	0
Student Survey (for 7 outcomes)										
Favorable	29%	0	14%	0	0	14%	0	0	0	0
Unfavorable	0	0	14%	0	14%	0	29%	0	0	0
Infractions Data (for 12 outcomes)										
Favorable	0	0	0	0	0	0	0	0	8%	0
Unfavorable	8%	8%	8%	8%	8%	8%	0	8%	0	0

associated with worse outcomes (i.e. more frequent negative events). The analyses summarized here included nine outcomes from the staff survey, seven outcomes from the student survey and 12 infraction categories (6 binary and 6 count outcomes); those values were used as the denominators in calculating the

percentages. For example, there was one favorable association between psychologist service minutes and the staff survey outcomes: the percentage was calculated as 1/9 = 11%.

Several conclusions can be drawn from these summaries. First, for the survey outcomes, concurrent associations were much more common than lagged associations. This finding is not surprising: concurrent associations reflect more temporally proximal effects, whereas lagged associations would require that services during one school year be associated with outcomes the following school year, which is an ambitious expectation.

Second, staff survey outcomes were often favorably associated with service levels. Among concurrent associations involving staff survey outcomes, 17 of 45 associations (38%) were favorable, suggesting that better staff experiences and perceptions regarding safety and student behaviors often followed higher levels of student mental health services. The number of these favorable concurrent associations greatly exceeds what would be expected by chance: 5% of 45 possible association would be 2.25. *This predominance of favorable associations supports a central premise of this study, that improving mental health services for students in need has the potential to improve the school climate broadly.* This relationship was not as strong for lagged associations involving staff survey outcomes: only 4 of 45 lagged associations (9%) were favorable, but this is still roughly twice as many as would be expected by chance. It is also worth noting that unfavorable associations between service levels and staff survey outcomes were rare, below the level expected by chance.

Third, service levels were less often associated with student survey outcomes than was seen for staff survey outcomes. Also, unfavorable associations were approximately as common as favorable associations. One possible explanation for the difference between staff and student results may be that about half of the staff survey outcomes were based on the respondent's personal experiences or feelings and half were based on perceptions of student behaviors, whereas all but one of the student survey outcomes was based on the respondent's personal experiences or feelings. Thus, the staff survey may reflect the broader school climate to a larger degree, whereas

the student survey reflects the experiences of the students who were sampled and completed the survey.

Fourth, service levels were associated with student infractions (as reported in administrative data) at a rate just slightly greater than would be expected by chance. Among concurrent associations involving infractions data, four of 60 associations (7%) were statistically significant, of which one was favorable (2%) and three were unfavorable (5%). Among lagged associations, 4 of 60 associations (7%) were unfavorable and there were zero favorable associations. Unfavorable associations were more common than favorable ones, for concurrent but especially for lagged associations. We posit that this pattern of results is best seen as suggesting that higher levels of services are delivered in schools with more problems of the type that relate to the infractions that we examined—and that the services were not sufficient to mitigate those problems. Associations between service levels and infractions barely exceeded chance levels and we recommend not placing undue emphasis on these findings.

4.1.4 Cost and Cost-Effectiveness

The CEA analysis has considerable practical use to administrators. The results indicate the likely costs of adopting the Expanded or Enhanced approaches over and above TAU, and what that additional cost would yield in terms of reducing victimization. The estimates showed that, although both interventions are effective compared to TAU, schools considering either approach would optimally choose the Expanded approach because it is more cost-effective.

Whether a TAU school should adopt either the Enhanced or Expanded approach at all depends on the districts' willingness to pay for reducing victimization—that is, the maximum amount that a district considers would be a worth a one unit reduction in the outcome.

An illustrative hypothetical example is of a school with 1,000 students that has been awarded a $6,800 grant to improve mental health. Our findings suggest using that grant on the Expanded Approach would result in 300 fewer victimization events than under treatment as usual.

One of the limitations of the field is that there is no accepted standard for how many averted victimization events would be considered 'worth it'. (In contrast, the health care field has commonly accepted levels of what decision-makers should be willing to pay for a given improvement in a standardized measure of health.) District administrators and other decision makers must therefore determine whether the improvement in victimization outweighs the cost of the program.

The base case cost-effectiveness analysis does not include start-up costs. Nevertheless, start-up costs – those costs incurred regardless of how many students benefit from the program – would have to be budgeted for. The findings from a sensitivity analysis that allocated start-up costs across study arms did not change the primary conclusions regarding the trade-off between spending resources on the study conditions and improving outcomes.

This economic analysis faces several limitations. The small number of study schools limits the variability of costs across schools within treatment conditions and the degree to which potential confounding differences between schools could be adjusted for in analyses. As for the main outcomes analyses, these results may have limited generalizability beyond the sample of schools for the years for which data are available in the current study.

4.1.5   Process Evaluation

Provider Survey Findings. In the provider survey, service providers were asked about their attitudes toward EBP and their school environments. Given the small number of respondents to the provider survey, especially at Time 4, the provider survey results were not analyzed relative to treatment group, service levels, or outcomes, but rather were examined regarding respondents practice of SBMH and other student support services. Providers were asked about their attitudes toward EBPs, which, in our study, were DBT and SPARCS; however, the questions broadly referenced "evidence-based practices" rather than DBT and SPARCS specifically. Notably, many providers endorsed the concept of EBPs, especially when they were intuitive. Scores of their willingness to try EBPs were lower if those practices (hypothetically) diverged from typical therapies or typical practice. This distinction, while common among mental health providers, is important and indicates that

4-13

providers may have been less likely to adopt DBT and SPARCS if they were different from things they typically do. This possibility is especially relevant for counselors, who often spend little time doing structured, manualized Tier 2 groups, but typically spent more time doing administrative duties or universal Tier 1 social and emotional learning programs. This concept was echoed in some of the qualitative data presented in Section 3.

The provider survey showed several significant correlations between perceptions of organizational readiness and attitudes toward EBP. Organizational readiness scales assessed the school's culture of and capacity to adopt new EBPs. Providers varied widely in the extent that they report their school has a culture and capacity for EBPs. Further, correlational analyses show that this is significantly correlated with providers' report of their openness to try new EBPs. These attitudes around new therapies and services could play a vital role in how much and how well SBMH is implemented. Qualitative and mixed methods data further explain this connection.

Qualitative Findings. We completed in-depth interviews with a small number of providers who were willing to discuss their experiences with SBMH at their schools. They represented non-SBMH, TAU, Expanded, and Enhanced schools and included counselors, therapists, psychologists, and school social workers. Providers were asked about their perceptions of the effectiveness of SBMH interventions. Almost all providers interviewed believed that SBMH, DBT, and SPARCS are effective ways to address individual student behavioral or mental health problems. Enhanced school providers were enthusiastic about the potential of DBT/SPARCS in helping students and felt that CMS student support administrators have been effective in assisting in the implementation of these therapies. In commentary from schools where SBMH was not available (non-SBMH schools) there was a desire for this service. These findings reflect similar findings regarding SBMH from other studies (Paluta, 2015; Powers et al., 2013)

Despite these positive assessments, there was little indication from the providers that the implementation of DBT and SPARCS in Enhanced schools or SBMH in other schools had resulted directly in demonstrable improvements in school safety. Conditions related to school safety, including student

aggression, suicidal thoughts, depression, and posttraumatic stress disorder, was described as becoming more prevalent and physical aggression and other inappropriate behaviors were noted as appearing in students as early as kindergarten and first grade. This finding of unmet need has been reported in other research investigating SBMH implementation (Paluta, 2015; Powers et al., 2013).

While respondents in Enhanced schools felt that DBT and SPARCS can address the needs of students, they also expressed concern about its practicality in a school setting. They believed that the large unmet need for behavioral/mental health services overwhelmed the ability of existing staff to address them with either currently available resources or with the addition of DBT/SPARCS. The time required both for DBT and SPARCS training added to existing work responsibilities and the limited number of students who are served by DBT or SPARCS are perceived as critical barriers by many student services staff. Conditions in schools require student service providers to spend much of their time in administrative and other activities that do not provide direct support to students. The lack of time for implementation by key staff has also been found to be a significant obstacle in SBMH implementation (Paluta, 2015; Reinke et al., 2011).

Providers also described a number of important facilitators that they identified as helping their practice of student support services and helping the overall mental health and well-being of the students at their schools. In particularly, they noted that collaboration and capacity developed by expanding SBMH services to be helpful and provided important details. For example, several respondents in Expanded and Enhanced schools noted that student services facilitators were important in allowing counselors to spend more time providing direct services to students, including in the SPARCS group.

Mixed Methods Findings. We used mixed methods analysis to combine the qualitative data findings from the provider interviews with quantitative data from the service level logs. Based on provider logs across all 3 school years with study-based services, we categorized schools as high, moderate, and low implementers. Then we reviewed provider interviews in the context of the quantitative categorization of the school of that interview respondent (e.g. If Respondent 1 provided services at

4-15

school X, and school X was a moderate implementation school, we reviewed Respondent 1's interview data as data related to a moderate implementation school). In combining qualitative interview data with quantitative service level data, we identified some findings that were suggestive of differences between providers at high and low implementing schools. Whereas some providers at high implementation schools reported more service facilitators than those at low implementation schools, the most interesting result of our mixed methods analysis was that challenges tended to be very similar across high implementing schools and low implementing schools, at least according to provider response. This suggests that challenges may not fully explain why some schools were low implementers. It is important to note that our data may be subject to response bias because our respondents to interviews were most likely the most motivated providers in that school. Respondents may have been more likely to have a positive experience with SBMH then other providers.

4.2 IMPLICATIONS AND TAKEAWAY LESSONS

A key takeaway from service level findings is that schools vary significantly even when districts attempt to implement a systematic change in their levels of service. While these differences in service levels may be due to differences in school climates, school leadership, and specific school conditions, they likely primarily reflect differing levels of student need for services. Another key takeaway regarding service levels is that the number of students seen by different types of providers did not significantly change across time or across our attempts (via the treatment conditions) to increase service availability. There may be multiple reasons for that finding but our qualitative data suggests that the training and consultation requirements of our Enhanced condition took up a significant amount of time. Providers spent time learning evidence-based therapies in very intensive trainings, which may have taken away from the amount of time that they were able to provide direct services to students.

An important lesson is seen in findings that suggested that some outcomes for the SBMH groups improved compared to the non-SBMH group and, within the SBMH groups, for Enhanced and (to a lesser degree) Expanded compared to TAU. These results aligned closely with our hypotheses that schools

that incorporated Enhanced SBMH treatments would show the most school safety and school climate improvement, with Expanded SBMH service schools showing the next highest levels of improvement, and TAU showing the least of the three SBMH groups. There was also some support of our hypothesis that SBMH services in any SBMH group support more improvement over time than that in schools with no SBMH therapy services. Our results supporting these hypotheses were not consistent or totally compelling and we view them as suggesting that the approach is promising rather proven, at least within this study.

Our findings on improvements in staff experiences and perceptions regarding safety and student behaviors often followed higher levels of student mental health services, mainly within the same year (concurrent analysis) rather than the following year (lagged analysis). This relationship was not seen for the student survey outcomes, suggesting the staff survey may reflect the broader school climate to a larger degree, whereas the student survey reflects the experiences of the students who were sampled and completed the survey. This reflects the difficulty that exists in trying to change individual experiences for students at a school by changing the behaviors of a few students.

This relationship was also not seen in the student infractions outcomes; to the contrary, unfavorable associations were more common than favorable ones, for concurrent but especially for lagged associations. We posit that this pattern of results is best seen as suggesting that higher levels of services are delivered in schools with more problems of the type that relate to the infractions that we examined—and that the services were not sufficient to mitigate those problems. Associations between service levels and infractions barely exceeded chance levels and we recommend not placing undue emphasis on these findings. Higher levels of services may be delivered in schools where problems that are serious enough to require administrative discipline, which was best reflected in infraction data, as opposed to more commonplace classroom behaviors, which were best reflected in the staff survey data. There could also be instances in which overall climate and perception of behavior improved by treatment condition according to staff, but students continue to experience covert or less obvious victimization such as bullying, which is reflected in their student survey responses.

Study results indicate that the Enhanced and Expanded interventions can improve aggression and victimization outcomes at increased costs. Further research is needed on administrators' willingness to pay more for services in hopes of reducing student aggression and victimization. Future cost-effectiveness studies with a larger sample size can help administrators better understand the anticipated costs of school-based mental health interventions.

4.3 CHALLENGES AND LIMITATIONS

4.3.1 Implementation

Understanding some findings of our study may be bolstered by consideration of limitations and challenges related to implementation of services in the treatment conditions. First, providers reported both during the training and in the provider interviews that trainings and ongoing consultation for DBT and SPARCS were both intensive and time-consuming. Trainings and implementation of evidence-based treatments may have taken providers away from their daily duties or direct services typical in Expanded or TAU schools, though Enhanced schools did not show *less* service provision based on provider logs. The fact that service levels in Enhanced schools stayed more or less the same as their Expanded and TAU counterparts, while adding EBPs, could account for some of the positive findings in safety and climate outcomes in Enhanced schools.

Staff turnover was another challenge for implementation. For example, a number of providers who were trained in DBT and SPARCS in 2016–2017 did not return for the 2017–2018 school year, leading some schools to have fewer SPARCS trained staff than in the first year. For the DBT-trained staff, the model shifted from school-based psychologist and therapist completing groups to two DBT-trained psychologists providing DBT services to all of the eight Enhanced schools in 2017–2018 and 2018–2019. The hours those psychologists spent in each school was added to the overall psychologist service level totals for those schools. These challenges may have resulted in less buy-in within each school in the SPARCS and DBT models.

As with many school districts, CMS underwent redistricting before the 2018–2019 school year affecting that year of implementation and Time 4 outcome measurement. This resulted in some communities and students moving to new

schools which likely made lagged analyses less valid at Time 4 because of the new makeup of the student body in some schools. In addition, four schools ceased to exist altogether, including one school in the Enhanced condition, two schools in the Expanded condition, and one school that had already been dropped form the sample due to changing conditions. Another school had also been dropped from the sample due to changing from non-SBMH to TAU. Moreover, these four schools described above that closed were K–8 schools.

From the beginning, we had been unsure of how combining K–8 schools in our sample along with middle schools could affect results. We ensured that the randomized SBMH conditions were balanced in the number of K–8 vs. middle schools they had. The K–8 schools had been created by CMS several years prior to the study beginning to address failing schools and promote improvement. They tended to be much smaller in enrollment but have students with more needs. After incorporating the type of school (K–8 vs. middle) as a variable in some of our outcome analyses and our implementation measures, we determined that the difference in both outcomes and service levels were not statistically different between the two types of schools and did not appear to affect our analyses. However, there may have been qualitative differences between the two types of schools that we were not able to measure.

4.3.2   Implementation Measurement Challenges

A significant challenge to our study was that we were not adequately able to capture adherence to the evidence-based treatment models of DBT and SPARCS. Initial fidelity measures during training tracked providers until they showed sufficient fidelity to the model, but adherence was difficult to measure thereafter. While we attempted to measure these fidelities, providers were not required by the district to complete them and therefore exhibited did not regularly complete fidelity measures. Fidelity data could have explained our SBMH implementation data in a way that service logs and interviews could not, especially in its utility identifying between-school differences in implementation within the Enhanced treatment condition. Fidelity data could have helped us describe which schools were performing less effective implementation.

Similarly, tracking counselor data throughout the second and third year of implementation in the study was challenging.

Counselors completed logs on a regular basis during the first implementation year (2016–2017) but completed fewer during the following year (2017–2018). An initiative to establish a district-wide data collection system for counselors was not successful and counselors did not complete any service logs for the final year (2018–2019). It is possible that these data may have helped describe important differences between Enhanced and Expanded schools, which had student services facilitators (SSFs) designed to free up counselor time for direct services. SSF tracking would have improved our ability to describe between-school differences in these treatment conditions, as well. The SSFs worked full time between two to three schools evenly across the Enhanced and Expanded conditions, but interviews indicated that SSFs were helpful for some providers but might not have been for others. More information could have assisted the research team in disentangling these differences.

To obtain more information about implementation, we turned to qualitative provider interviews, but the number of respondents was small. There are several limitations to the generalizability of the provider interview data that should be considered when evaluating the findings of the qualitative data. The voluntary nature of the interview sample restricts input to only those providers who chose to participate in the research. The individuals who did not chose to participate in the interviews may have different perspectives regarding SBMH and DBT or SPARCS than those who did participate. While the participants were willing to speak confidentially with the researcher over the telephone, there was no opportunity to conduct interviews in person within the context of the school environment, or as a participant observer. These limitations inhibited the possibility of the observer establishing a rapport with respondents that may have allowed for more sensitive conversations.

4.3.3 Outcome Measurement Challenges

The primary challenge in our design and outcome measurement was that we used a repeated cross-sectional design to measure staff and student outcomes at the school level. This differed from a longitudinal design in that individual students and staff were not tracked over time to see individual changes in outcomes. Students are often not tracked over time because it can be very challenging to follow students, obtain parental

consent to follow students, and obtain a large enough sample size to adequately power school-level models (Trudeau, 2004; Trudeau, Williams & Murray, 2010). This design of repeated cross-sectional measurement is therefore common among school-level studies. However, it does not account for the fact that some of the same students and staff are likely surveyed across time points, making an unknown portion of samples include the same individuals across time. The problem with this is that lagged models do not account for this repeated measures variability and instead treat it as unexplained variability, leading to larger standard errors and lower statistical power (Kwok, West & Green, 2007).

The inability to track individual students or staff over time also means that an essential piece of information about our theory of change is missing, which is whether those students most in need of SBMH actually improved based on their receipt of services. While we assume that they do, and that this is the mechanism for improvement in the TAU, Enhanced, and Expanded conditions, we cannot prove that it is, based on the data available. We had hoped to have extensive data about improvement in symptoms of SBMH students including aggressive and externalizing behavior that affects school safety, but we were unable to obtain a significant number of parent consents to allow us to access those students' data. This is a common problem among SBMH programs, and some students fail to access services because of lack of parental approval. We encountered this problem in trying to obtain parental research consent for SBMH students in our study.

Another measurement challenge came in the administrative data available from CMS that included student infractions. About 10% of students changed schools, and often treatment condition within and across the years of the study. This number is not surprising, but we know that students who change schools are often the students who are most transient and in need of services due to environmental stressors like homelessness and housing instability, parental incarceration, and caregiver changes (Welsh, 2017). As such, the students who changed schools and changed treatment conditions could be some of those most likely to receive infractions, and therefore, most likely to affect school climate. This movement between schools is challenging for school-level designs in that multiple membership models and other statistical solutions are

extremely challenging with only 34 schools. These considerations led us to not examine administrative infraction data at the individual level, but instead assigned students to the first school they attended that school year, even if they changed school later in the school year. This made interpretation challenging.

4.4 RECOMMENDATIONS FOR FUTURE INTERVENTIONS AND RESEARCH

This study moves the field of SBMH and mental health responses to school safety issues forward in several ways. First, this study offers promise that some levels of treatment of student mental health problems may affect student safety and school climate in some cases. It appears that that effect is more likely to be endorsed by school staff than students, which may indicate that students do not perceive as much school safety improvement and actually experience more aggression and victimization than their school's staff perceive.

Second, this study describes the significant challenges in introducing intensive EBPs into real-world schools and sustaining implementation over several years. Not only was the introduction of SPARCS and DBT time-consuming, but it may have pulled providers away from interventions that they find to be effective in their individual school contexts. Implementation of SBMH programming may be most helpful if it is done in response to the particular needs of each school and with the buy-in of and cooperation from providers and administrators. Moreover, programs designed specifically for schools rather than for community or clinical settings may assist in balancing clinical needs of the students with feasibility concerns in school settings.

Third, despite measurement challenges, discipline infractions did not appear to be improved by the introduction of EBPs into SBMH or overall levels of services of SBMH. This suggests that more work is needed to understand the needs of students, staff, and school contexts in schools where large numbers of infractions occur. Last, although SSFs were an important part of the model designed to increase counselors' direct services to students, that did not appear to occur. More information is needed to understand how SSFs might work in some school contexts, depending on the school.

This study provides several suggestions for future research on school safety and school mental health service provision. Results suggest that increasing and improving the mental health care that high-need students receive at school has the potential to improve safety and overall climate for school staff and, to a lesser degree, for students. This study was situated in a single urban/suburban school district and used single EBPs for Tier 2 (SPARCS) and Tier 3 (DBT) student needs. The field would benefit from additional studies conducted in other varied settings and using other EBPs. Broad studies involving many schools, staff, and students—such as this one—provide many advantages in terms of possible research design and analysis. Complementary advantages would be gained from focused study of service implementation and outcomes, both for students who participate in services and the broader student body and staff, in a small number of schools. Such focused study could help shed light on the interplay of services and outcomes in specific schools.

The study provides direct implications for service provision in schools around the country. First, schools need better mechanisms for collection of mental health service provision data. Providers are often overburdened by paperwork and administrative duties and documentation of each student contact is very challenging. Even when schools and districts attempt to introduce new service logs or data collection efforts, uptake may be slow and implementers may need significant training and technical assistance. In addition to data about how many services are provided, data should include information about the fidelity of those services to the intended service model. Second, future research should continue to examine training and implementation of EBPs in schools, including DBT and SPARCS. These studies should measure individual student's treatment outcomes so that participants' response to treatment can be understood. Studies should also explore training and implementation of other, less intensive EBPs that fit the needs of students at different schools. The current number of school psychologists, school social workers, and school counselors is below the national recommended average in schools in North Carolina, where we did our study, and in many other states around the country. In order to meet the needs of students, schools need more providers. Additional research is needed on how best to increase the availability of service providers, the

4-23

benefits that accrue to students and schools, and the advantages and disadvantages of different approaches to promote school safety and student well-being.

References

Aarons, G. A., Glisson, C., Hoagwood, K., Kelleher, K., Landsverk, J., & Cafri, G. (2010). Psychometric properties and U.S. National norms of the Evidence-Based Practice Attitude Scale (EBPAS). *Psychological Assessment, 22*(2), 356–365. https://doi.org/10.1037/a0019188

Austin, M. J., & Claassen, J. (2008). Implementing evidence-based practice in human service organizations: Preliminary lessons from the frontlines. *Journal of Evidence-Based Social Work, 5*(1-2), 271–293. https://doi.org/10.1300/J394v05n01_10

Ballard, K. L., Sander, M. A., & Klimes-Dougan, B. (2014). School-related and social-emotional outcomes of providing mental health services in schools. *Community Mental Health Journal, 50*(2), 145–149. https://doi.org/10.1007/s10597-013-9670-y

Bradshaw, C. P. (2015). Translating research to practice in bullying prevention. *American Psychologist, 70*(4), 322-332. https://doi.org/10.1037/a0039114

Brand, S., Felner, R., Shim, M., Seitsinger, A., & Dumas, T. (2003). Middle school improvement and reform: Development and validation of a school-level assessment of climate, cultural pluralism, and school safety. *Journal of Educational Psychology, 95*(n3), 570–588.

Brand, S., Felner, R. D., Seitsinger, A., Burns, A., & Bolton, N. (2008). A large scale study of the assessment of the social environment of middle and secondary schools: The validity and utility of teachers' ratings of school climate, cultural pluralism, and safety problems for understanding school effects and school improvement. *Journal of School Psychology, 46*(5), 507–535. https://doi.org/10.1016/j.jsp.2007.12.001

Bruns, E. J., Walrath, C., Glass-Siegel, M., & Weist, M. D. (2004). School-based mental health services in Baltimore: Association with school climate and special education referrals. *Behavior Modification, 28*(4), 491–512. https://doi.org/10.1177/0145445503259524

Buhs, E. S., Ladd, G. W., & Herald, S. L. (2006). Peer exclusion and victimization: Processes that mediate the relation between peer group rejection and children's classroom engagement and achievement? *Psychology (Savannah, Ga.), 98*(1), 1–13. https://doi.org/10.1037/0022-0663.98.1.1

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology, 13*(1), 3-21. https://doi.org/10.1007/bf00988593

Cowan, K. C., Vaillancourt, K., Rossen, E., & Pollitt, K. (2013). *A framework for safe and successful schools [Brief]*. National Association of School Psychologists.

Douglas, K. A., & Morris, C. A. W. (2017). Assessing counselors' self-efficacy in suicide assessment and intervention. *Counseling Outcome Research and Evaluation, 6*(1), 58-69. https://doi.org/10.1177/2150137814567471

Elliott, D. S., Hamburg, B. A., & Williams, K. R. (1998). *Violence in American schools: a new perspective*. Cambridge University Press. http://www.loc.gov/catdir/description/cam029/98020495.html

Fabelo, T., Thompson, M. D., Plotkin, M., Carmichael, D., Marchbanks III, M., & Booth, E. A. (2011). *Breaking schools' rules: A statewide study of how school discipline relates to students' success and juvenile justice involvement*. The Council of States Governments Justice Center and the Public Policy Research Institute.

Farrell, A. D., Sullivan, T. N., Esposito, L. E., Meyer, A. L., & Valois, R. F. (2005). A latent growth curve analysis of the structure of aggression, drug use, and delinquent behaviors and their interrelations over time in urban and rural adolescents. *Journal of Research on Adolescence 15*(2), 179-204.

Gastic, B. (2008). School truancy and the disciplinary problems of bullying victims. *Educational Review, 60*(4), 391-404.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine. https://books.google.com/books?id=oUxEAQAAIAAJ

Goodenow, C. (1993). The psychological sense of school membership among adolescents: Scale development and educational correlates. *Psychology in the Schools, 30*(1), 79–90. https://doi.org/10.1002/1520-6807(199301)30:1<79::aid-pits2310300113>3.0.co;2-x

Greene, M. B. (2005). Reducing violence and aggression in schools. *Trauma, Violence & Abuse, 6*(3), 236-253. https://doi.org/10.1177/1524838005277406

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., Duda, S. N., & Consortium, R. E. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics, 95*, 103208. https://doi.org/10.1016/j.jbi.2019.103208

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics, 42*(2), 377-381. https://doi.org/10.1016/j.jbi.2008.08.010

Hawker, D., & Boulton, M. (2000). Twenty years' research on peer victimization and psychosocial maladjustment: A meta-analytic review of cross-sectional studies. *Journal of Child Psychology and Psychiatry, 41*, 441-455. https://doi.org/10.1111/1469-7610.00629

Hoagwood, K. E., Jensen, P. S., Acri, M. C., Olin, S. S., Lewandowski, R. E., & Herman, R. J. (2012). Outcome domains in child mental health research since 1996: Have they changed and why does it matter? *Journal of the American Academy of Child and Adolescent Psychiatry, 51*(12), 1241-1260.

Hymel, S., & Swearer, S. M. (2015). Four decades of research on school bullying: An introduction. *American Psychologist, 70*(4), 293–299. https://doi.org/10.1037/a0038928

Mellin, E. A., Taylor, L., & Weist, M. D. (2014). The expanded school mental health collaboration instrument [school version]: Development and initial psychometrics. *School Mental Health, 6*(3), 151–162. https://doi.org/10.1007/s12310-013-9112-6

Mellin, E. A., Taylor, L., Weist, M. D., & Lockhart, N. C. (2016). The expanded school mental health collaboration instrument [community version]: Development and initial psychometrics. *School Mental Health, 8*(2), 305–318. https://doi.org/10.1007/s12310-015-9164-x

Multisite Violence Prevention Project. (2004). *Description of measures: Cohort-wide student survey (Unpublished)*. Available from the Centers for Disease Control and Prevention, National Center for Injury Prevention and Control.

Nakamoto, J., & Schwartz, D. (2010). Is peer victimization associated with academic achievement? A meta‐analytic review. *Social Development, 19*(2), 221–242. https://doi.org/10.1111/j.1467-9507.2009.00539.x

R-3

National Center for Education Statistics (NCES), & U.S. Department of Education. (2000). *Public school teacher questionnaire: Schools and staffing survey—1999–2000 school year*. National Center for Education Statistics, Institute of Education Sciences. Retrieved September 16, 2020, from http://nces.ed.gov/surveys/sass/

Orpinas, P. (2009). *Measurement manual: Aggression, victimization, and social skills scales*. University of Georgia. Retrieved September 16, 2020, from https://www.yumpu.com/en/document/read/14796920/measurement-manual-college-of-public-health-university-of-

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

Padgett, D. (2008). *Qualitative methods in social word research (2nd ed.)*. Sage.

Paluta, L. (2015). *Examining the processes and outcomes of a school-based mental health pilot (Doctoral dissertation)*. Retrieved from OhioLINK. (osu1429606299).

Portnow, S., Downer, J. T., & Brown, J. (2018). Reductions in aggressive behavior within the context of a universal, social emotional learning program: Classroom- and student-level mechanisms. *Journal of School Psychology, 68*, 38–52. https://doi.org/10.1016/j.jsp.2017.12.004

Powers, J. D., Edwards, J. D., Blackman, K. F., & Wegmann, K. M. (2013). Key elements of a successful multi-system collaboration for school-based mental health: in-depth interviews with district and agency administrators. *The Urban Review, 45*(5), 651-670. https://doi.org/10.1007/s11256-013-0239-4

Powers, J. D., Wegmann, K. M., Blackman, K. F., & Swick, D. C. (2014). Increasing access to mental health services in schools through community-engaged research: results from a one-year pilot project. *School Social Work Journal, 39*(1), 73-89(17).

Reinke, W. M., Stormont, M., Herman, K. C., Puri, R., & Goel, N. (2011). Supporting children's mental health in schools: teacher perceptions of needs, roles, and barriers. *School Psychology Quarterly, 26*(1), 1-13. https://doi.org/10.1037/a0022714

Ringeisen, H., Henderson, K., & Hoagwood, K. (2003). Context matters: Schools and the "research to practice gap" in children's mental health. *School Psychology Review, 32*(2), 153-168.

Sanchez, A. L., Cornacchio, D., Poznanski, B., Golik, A. M., Chou, T., & Comer, J. S. (2018). The effectiveness of school-based mental health services for elementary-aged children: a meta-analysis. *Journal of the American Academy of Child and Adolescent Psychiatry, 57*(3), 153-165. https://doi.org/10.1016/j.jaac.2017.11.022

Sullivan, A. L., Klingbeil, D. A., & Van Norman, E. R. (2013). Beyond behavior: multilevel analysis of the influence of sociodemographics and school characteristics on students' risk of suspension. *School Psychology Review, 42*(1), 99-114.

Sullivan, T. N., Farrell, A. D., & Kliewer, W. (2006). Peer victimization in early adolescence: Association between physical and relational victimization and drug use, aggression, and delinquent behaviors among urban middle school students. *Development and Psychopathology, 18*(1), 119–137. https://doi.org/10.1017/S095457940606007X

Sullivan, T. N., Helms, S. W., Bettencourt, A. F., Sutherland, K., Lotze, G. M., Mays, S., Wright, S., & Farrell, A. D. (2012). A qualitative study of individual and peer factors related to effective nonviolent versus aggressive responses to problem situations among adolescents with high incidence disabilities. *Behavioral Disorders, 37*(3), 163–178. https://doi.org/10.1177/019874291203700304

U.S. Department of Education. (2014). *Guiding principles: A resource guide for improving school climate and discipline.* U.S. Department of Education.

Weist, M. D., Youngstrom, E. A., Stephan, S., Lever, N., Fowler, J., Taylor, L., McDaniel, H., Chappelle, L., Paggeot, S., & Hoagwood, K. (2014). Challenges and ideas from a research program on high-quality, evidence-based practice in school mental health. *Journal of Clinical Child and Adolescent Psychology, 43*(2), 244-255. https://doi.org/10.1080/15374416.2013.833097

Welsh, R. O. (2017). School hopscotch: a comprehensive review of K–12 student mobility in the United States. *Review of Educational Research, 87*(3), 475-511.

Yeager, D. S., Fong, C. J., Lee, H. Y., & Espelage, D. L. (2015). Declines in efficacy of anti-bullying programs among older adolescents: Theory and a three-level meta-analysis. *Journal of Applied Developmental Psychology, 37*, 36–51. https://doi.org/10.1016/j.appdev.2014.11.005

Appendix A:
Interview Protocol

1. Compared to previous years do you think that school safety (bullying, fights, drug/alcohol use, sexual harassment, etc.) has changed in CMS?

    a. If "Yes", what changes have you seen that support your opinion that safety has improved
        i. To what do you attribute the improved safety?
    b. If "No", has school safety gotten worse or stayed the same?
        i. What would need to happen for school safety to improve?

2. Do you believe that the counseling or therapeutic services available in CMS schools are being effective in helping students or not?

    a. What evidence or data do you have that supports your belief?
    b. Have you seen a change in the level of services available to students with emotional/behavioral issues in CMS?

3. How would you describe the capacity of CMS student support services (Social Workers, Counselors, Psychologists) and community therapists to serve students in need?

    a. Have you observed any recent changes in how student support services staff work together or with community therapists at your school?

4. **(Enhanced Schools Only) Have you participated in the SPARCS/DBT process?**

    a. If "Yes", In what ways have your participated?

        i. In what ways do you communicate with Social Workers, Counselors, Psychologists about SPARCS/DBT?

        ii. Have you discussed DBT/SPARCS with the community therapist assigned to your school?

            1. What have been the greatest challenges in implementing SPARCS/DBT?

            2.

    b. Do you believe that SPARCS and/or DBT have been effective in helping students?

    c. What would make it easier to implement DBT/SPARCS?

        i. Probe: Are your efforts to implements DBT/SPARCS supported by administrators and supervisors?

        ii. What components of DBT/SPARCS are most difficult to implement? Why?

        iii. Do you believe DBT/SPARCS is being implemented in the way it was designed?

        iv. Probe: Have you encountered challenges in adhering to the guidelines provided in your training for DBT/SPARCS?

    d. Would you like additional training in DBT/SPARCS?

        i. If "Yes" then,

        ii. What specific type of instruction regarding DBT/SPARCS would you like to receive?

        iii. In what formats would you like to see the technical assistance provided? That is, would you prefer in-person training, web conferences, phone calls, e-mail, a knowledge base, or something else?

   e. Do you believe that sufficient data is being collected to accurately assess the effectiveness of DBT/SPARCS?

      i. If "Yes" then, which data points do you believe are the most informative for researchers?

5. When you have questions about addressing the needs of students with behavioral/psychological issues which student services staff member (counselor, psychologist, social worker) or community therapist at your school would you ask?

   a. In what ways have you found them helpful in the past?
   b. Do you think that they would be willing to participate in an interview addressing the questions we have discussed today?

Appendix B: Provider Survey Correlations

	EBPAS Requirement	EBPAS Appeal	EBPAS Openness	EBPAS Divergence	Suicide Self-Efficacy Characteristics	Suicide Self-Efficacy SI	Org Capacity	Org. Culture	Org. Staff	Org. Impl.
EBPAS Requirement	1									
EBPAS Appeal	0.38* (<.001)	1								
EBPAS Openness	0.25* (<.001)	0.42* (<.001)	1							
EBPAS Divergence	-0.07 (0.32)	-0.07 (0.39)	-0.03 (0.73)	1						
Suicide Self-Efficacy Characteristics	0.11 (0.15)	0.12 (0.13)	0.19* (<.001)	0.11 (0.15)	1					
Suicide Self-Efficacy SI	0.06 (0.40)	0.04 (0.57)	0.10 (0.18)	0.10 (0.20)	0.71* (<.001)	1				
Organizational Capacity	-0.10 (0.22)	-0.04 (0.59)	0.27* (<.001)	0.15 (0.05)	0.15* (0.04)	0.05 (0.51)	1			
Organizational Culture	0.01 (0.88)	0.06 (0.42)	0.30* (<.001)	0.14* (0.07)	0.26* (<.001)	0.13 (0.09)	0.64* (<.001)	1		
Organizational Staff	0.07 (0.37)	0.12 (0.13)	0.39* (<.001)	0.10 (0.19)	0.24* (<.001)	0.14 (0.08)	0.70* (<.001)	0.67* (<.001)	1	
Organizational Implementation	0.03 (0.67)	0.12 (0.12)	0.40* (<.001)	0.09 (0.28)	0.14 (0.07)	0.06 (0438)	0.75* (<.001)	0.72* (<.001)	0.79* (<.001)	1

Note: Values are correlation coefficients, with p-values in parentheses. EBPAS, Evidence-Based Practice Attitude Scale.