Received: 17 December 2019 | Revised: 25 May 2020 | Accepted: 3 June 2020

DOI: 10.1002/pits.22420

RESEARCH ARTICLE

WILEY

Implementing a comprehensive behavioral health model for social, emotional, and behavioral development in an urban district: An applied study

Jill Battal[1] | Melissa M. Pearrow[2] | Amy J. Kaye[3]

[1]Behavioral Health Services Department, Boston Public Schools, Boston, Massachusetts, USA

[2]Department of Counseling and School Psychology, University of Massachusetts Boston, Boston, Massachusetts, USA

[3]Children's Hospital Neighborhood Partnership, Boston Children's Hospital, Boston, Massachusetts, USA

Correspondence
Jill Battal, Behavioral Health Services Department, Boston Public Schools, 2300 Washington Street, Boston 02119, Massachusetts, USA.
Email: jsnyder2@bostonpublicschools.org

Abstract

Urban communities are disproportionality impacted by issues such as community violence, leaving children at higher risk of trauma, and increased adverse childhood experiences. This applied study addresses a research-to-practice gap by demonstrating the longitudinal impact of a district-initiated, school-wide multitiered systems of support approach to universal instruction, intervention, and assessment—the Comprehensive Behavioral Health Model. Results from more than 1200 students over a 3-year period indicate that this universal, integrated approach is associated with improved outcomes for students with a demonstrated level of risk, with the greatest impact on students with risk of internalizing behaviors. The findings of this applied study highlight the benefits of a systemic approach and enhance the discourse on robust universal screenings that can identify needs and assets of children within urban communities.

KEYWORDS

CBHM, MTSS, universal behavior screening, urban schools

1 | INTRODUCTION

The behavioral health needs of children have been characterized as a "silent epidemic" with grave implications for students, families, and communities (Anderson & Cardoza, 2016). Despite increasing familiarity with the estimate that 20% of U.S. children meet criteria for behavioral health disorders (Caldarella, Young, Richardson, Young, & Young, 2008; Perou, Bitsko, & Blumberg, 2013), our society's ability to respond to this level of need continues to

fall short with estimates that only one-third of youth with mental health disorders ever receive any treatment, and increased barriers to accessing services among Black and Hispanic youth. Among the third of students who do receive services, schools play an integral role, providing services to over half of students identified with behavioral disorders, and over a third of all students receiving services (Merikangas et al., 2011).

The behavioral health needs of students within urban communities provide additional urgency with which to approach this study. Children living in metropolitan areas are 20% more likely to experience adverse childhood experiences (ACEs; e.g., parental separation, childhood neglect, unemployment) than children in nonurban populations (Hunt, Slack, & Berger, 2017; Mersky, Topitzes, & Reynolds, 2013). Exposure to ACEs results in increased risk for more severe symptomology, such as depression, substance use, and antisocial behaviors, as well as for a plethora of other negative outcomes, such as physical health disparities, diminished life satisfaction, and poor academic performance (e.g., Anda et al., 2002; Chapman et al., 2004; Lassen, Steele, & Sailor, 2006; Mersky et al., 2013; Reiss, 2013; Schilling, Aseltine, & Gore, 2007; Sugai & Horner, 1999). At the same time, families in urban communities have historically faced significant barriers to accessing behavioral health services. Early detection and services reduce later incidence of behavioral health problems, and the school setting offers a convenient location that reduces a variety of access barriers (S. M. Blake, Ledsky, Goodenow, & O'Donnell, 2001; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011).

The prevalence of behavioral health disorders among children indicates a significant—and as of yet, unmet—need for targeted and intensive services. However, this prevalence also compels the field to consider a public health approach to addressing this need. A public health approach focuses on prevention and early intervention, placing the emphasis on promotion of population-based well-being, early assessment and interventions, comprehensive services, and positive outcomes (Doll & Cummings, 2008; Harrison et al., 2004; Hess, Short, & Hazel, 2012). The education system is a highly efficient setting in which to promote behavioral health, as it is a setting in which children are engaged on a regular and consistent basis over the majority of their formative years, and it has reasonable access to most parents and guardians (Weissberg, Caplan, & Sivo, 1989). Furthermore, behavioral health is closely linked to academic success in the educational setting (Durlak et al., 2011).

1.1 | The Comprehensive Behavioral Health Model (CBHM)

The need for a public health approach to preventing behavioral health difficulties has led many schools to implement multitiered systems of support (MTSS), a best practice model that includes community behavioral health services incorporated through an interconnected systems framework (ISF; Barrett, Eber, & Weist, 2013), tiered instruction in social-emotional skills, tiered behavioral health assessment procedures, and data based decision-making (Cook et al., 2015; Florida's Positive Behavior Support Project, 2011; Kansas State Department of Education, 2017; Stockslager, Castillo, Hines, Batsche, & Curtis, 2013). The CBHM is one such MTSS model for social, emotional, and behavioral health developed to ensure access to high-quality behavioral health services across schools in a large urban school district (for more information see Pearrow, Amador, & Dennery, 2016). The CBHM framework was developed by Boston Public Schools. Implementation began in 10 Boston schools during the 2012–2013 school year, and 10 new schools were trained and coached on the model each year.

Boston has a compelling need for a comprehensive approach to behavioral health services to address the widespread needs of students. The Health of Boston's Children report, authored by the Boston Public Health Commission, noted that one out of every five Boston children experienced two or more ACEs. Like so many large districts, Boston schools vary considerably in configuration, student population, and available resources and supports. The 125 schools in Boston educate approximately 57,000 students speaking over 70 languages and from diverse backgrounds—approximately 42% identify as Latinx and 35% as Black. Additionally, 70% are identified as economically disadvantaged by the state department of education (MA DESE). The core components of the CBHM are described below.

1.1.1 | Instruction

In an MTSS model, high-quality core instruction implemented with fidelity for all students is considered the first layer (Tier 1) of prevention for academic and behavioral difficulties in school (Harn, Chard, Biancarosa, & Kame'enui, 2011). For this reason, schools are increasingly incorporating evidence-based, high-quality core instruction at this tier for social, emotional and behavioral skills in classroom settings through the use of school-wide positive behavioral interventions and supports frameworks (SWPBIS; Horner et al., 2009), direct instruction on social and emotional learning (SEL) skills such as in the Second Step® (Gottfredson et al., 2010; Low, Cook, Smolkowski, & Buntain-Ricklefs, 2015) and other similar curricula (Durlak et al., 2011; Taylor, Oberle, Durlak, & Weissberg, 2017), or an integration of these two approaches. For students who require additional support to master certain social, emotional, and behavioral skills, additional "tiers" of evidence-based instruction and/or supports are provided, based on the identified level of need (Freeman et al., 2016). Despite variation among school-level goals and resources in CBHM, one common concern is limited fidelity to universal (Tier 1) instruction. As such, CBHM implementation has focused on supporting schools to design and implement high-quality core instructional programming for social, emotional, and behavioral health at the Tier 1 level.

1.1.2 | Assessment

A basic tenant of MTSS is matching students with appropriate instruction based on their individual level of need (Harn et al., 2011). To do this, schools must have accurate data sources to understand students' levels of need across a variety of domains and skills. In an MTSS model, universal screening provides one such data source and is critical to a truly systemic approach to prevention and early intervention (Cook, Volpe, & Livanis, 2010). In addition to universal screening, MTSS models utilize progress monitoring assessment procedures to ensure that Tier 2 and Tier 3 instruction is resulting in desired student outcomes (Bolt, Ysseldyke, & Patterson, 2010).

Universal screening is widely accepted as a best practice (Cook et al., 2010), yet remains uncommon in most educational systems with only between 2% and 12% of schools utilizing behavioral health screening instruments (Bruhn, Woods-Groves, & Huddle, 2014; Romer & McIntosh, 2005). In the absence of universal screening procedures, schools, and districts most often rely upon teacher referral or discipline data to identify student need. Relying upon teacher referrals alone is problematic because some teachers over-refer and others under-refer (Kalberg, Lane, & Menzies, 2010). Additionally, discipline data have the tendency to over-identify students of color (C. P. Bradshaw, Mitchell, O'Brennan, & Leaf, 2010; Carter, Fine, & Russell, 2014), reducing opportunities for learning (Scott & Barrett, 2004) and increasing the risk for truancy and subsequent school failure (Gregory, Skiba, & Noguera, 2010; Skiba, Michael, Nardo, & Peterson, 2002; Skiba et al., 2011). As opposed to strategies that primarily recognize acting out behaviors, universal screening systematically identifies students at risk for a range of concerns (Cook et al., 2010; Davis, Young, Hardman, & Winters, 2011).

CBHM uses the Behavior Intervention Monitoring Assessment System, Second Edition (BIMAS-2; McDougal, Bardos, & Meier, 2011) for its universal screening and progress monitoring procedures. In CBHM schools, teachers complete the BIMAS-2 for every student in their class two times a year (e.g., Fall and Spring). The results are used by schools to understand student strengths and needs at the school-wide and grade levels, as well as identify students at risk for social, emotional, and behavioral difficulties. At the Tier 1 level, CBHM school-wide teams use the universal screening data to ensure that the instructional programming being implemented is effective in increasing students' mastery of social, emotional, and behavioral skills and learning goals. At Tiers 2 and 3, the BIMAS is used to monitor student progress. The BIMAS tool was selected by the district following a pilot phase exploring available tools. Multiple stakeholders, including teachers, parents, school leaders, school psychologists, and district staff reviewed tools for usability and strong psychometrical considerations (Pearrow, Snyder, & Kaye, 2017). In particular, reliability, validity, multirater usability, progress monitoring, and sensitivity of scores to change over time supported the use of the BIMAS tool within this framework (Jenkins et al., 2014).

1.1.3 | Implementation considerations

Empirical accounts of MTSS emphasize that district and school-level leadership, capacity building among staff, and clear communication and collaboration are each essential to ensuring buy in, and smooth and effective implementation (Florida's Positive Behavior Support Project, 2011; Forman & Crystal, 2015; Kansas State Department of Education, 2017).

CBHM considers these points in its model by involving coaching at both the district and school levels. Within a school, a school psychologist or school social worker serves as the "internal" coach or primary coordinator of MTSS activities. They are prepared to fulfill this role through coaching provided by a district level (i.e., "external") coach. External coaches provide technical assistance to not only internal coaches but also to school administrators and Tier 1 team members. Additional professional learning opportunities are provided throughout the year to bring together a variety of school-level stakeholders to hear from other school teams what has been working and why (Pearrow et al., 2016).

1.2 | PURPOSE

The purpose of this applied study is not to revalidate the empirically supported interventions (e.g., SWPIS, SEL instruction, universal screening) and implementation considerations (e.g., coaching, technical assistance) comprising CBHM, but rather to shed light on the utility of these procedures in the field. The procedures, samples, implementation, and evaluation components presented are intended to illustrate the adaptation and application of empirically based prevention and intervention procedures to improve the behavioral health outcomes of students in the context of an urban public school district. The following evaluation questions guided the study.

- What are the needs of students within an urban, highly diverse district, as measured by the number of students identified as at-risk for various social, emotional, and behavioral difficulties?
- How does CBHM impact *all* students within a 3-year implementation period?
- How does CBHM impact those students identified in the first year as at-risk for social, emotional and behavioral difficulties within a 3-year implementation period?

2 | METHOD

2.1 | School-level implementation

The CBHM was conceptualized to provide a *framework* that could be applied across a variety of grade configurations, programming, and specific needs of schools within a large urban district. As such, specific programming may differ from school to school, while certain key elements anchor the framework across each school setting. Each school was invited to complete a proposal to voluntarily participate in this district initiative, and all participating schools were asked to sign an agreement with the district articulating their commitment to implement the following: (a) develop or refine SWPBIS structures (e.g., core values, behavior matrix, lesson plans, reinforcement systems, consequence system, data system) for their school community, (b) provide direct SEL instruction for a minimum of 30 min per week, and (c) complete universal screening using the BIMAS-2. The agreement additionally emphasized data-based decision making such that teams within the building (e.g., Tier-1 teams, grade-level teams, student support teams) would each *use* universal screening data to inform instructional decisions at each level. Despite variation among school-level goals and resources in CBHM, one common concern was limited fidelity of universal (Tier 1) instruction. As such, CBHM implementation focused on supporting schools to design and implement high-quality core instructional programming for social, emotional, and behavioral health at the Tier 1 level.

Although the framework has implications for additional supports and structures at Tier 2 and 3 levels, implementation over the years examined primarily focused on Tier 1 structures. Tier 2 interventions included the use of check-in/check-out (Hunter, Chenier, & Gresham, 2013) by school-based staff. Tier 3 interventions were provided by behavioral health service providers (e.g., school psychologists and school social workers) who received central training on evidence-based intervention programs including CBT-based interventions such as Coping Cat (Kendall, 1994) and Strong Start (Merrell, Whitcomb, & Parisi, 2009).

2.2 | District-level implementation

District-level structures were developed to support CBHM implementation within schools. Specifically designed structures supporting the CBHM framework included key competency drivers, and organization drivers designed to support school-based implementation. Competency drivers included (a) professional development and coaching for school psychologists who serve as the "internal coach" or primary coordinator of CBHM activities within the school, (b) professional development and consultation for school psychologist and school leader dyads, (c) professional development for Tier 1 school teams to develop an action plan for CBHM implementation, and (d) professional development for teachers that is developed and coordinated by school leaders, school psychologists, and Tier 1 teams. Organizational drivers included the (a) selection and (b) purchasing of key components of the CBHM framework, such as the BIMAS-2 and SEL Curricula (Second Step and Dialectical Behavior Therapy Skills in Schools), as well as (c) the development of protocols for data-based decision making following the BIMAS-2 universal screening. Leadership for this model was initially conceived within the behavioral health services department, with the supervisor of school psychologists and school social workers responsible for serving as "internal coaches." However, throughout implementation, leadership has intentionally been more broadly distributed across additional stakeholders, such as school leaders, teachers, and school-based Tier 1 team members, central office administrators, families, and partnering agencies. In addition, district-level commitment was obtained by the School Committee, who incorporated into district policy that an additional 10 schools would be selected each year to install this framework across the district (Pearrow et al., 2016).

2.3 | Fidelity of implementation

The fidelity of implementation at both the school and district level was monitored in various ways over the course of this study. Halfway through the first school year included in this analysis, district capacity was increased to support a full-time position responsible for coordinating data, evaluation and research related to this study (Pearrow et al., 2017). Through this capacity, concrete process and outcome evaluation indicators at both the district and school levels were identified. As is typical in program evaluation, these indicators shifted as implementation and data collection structures were refined, and while the ultimate indicators are richer as a result, they are also not available for all years of implementation. Early on, district staff supporting CBHM identified the need for a measure of school-level implementation, and reviewed multiple tools and indicators (e.g., screening completion rates). While multiple measures and indicators are now identified and consistently collected, only one primary indicator was available for all schools over the course of this study. Although SEL curricula were purchased and professional development was provided for all teaching staff, there was not a quantitative measure of SEL curricula use within classrooms to help inform these data. Therefore, the primary indicator used to measure school level implementation for this analysis is the Tiered Fidelity Inventory (TFI; Algozzine et al., 2014).

The TFI is a valid and reliable measure used to assess the extent to which schools are applying the core features of SWPBIS (Algozzine et al., 2014). The TFI is divided into three sections to assess: Tier 1 universal interventions and supports, Tier 2 targeted interventions and supports, and Tier 3 intensive interventions and supports. For the purposes of the present study, CBHM schools completed the Tier 1 section only, given the initial

focus of CBHM coaching on building Tier 1 universal intervention and supports. The Tier 1 section of the TFI includes 15 items rated on a 3-point Likert scale. The TFI yields a Tier 1 ratio score in the form of the overall percentage. A criterion of 70% implementation has been described as indicative of adequate fidelity, while a score of 100% suggests near-perfect fidelity (Mcintosh et al., 2017). Literature also suggests that implementing SWPBIS constitutes a systemic change that can take between 3 and 5 years to yield true impacts on a school environment and the associated student-level change (C. Bradshaw, Koth, Bevans, Ialongo, & Leaf, 2008).

TFI Tier 1 data was available for all schools included in this sample during years 2 (2014–2015) and 3 (2015–2016). During Year 2, Tier 1 ratio scores ranged from 20% to 73%, with an average score across all schools of 47%. During Year 3, Tier 1 ratio scores ranged from 23% to 87%, with an average score across all schools of 53%. Although the average TFI Tier 1 ratio score fell below the 70% criterion in both Years 2 and 3, the trendline overall shows improving fidelity. In addition, a few individual schools in both years demonstrated "adequate fidelity" at the 70% or higher criterion. This information, while limited, provides evidence that CBHM implementation at both the school and district levels was occurring to some extent.

2.4 | Participants

The current applied study presents data for students attending schools who were implementing CBHM during the 2013–2014 school year, and it follows these students for 3 years of implementation. During the Fall of 2013, there were 20 participating schools representing each neighborhood of the city and educating 9025 students between kindergarten and grade 12. Fourteen of the 20 schools completed teacher BIMAS ratings during Fall 2013, resulting in a total of 3060 students screened at Time 1. Over a 3-year academic period, screenings were consistently completed at each time point (Fall 2013 = Time 1; Fall 2014 = Time 2; Fall 2015 = Time 3) for 1428 of these students.

The vast majority of the 1428 students screened were in the elementary grades (87%). Given this imbalance of distribution, the analysis was restricted to students in elementary school (grades kindergarten through five) with screening data from all three data collection periods, for a total *n* of 1247 students. Included in this sample were 656 (52.6%) male and 591 (47.4%) female students. More than half of students were enrolled in kindergarten ($n = 311$, 24.9%) or first grade ($n = 363$, 29.1%) during the Fall of 2013. An additional 252 (20.2%) students were enrolled in second grade, 222 (17.8%) in third grade, 72 (5.8%) in fourth grade, and 27 (2.2%) in fifth grade. In total, the students included in this sample were representative of district demographics, with 439 (35.2%) identified as African American, 405 (32.5%) Hispanic/Latino, 231 (18.5%) Asian/Pacific Islander, 143 (11.5%) White, and 29 (2.3%) multiple or other races. Within the district, 35% of students identify as African American, 40% as Hispanic/Latino, 9% as Asian, and 13% as White. Teacher demographics for the specific group of staff responsible for entering ratings is not available to report. Teacher demographics are collected at the district level, but not available within the BIMAS system, and impossible to retroactively match with the information available. Given the intentional representativeness of the schools participating in CBHM (Pearrow et al., 2016), as well as the representativeness of the student data included in this analysis, it is suspected that this teacher group reflects broader district teaching demographics: 3924.8 FTE (55%) identified as White, 1963.6 FTE (27%) identified as African American, 893.9 (12%) identified as Hispanic, and the majority (5,366.7, 75%) identified as female (MA DESE).

2.5 | Measures

2.5.1 | Behavior Intervention Monitoring Assessment System

After piloting multiple tools and obtaining feedback from key stakeholders, the BIMAS-2 (McDougal et al., 2011) was selected due to its strong psychometric properties, user-friendly nature, and opportunity to identify problems,

intervene early and disrupt racial inequity in student discipline (J. Blake, Gregory, James, & Hasan, 2016). This system allowed for an accessible, data- and norm-based strategy for organizing resources while capturing changes in both problem and adaptive behaviors. It also allowed for continuous data collection from kindergarten until graduation, a critical aspect of this initiative, while still providing relevant information that accounts for the student's developmental level.

The BIMAS-2 is a brief measure with versions that can be completed by teachers, parents, support staff (including clinicians), or students themselves. For the purposes of this project, only teacher ratings were included as this is the version primarily used by CBHM schools in the sample, and the data consisted of age-based T-scores. During a Fall time period consistent across all CBHM schools, elementary school teachers in participating schools completed the BIMAS-2 screening for all children in their classrooms, as students generally have one primary classroom teacher. Teachers indicated the frequency of behaviors for each student in classrooms using a 5-point scale, ranging from *Never* to *Very often*, and all teacher ratings on the 34-item BIMAS were completed electronically via the web-based system. The BIMAS yields T-scores (mean = 50; *SD* = 10) and percentile ranks for two adaptive (Social and Academic Functioning) and three problem (Conduct, Negative Affect, and Cognitive/Attention) scales. The Conduct scale screens for externalizing concerns including aggression (physical or verbal), defiance, anger management, and engaging in risky behaviors. The Negative Affect scale screens for internalizing concerns such as depression, anxiety, and withdrawal. The Cognitive/Attention scale screens for difficulties with attention, concentration, and focus. The Social scale monitors positive relationships with peers such as the ability to recognize body language and social cues, and to express emotions. The Academic Functioning scale monitors behaviors related to academic competence, such as the ability to follow directions, come to class prepared, and sustain effort on academic tasks.

Higher T scores on the behavior concern scales indicate greater levels of concern, with T scores greater than 70 indicating "High Risk," T scores between 60 and 69 indicating "Some Risk" and T scores less than 60 indicating "Low Risk." Higher T scores on the two adaptive scales indicate strengths, with T scores of 60 or greater indicating "Strength," T scores between 41 and 59 indicating "Typical" functioning, and T scores of 40 or lower indicating "Concern."

Reliability analyses reported in the BIMAS Technical Manual (McDougal et al., 2011) indicated good levels of internal consistency for all scales, with Cronbach's α values between .81 and .91. When BIMAS-2 was taken twice within a 2- to 4-week period, correlation coefficients ranged from .79 to .96, indicating strong temporal stability. Finally, correlation coefficients for teacher and self-reports on the BIMAS ranged from .54 to .64, and those for teacher and parent-reports from .79 to .86, indicating strong inter-rater reliability. The BIMAS Technical Manual (McDougal et al., 2011) also indicated good content validity, with moderate to strong positive correlations (*r*s = .38 to .78, all $p < .001$) with several measures of similar constructs and moderate to strong negative correlations (*r*s = −.31 to −.71, all $p < .001$) with measures of distinct constructs. Discriminant function analyses indicated that the BIMAS had good overall correct classification rates (82.5%–85.2%), indicating its appropriateness for use as a screening measure. Finally, a progress monitoring study indicated that the BIMAS-2 was able to capture significant decreases in pre- to post-treatment with large effect sizes across subscales (*d*s = 1.0 to 1.5 for problem scales and *d*s = −.7 to −.8 for adaptive scales).

2.6 | Data analysis

To answer the specified evaluation questions, multiple analyses were conducted. First, to understand the risk levels of the student population, descriptive statistics for BIMAS-2 scores were explored. To examine the impact of the CBHM on all students as well as by risk level, a mixed between-within subjects analysis of variance (ANOVA) was performed to examine both the main effect of time, as well as the interaction between time and risk level for each BIMAS-2 scale. The BIMAS-2 does not provide an overall index score to summarize all scores, so results must instead be examined by scale. Across analyses, "time" served as the within-subjects factor with three levels of the

dependent variable (T score at Fall 2013 = T1, Fall 2014 = T2, Fall 2015 = T3), and the risk level from T1 served as the between-subjects factor. Post-hoc analyses including Tukey's HSD and one-way ANOVAs by scale were performed to ascertain effect size statistics for each group. Effect sizes were explored using partial η^2 with 0.01 being considered a small effect, 0.06 a moderate effect, and 0.14 a large effect (Cohen, 1998).

3 | RESULTS

Descriptive statistics representing students' social, emotional, and behavioral strengths and needs are summarized in Table 1, which shows the mean, standard deviation, and range on each BIMAS scale for each time period. In each case, the mean score falls within the average range. The observed variability is consistent across scales, with slightly greater variation observed in student scores on the Cognitive/Attention scale at each time point.

3.1 | Student strengths and needs

Table 2 presents the distribution of students across the BIMAS-2 defined risk levels at each time point. At T1, the majority of students fell within the "Low Risk" range on BIMAS concern scales (Conduct, Negative Affect, Cognitive/Attention). Roughly 20% of students emerged with "Some Risk" on each these scales, and between 7% and 8% of all students were identified with "High Risk" on each of the BIMAS-2 concern scales. Teacher ratings identified the most students at "High Risk" on the Negative Affect (n = 103, 8.3%) scale, suggesting slightly increased concern around internalizing behaviors such as depression, anxiety, and withdrawal. On the BIMAS-2 adaptive scales (Social, Academic Functioning), the majority of students fell within the "Typical" or "Strength" range, with teacher ratings more frequently indicating "Strength" in academic functioning than social skills. Teacher ratings indicated "Concern" for just under one-third of the sample on each scale. Minimal changes in this overall distribution were observed, despite increases in the number of students rated in the "Low Risk" or "Strength" categories between T1 and T3 for four out of five scales, with the exception of the Cognitive/Attention scale.

Overall, more than a quarter of the students demonstrated risk in at least one area, with the highest number of students demonstrating risk for negative affect. Approximately one in eight (12.2%) students demonstrated risk in all three areas of *problem* behaviors, while an additional 16.4% demonstrated risk in all *adaptive* behaviors. In addition, 5.7% demonstrated risk in all five BIMAS-2 scales.

TABLE 1 Descriptive statistics and main effects for behavior intervention monitoring assessment (BIMAS) scores over time

		Time 1 Fall 13	Time 2 Fall 14	Time 3 Fall 15		
	n	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	*F*	Partial η^2 [a]
Conduct	1247	54.17 (9.27)	53.44 (9.10)	53.63 (9.75)	3.67*	0.01
Negative Affect	1247	53.98 (10.05)	52.26 (10.37)	52.55 (10.85)	10.46*	0.02
Cognitive/Attention	1247	52.86 (11.44)	53.46 (11.55)	54.01 (12.30)	6.13*	0.01
Social	1247	45.26 (9.64)	46.58 (10.61)	47.00 (10.67)	14.95*	0.02
Academic Functioning	1247	46.71 (10.34)	47.71 (10.93)	47.24 (10.78)	5.75*	0.01

[a]0.01 = small, 0.06 = moderate, 0.14 = large (Cohen, 1998).

*p < .01.

TABLE 2 Student strengths and needs at each screening period

	Fall 2013		Fall 2014		Fall 2015	
Group	***n***	**%**	***n***	**%**	***n***	**%**
Conduct						
Low Risk	892	71.5	942	75.5	913	73.2
Some Risk	264	21.2	222	17.8	238	19.1
High Risk	91	7.3	83	6.7	96	7.7
Negative Affect						
Low Risk	879	70.5	916	73.5	895	71.8
Some Risk	265	21.3	243	19.5	256	20.5
High Risk	103	8.3	88	7.1	96	7.7
Cognitive/Attention						
Low Risk	917	73.5	906	72.7	849	68.1
Some Risk	239	19.2	215	17.2	251	20.1
High Risk	91	7.3	126	10.1	147	11.8
Social						
Strength	84	6.7	135	10.8	152	12.2
Typical	813	65.2	800	64.2	791	63.4
Concern	350	28.1	312	25.0	304	24.4
Academic Functioning						
Strength	150	12.0	211	16.9	193	15.5
Typical	735	58.9	720	57.7	723	58.0
Concern	362	29.0	316	25.3	331	26.5

3.2 | Outcomes for all students

Overall, as seen in Table 1, while the large sample size contributed to statistically significant main effects on four out of five of the BIMAS-2 scales (e.g., Conduct, Negative Affect, Social and Academic Functioning), the changes observed are not indicative of clinically significant differences in functioning. When examining the whole sample, average scores on each BIMAS-2 scale remained within the "Low Risk" ranges at each time point, with consistent variation observed at each time point. Furthermore, the effect sizes were negligible, with only the change on the social scale rising to the threshold of a "small" effect (partial $\eta^2 = 0.02$).

For each BIMAS scale, the mixed between within-subjects ANOVA revealed a significant interaction effect. In other words, the amount of change in students' scores was affected by initial risk level at T1. These results suggest that the impact of CBHM cannot be accurately understood by examining trends for all students. In each case, statistically significant interaction effects rendered interpretation of the overall effect on all students meaningless. Thus, the outcomes for each risk level are described below, as well as in Table 3.

3.3 | Outcomes by risk level

The data for the third research question consisted of scales that identified concerns and those that identified adaptive functioning, and then the data were analyzed by risk level. All five analyses are described below.

3.3.1 | Concern scales

There was a significant interaction effect between risk group and time for the Conduct, Negative Affect, and Cognitive/Attention scales (see Table 3). Based on a review of estimated marginal means and post hoc analyses (see Table 4), students in the "Some Risk" and "High Risk" groups in each of the subscales showed significant improvement in comparison to those in the "Low Risk" groups whose scores migrated slightly toward the meanwhile remaining within the "Low Risk" range. Although the changes in the "Low Risk" groups were statistically significant, the effect sizes were small (partial $\eta^2 < 0.06$) and not representative of clinically meaningful differences in functioning. Alternatively, the effect sizes associated with changes in the "Some Risk" and "High Risk" groups were large (partial $\eta^2 > 0.21$) and represent clinically meaningful improvements in functioning.

3.3.2 | Adaptive scales

There was a significant interaction effect between risk group and time for both the Social and the Academic Functioning scales (see Table 3). Based on a review of estimated marginal means and post hoc analyses (see Table 4), students in the "Concern" risk groups showed significant improvement, while students in the "Strength" risk groups showed significant decreases in scores. Students in the "Typical" risk groups showed no significant changes over time. The effect sizes for students in both the "Strength" and "Concern" groups across scales were very large (partial $\eta^2 > 0.22$). However, it is important to note that while students in the "Strength" group at T1 decreased, they remained on average within the "Typical" range. Therefore, this change does not indicate any presence of clinical concern. On the other hand, students in the "Concern" group on average moved into the "Typical" range, indicating a meaningful reduction in risk for students identified early.

3.3.3 | Concern scales "High Risk" group

Across all scales, as seen in Table 3, the most improvement was observed for the most at-risk students (e.g., those in the "High Risk" group). On all three BIMAS-2 concern scales, effect sizes for students in the "High Risk" group at T1 were substantial (0.41 Conduct, 0.49 Negative Affect, 0.27 Cognitive/Attention). Furthermore, these large effects translated into clinically meaningful changes in functioning, with students in the "High Risk" group on average moving into the "Some Risk" (Conduct and Cognitive/Attention) or "Low Risk" (Negative Affect) ranges over time.

TABLE 3 Interaction of effect of time and risk level on behavior intervention monitoring assessment (BIMAS-2) scores

Scale	Wilks' λ	*F*	Partial η^2 [a]
Conduct	0.775	84.58*	0.12
Negative Affect	0.710	115.95*	0.16
Cognitive/Attention	0.873	43.86*	0.07
Social	0.741	100.60*	0.14
Academic Functioning	0.774	85.01*	0.12

[a]0.01 = small, 0.06 = moderate, 0.14 = large (Cohen, 1998).

*$p < .01$.

TABLE 4 Changes in behavior intervention monitoring assessment schedule (BIMAS) subscale T scores by risk group over time

		Time 1 Fall 13	Time 2 Fall 14	Time 3 Fall 15		
	n	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	*F*	Partial η^2 [a]
Conduct						
High Risk	91	74.35 (3.00)	64.78 (9.83)	64.86 (9.338)	63.01*	0.41
Some Risk	264	63.22 (2.74)	56.35 (9.23)	58.01 (10.19)	69.30*	0.21
Low Risk	892	49.44 (5.36)	51.42 (7.82)	51.19 (8.30)	25.99*	0.03
Negative Affect						
High Risk	103	74.23 (3.75)	58.17 (11.97)	59.90 (11.30)	97.38*	0.49
Some Risk	265	63.08 (2.65)	54.71 (10.88)	55.12 (11.02)	92.10*	0.26
Low Risk	879	48.86 (6.35)	51.26 (9.67)	50.91 (10.23)	22.09*	0.03
Cognitive/Attention						
High Risk	91	73.62 (2.89)	65.27 (10.30)	66.22 (9.65)	33.14*	0.27
Some Risk	239	64.37 (2.38)	60.11 (9.85)	61.46 (10.49)	20.44*	0.08
Low Risk	917	47.79 (8.51)	50.56 (10.60)	50.85 (11.35)	42.58*	0.04
Social						
Concern	350	34.36 (5.14)	40.73 (9.47)	42.99 (10.52)	118.02*	0.25
Typical	813	47.73 (4.92)	48.38 (9.91)	48.10 (10.25)	1.39*	0.00
Strength	84	66.73 (4.31)	53.64 (11.19)	53.07 (10.00)	76.39*	0.48
Academic						
Concern	362	34.72 (5.40)	40.92 (9.28)	41.66 (9.57)	98.45*	0.21
Typical	735	48.94 (4.78)	49.13 (9.98)	48.25 (10.27)	3.02*	0.00
Strength	150	64.72 (3.13)	57.15 (9.34)	55.77 (8.778)	81.61*	0.35

[a]0.01 = small, 0.06 = moderate, 0.14 = large (Cohen, 1998).

*$p < .01$.

3.3.4 | Concern scales "Some Risk" group

Similarly, effect sizes for students in the "Some Risk" group at T1 were large for Conduct (0.21) and Negative Affect (0.26), and moderate for Cognitive/Attention (0.08); however, in this case effect sizes suggest about half the amount of improvement noticed among students in the "High Risk" range. Still, the large and moderate effect sizes constituted clinically meaningful changes in functioning. Students in the "Some Risk" category on average moved into the "Low Risk" category for Conduct and Negative Affect, and within one *T* score of the "Low Risk" range for Cognitive/Attention.

3.3.5 | Adaptive scales "Concern" group

Likewise, on both BIMAS-2 adaptive scales, effect sizes for students in the "Concern" group were large (0.25 Social and 0.21 Academic Functioning) and clinically meaningful, with students on average moving into the "Typical" range. The fact that the effect sizes here are about half as strong as those observed on concern scales may in part be due to the fact that the risk group on the BIMAS adaptive scales is represented by the one "Concern" category, while the BIMAS concern scales break the risk group into two levels (e.g., "High Risk" and "Some Risk").

3.3.6 | Both adaptive and concern scales not at-risk groups

In comparison to at-risk students, students who were not identified with risk at T1 showed very minor changes in BIMAS scales. In general, students with "Low Risk" on the concern scales or "Typical"/"Strength" on the adaptive scales remained within the "Low Risk," "Typical," or "Strength" risk ranges. On the adaptive scales, results seem to suggest a decline in student outcomes for students identified with "Strength" at T1; however, it is important to note that while student scores from the "Strength" group at T1 decreased, they remained on average within the "Typical" range. Therefore, this change does not indicate any presence of clinical concern. Across all scales, follow up pairwise comparisons confirmed that scores across risk groups changed significantly between T1 and T2 ($p < .001$), and were maintained without significant change into T3 ($p > .200$).

4 | DISCUSSION

The purpose of this applied study was to explore the utility of a comprehensive, systemic MTSS approach for behavioral health prevention under applied conditions in the context of a highly diverse, urban school district. The findings indicate significant improvements within the span of 1 year for students in CBHM schools with elevated levels of risk, without deleterious effects on students with normative social, emotional, and behavioral health. Moreover, the overall gains demonstrated over the first year are sustained into the second year.

The interaction between risk level and time was a meaningful finding. What may at first appear to suggest minimal changes at the overall level can conceal changes in functioning for students presenting level of risk, or in other words, the students whose behaviors most interfere with their learning demonstrated the most improvement. This makes sense because there would not necessarily be any expected or desired change in behavior among students without behavioral concerns. On the other hand, it would be expected and desired that students with some or high risk are provided with *early intervention* in the form of additional instruction and/or support to build skills and mitigate risk. Beyond answering the specific evaluation questions posed, this applied study makes several unique contributions to the literature.

The current data support the claim that implementation of an MTSS model in an applied setting can have a positive effect on students' social, emotional, and behavioral functioning. While previous research highlights comparable levels of impact among similar researcher-led school-based interventions, these findings have not previously been demonstrated when these same interventions have been district- and school-led (Burns, Appleton, & Stehouwer, 2005; Burns & Symington, 2002), and there is generally a dearth of research on these large scale, district-led initiatives. The present findings, however, address this gap in the literature and suggest that large scale, district-led approaches grounded in MTSS and ISF, where schools and community agencies partner to address student needs, can result in tremendous gains in real-world settings, particularly for students who are frequently not referred or identified as in need of services due to the internalizing presentation of symptoms.

4.1 | Prevalence of risk in urban context

Although universal screening is a recognized best practice, its use in school districts is not common (Bruhn et al., 2014; Romer & McIntosh, 2005), particularly in large urban districts. Universal screening methods utilized in this applied study identified the prevalence rates of the behavioral health needs of elementary-aged students in a large, urban school district. The high rates found suggest a significant need for a broad-reaching, systemic approach to enhancing the SEL environment for all students while also identifying students in need of supplementary supports. In this sample, the rate of students who demonstrated a level of risk in all five scales approximates that less than 6% of students were in need of Tier 3 services. These most vulnerable students may experience the greatest

benefit from additional individualized family and community-based supports. In this way, using systematic screening methods for identification provides districts with effective strategies for allocating resources in a more equitable way to those with the greatest level of needs.

Teachers in this sample more frequently endorsed concerns around internalizing behaviors than other areas of concern, and interestingly enough, the greatest improvements were realized for these students identified early with risk for internalizing concerns. Certainly, additional investigations are necessary to fully understand the cause of this change; however, this applied study suggests preliminary evidence for the role of school-wide preventative supports in addressing students' internalizing concerns. The unique benefits for students who present with risk for internalizing concerns (e.g., anxiety, depression) was an unforeseen outcome. The effect sizes associated with changes for these students in particular are comparable to more intensive one-on-one and group therapies for students who struggle with negative affect. For example, in meta-analytic studies in the area of depression among youth, the effect sizes for psychotherapy to address adolescent depression was $d = 0.34$ (Weisz, McCarty, & Valeri, 2006). In meta-analytic studies using cognitive-behavioral therapy to address anxiety with children, group interventions yielded a medium effect on student outcomes, while individually delivered services demonstrated a large effect (Reynolds, Wilson, Austin, & Hooper, 2012). In the present applied study, the effect sizes realized for students with internalizing concerns exceeds these estimates.

4.2 | Limitations and future directions

There are several limitations of the present study that reflect the challenges of research in applied settings and indicate the need for future research in this area. Out of the potential 9025 students enrolled in CBHM schools, data were available on the three identified screening time points for only 16% students. For a variety of reasons, the data were not completed at all three time points for all possibly included students. In most cases, school-level factors contributed to inconsistent or incomplete implementation of universal screening. This is a reality of implementing best practices such as screening, which has been supported by literature in a variety of settings (Durlak & DuPre, 2008). In other cases, students may have transferred to non-CBHM schools that did not have access to the screening system, or transferred out of the district. These reasons, in addition to several others, resulted in less than half of the potential students being included. As such, the findings are based on a limited sample of students, and although representative of the demographics of the district's population, this sample may be narrow in some ways. For example, it includes students who were in classrooms for 3 years with teachers who completed the screening tools, which might suggest something unique about the teacher and their commitment to CBHM and behavioral health. Additionally, in requiring three consistent screening data points, these results do not estimate the impact on students who transition out of a CBHM school or out of the district altogether, perhaps sometimes for reasons related to behavioral health challenges. Further research regarding the factors that impact a teacher's screening tool completion could help add to our understanding of ways to improve screening rates and thus lead to a larger number of students with longitudinal data.

This study is also limited in that, as with applied research in educational settings, it does not include random assignment of students or schools to participation in the CBHM intervention. As such, although these findings suggest strong evidence for the CBHM, firm conclusions cannot be drawn regarding causality. It is possible that improved student outcomes could be related to a variety of third variables, such as characteristics related to reasons the schools were chosen to participate in the intervention (e.g., readiness for change, systems in place before the intervention). Future analyses of CBHM outcomes is planned to examine student outcomes by school, taking into effect the relationship between outcomes and varying levels of fidelity to the various components of the model. Moving forward, additional indicators of fidelity at both the district and school-level implementation will be available to incorporate in analyses. In addition multilevel statistical modeling to account for the nested nature of each of these data sets (e.g., students, within classrooms, within schools) will additionally illustrate trends in student, school & district outcomes. This will provide

additional information regarding the key components of the model to determine how to most effectively and cost-efficiently implement these large-scale, district-led systemic initiatives.

5 | CONCLUSION

The current data address the significant gap in the literature demonstrating the application of systemic interventions in real-world, urban schools. In summary, they provide support for urban districts to engage in comprehensive and multitiered systemic interventions to address and improve student behavioral health. Differential risk factors contribute to the functioning of children in urban environments, and this study demonstrates the prevalence rates of risks as well as the unique way in which universal screening can contribute to the feedback loops that are a crucial component allowing for greater specificity of service delivery. The improvements demonstrated for students with some level of risk indicate the importance of focusing on the SEL environment. Given their potential for significant impact on student behavioral health, including much overlooked internalizing behaviors, the need for further investigation into this comprehensive and systems-based approach cannot be overstated.

ORCID

Jill Battal http://orcid.org/0000-0002-8583-5036
Melissa M. Pearrow https://orcid.org/0000-0003-2518-0813
Amy J. Kaye https://orcid.org/0000-0003-1776-1538

REFERENCES

Algozzine, B., Barrett, S., Eber, L., George, H., Horner, R., Lewis, T., … Sugai, G. (2019). *School-wide PBIS Tiered Fidelity Inventory*. OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. Retrieved from www.pbis.org

Anda, R. F., Whitfield, C. L., Felitti, V. J., Chapman, D. P., Edwards, V. J., & Dube, S. R. (2002). Adverse childhood experiences, alcoholic parents, and later risks of alcoholism and depression. *Psychiatric Services*, *53*, 1001–1009. https://doi.org/10.1176/appi.ps.53.8.1001

Anderson, M., & Cardoza, K. (2016). Mental health in schools: A hidden crisis affecting millions of students. Retrieved from http://www.npr.org/sections/ed/2016/08/31/464727159/mental-health-in-schools-a-hidden-crisis-affecting-millions-of-students

Barrett, S., Eber, L., & Weist, M. (2013). Advancing education effectiveness: Interconnecting school mental health and school-wide positive behavior support. Retrieved from http://www.pbis.org/common/cms/files/pbisresources/FinalMonograph.pdf

Blake, J., Gregory, A., James, M., & Hasan, G. W. (2016). Early warning signs: Identifying opportunities to disrupt racial inequities in school discipline through data-based decision making. *School Psychology Forum*, *10*, 289–306.

Blake, S. M., Ledsky, R., Goodenow, C., & O'Donnell, L. (2001). Receipt of school health education and school health service among adolescent immigrants in Massachusetts. *Journal of School Health*, *71*, 105–113.

Bolt, D. M., Ysseldyke, J., & Patterson, M. J. (2010). Students, teachers, and schools as sources of variability, integrity, and sustainability in implementing progress monitoring. *School Psychology Review*, *39*(4), 612–630.

Bradshaw, C., Koth, C., Bevans, K., Ialongo, N., & Leaf, P. (2008). The impact of school-wide positive behavioral interventions and supports (PBIS) on the organizational health of elementary schools. *School Psychology Quarterly*, *23*(4), 462–473.

Bradshaw, C. P., Mitchell, M. M., O'Brennan, L. M., & Leaf, P. J. (2010). Multilevel exploration of factors contributing to the overrepresentation of black students in office disciplinary referrals. *Journal of Educational Psychology*, *102*, 508–520.

Bruhn, A. L., Woods-Groves, S., & Huddle, S. (2014). A preliminary investigation of emotional and behavioral screening practices in K-12 schools. *Education and Treatment of Children*, *37*, 611–634.

Burns, M. K., Appleton, J. J., & Stehouwer, J. D. (2005). Meta-analytic review of responsiveness-to-intervention research: Examining field-based and research-implemented models. *Journal of Psychoeducational Assessment*, *23*(4), 381–394. https://doi.org/10.1177/073428290502300406

Burns, M. K., & Symington, T. (2002). A meta-analysis of prereferral intervention teams: Student and systemic outcomes. *Journal of School Psychology*, *40*(5), 437–447.

Caldarella, P., Young, E. L., Richardson, M. J., Young, B. J., & Young, K. R. (2008). Validation of the systematic screening for behavior disorders in middle and junior high school. *Journal of Emotional and Behavioral Disorders*, *16*, 105–117.

Carter, P., Fine, M., & Russell, S. (2014). *Discipline disparities overview*. Discipline disparities: A research-to-practice collaborative. Bloomington, IN: The Equity Project at Indiana University.

Chapman, D. P., Whitfield, C. L., Felitti, V. J., Dube, S. R., Edwards, V. J., & Anda, R. F. (2004). Adverse childhood experiences and the risk of depressive disorders in adulthood. *Journal of Affective Disorders*, *82*, 217–225. https://doi.org/10.1016/j.jad.2003.12.013

Cook, C. R., Burns, M. F., Browning-Wright, D., & Gresham, F. M. (2010). *Transforming school psychology in the RTI Era: A guide for administrators and school psychologists*. Palm Beach: LRP Publications.

Cook, C. R., Frye, M., Slemrod, T., Lyon, A. R., Renshaw, T. L., & Zhang, Y. (2015). An integrated approach to universal prevention: Independent and combined effects of PBIS and SEL on youths' mental health. *School Psychology Quarterly*, *30*(2), 166–183

Cook, C. R., Volpe, R. J., & Livanis, A. (2010). Constructing a roadmap for future universal screening research beyond academics. *Assessment for Effective Intervention*, *35*(4), 197–205.

Davis, S. D., Young, E. L., Hardman, S., & Winters, R. (2011). Screening for emotional and behavioral disorders. *Principal Leadership*, *11*(9), 12–17.

Doll, B., & Cummings, J. A. (2008). *Transforming school mental health services*. Bethesda, MD: National Association of School Psychologists and Corwin Press.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, *41*(3-4), 327–350. https://doi.org/10.1007/s10464-008-9165-0

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405–432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

Farmer, E. M., Burns, B. J., Phillips, S. D., Angold, A., & Costello, E. J. (2003). Pathways into and through mental health services for children and adolescents. *Psychiatric Services*, *54*, 60–66. https://doi.org/10.1176/appi.ps.54.1.60

Florida's Positive Behavior Support Project. (2011). Implementing a multi-tiered system of support for behavior: A practical guide. Retrieved from http://www.pbis.org/Common/Cms/files/pbisresources/RTIB_Guide_101811_final.pdf

Forman, S. G., & Crystal, C. D. (2015). Systems consultation for multitiered systems of supports (MTSS): Implementation issues. *Journal of Educational and Psychological Consultation*, *25*(2-3), 276–285. https://doi.org/10.1080/10474412.2014.963226

Freeman, J., Simonsen, B., McCoach, D. B., Sugai, G., Lombardi, A., & Horner, R. (2016). Relationship between school-wide positive behavior interventions and supports and academic, attendance, and behavior outcomes in high schools. *Journal of Positive Behavior Interventions*, *18*(1), 41–51. https://doi.org/10.1177/1098300715580992

Gottfredson, G. D., Pas, E. T., Nebbergall, A. J., Nese, J. F., Strein, W., & Shaw, F. (2010). *An experimental outcome evaluation of the second step elementary universal prevention program*. Technical report.

Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher*, *39*, 59–68. https://doi.org/10.3102/0013189x0935762

Harn, B. A., Chard, D. J., Biancarosa, G., & Kame'enui, E. J. (2011). Coordinating instructional supports to accelerate at-risk first-grade readers' performance: An essential mechanism for effective RTI. *The Elementary School Journal*, *112*(2), 332–355.

Harrison, P. L., Cummings, J. A., Dawson, M., Short, R. J., Gorin, S., & Palomeres, R. (2004). Responding to the needs of children, families, and schools: The 2002 multisite conference on the Future of school psychology. *School psychology review*, *33*, 12–33.

Hess, R. S., Short, R. J., & Hazel, C. E. (2012). *Childrens comprehensive mental health services in the schools'*. New York, NY: Routledge.

Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A. W., & Esperanza, J. (2009). A randomized, wait-list controlled effectiveness trial assessing school-wide positive behavior support in elementary schools. *Journal of Positive Behavior Interventions*, *11*, 133–144.

Hunt, T. K., Slack, K. S., & Berger, L. M. (2017). Adverse childhood experiences and behavioral problems in middle childhood. *Child Abuse & Neglect*, *67*, 391–402. https://doi.org/10.1016/j.chiabu.2016.11.005

Hunter, K. K., Chenier, J. S., & Gresham, F. M. (2013). Evaluation of check in/check out for students with internalizing behavior problems. *Journal of Emotional and Behavioral Disorders*, *22*(3), 135–148.

Jenkins, L., Demaray, M., Wren, N., Fredrick, S., Lyell, K., Magers, A., … Tennant, J. (2014). A critical review of five commonly-used social-emotional and behavioral screeners for elementary or secondary schools. *Contemporary School Psychology*, *18*(4), 241–254.

Kalberg, J. R., Lane, K. L., & Menzies, H. M. (2010). Using systematic screening procedures to identify students who are nonresponsive to primary prevention efforts: Integrating academic and behavioral measures. *Education and Treatment of Children*, *33*, 561–584.

Kansas State Department of Education. (2017). *Kansas multi-tier system of supports: Structuring guide for systems, 2017–2018 academic year*. Topeka, KS: Kansas MTSS Project, Kansas Technical Assistance System Network. Retrieved from https://ksdetasn.s3.amazonaws.com/uploads/resource/upload/1593/Systems_Structuring_Guide.pdf

Kataoka, S., Zhang, L., & Wells, K. (2002). Unmet need for mental health care among U.S. children: Variation by ethnicity and insurance status. *American Journal of Psychiatry*, *159*(9), 1548–1555.

Kendall, P. C. (1994). Treating anxiety disorders in children: Results of a randomized clinical trial. *Journal of Consulting and Clinical Psychology*, *62*(1), 100–110. https://doi.org/10.1037/0022-006X.62.1.100

Lassen, S. R., Steele, M. M., & Sailor, W. (2006). The relationship of school-wide positive behavior support to academic achievement in an urban middle school. *Psychology in the Schools*, *43*(6), 701–712.

Low, S., Cook, C. R., Smolkowski, K., & Buntain-Ricklefs, J. (2015). Promoting social–emotional competence: An evaluation of the elementary version of Second Step. *Journal of School Psychology*, *53*, 463–477.

Massachusetts Department of Elementary and Secondary Education (MA DESE). (n.d.). Staffing data by race, ethnicity, gender by full-time equivalents (2013-14). School/district profiles: Boston (00350000). Retrieved from http://profiles.doe.mass.edu/profiles/teacher.aspx?orgcode=00350000&orgtypecode=5&leftNavId=817&&fycode=2014

McDougal, J., Bardos, A., & Meier, S. (2011). *Behavior intervention monitoring assessment system technical manual*. Toronto, Canada: Multi-Health Systems.

Mcintosh, K., Massar, M., Algozzine, R., George, H., Horner, R., Lewis, T., & Swain-Bradway, J. (2017). Technical adequacy of the SWPBIS tiered fidelity inventory. *Journal of Positive Behavior Interventions*, *19*(1), 3–13.

Merikangas, K. R., He, J. P., Burstein, M., Swendsen, J., Avenevoli, S., Case, B., … Olfson, M. (2011). Service utilization for lifetime mental disorders in US adolescents: Results of the National Comorbidity Survey–Adolescent Supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry*, *50*(1), 32–45.

Merrell, K. W., Whitcomb, S. A., & Parisi, D. M. (2009). *Strong start, pre-K: A social & emotional learning curriculum*. Baltimore, MD: Paul H. Brookes Publication.

Mersky, J. P., Topitzes, J., & Reynolds, A. J. (2013). Impacts of adverse childhood experiences on health, mental health, and substance use in early adulthood: A cohort study of an urban, minority sample in the U.S. *Child Abuse and Neglect*, *37*(11), 917–925. https://doi.org/10.1016/j.chiabu.2013.07.011

Pearrow, M., Amador, A., & Dennery, S. (2016). Boston public school's comprehensive behavioral health model: An overview. *Communiqué*, *45*(3), 1–22.

Pearrow, M., Amador, A., Dennery, S., Cohen, M., Snyder, J., & Shachmut, L. (2016). Boston public school's comprehensive behavioral health model: Organizational structures. *Communiqué*, *45*(4), 16–17.

Pearrow, M., Snyder, J., & Kaye, A. (2017). Boston's comprehensive behavioral health model: Research and evaluation. *Communiqué*, *45*(6), 10–12.

Perou, R., Bitsko, R. H., Blumberg, S. J., Pastor, P., Ghandour, R. M., Gfroerer, J. C., … Centers for Disease Control and Prevention. (2013). Mental health surveillance among children—United States, 2005-2011. *MMWR Surveillance Summaries*, *62*, 1–3.

Reiss, F. (2013). Socioeconomic inequalities and mental health problems in children and adolescents: A systematic review. *Social Science & Medicine*, *90*, 24–31. https://doi.org/10.1016/j.socscimed.2013.04.026

Reynolds, S., Wilson, C., Austin, J., & Hooper, L. (2012). Effects of psychotherapy for children and adolescents: A meta-analytic review. *Clinical Psychology Review*, *32*, 251–262. https://doi.org/10.1016/j.cpr.2012.01.005

Romer, D., & McIntosh, M. (2005). The roles and perspectives of school mental health professionals in promoting adolescent mental health. In D. L. Evans, E. B. Foa, R. E. Gur, H. Hendin, C. P. O'Brien, M. E. P. Seligman, T. Walsh & Annenberg Foundation Trust at Sunnylands (Eds.), *Treating and preventing adolescent mental health disorders: What we know and what we don't know: A research agenda for improving the mental health of our youth* (pp. 597–615). New York, NY: Oxford University Press. https://doi.org/10.1093/9780195173642.003.0032

Schilling, E. A., Aseltine, R. H., & Gore, S. (2007). Adverse childhood experiences and mental health in young adults: A longitudinal survey. *BMC Public Health*, *7*, 30.

Scott, T. M., & Barrett, S. B. (2004). Using staff and student time engaged in disciplinary procedures to evaluate the impact of school-wide PBS. *Journal of Positive Behavior Interventions*, *6*, 21–27. https://doi.org/10.1177/10983007040060010401

Skiba, R. J., Horner, R. H., Chung, C. G., Rausch, M. K., May, S. L., & Tobin, T. (2011). Race is not neutral: A national investigation of African American and Latino disproportionality in school discipline. *School Psychology Review*, *40*, 85–107.

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

Stockslager, K., Castillo, J. M., Hines, C. M., Batsche, G. M., & Curtis, M. J. (2013). *The Florida Problem Solving/Response to Intervention Project: Year 4 evaluation report*. Retrieved from http://floridarti.usf.edu/resources/format/pdf/yr4_eval_report.pdf

Sugai, G., & Horner, R. H. (1999). Discipline and behavioral support: Preferred processes and practices. *Effective School Practices*, *17*(4), 10–22.

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development*, *88*, 1156–1171. https://doi.org/10.1111/cdev.12864

Weissberg, R., Caplan, M., & Sivo, P. (1989). A new conceptual framework for establishing school-based social competence promotion programmes. In L. S.Bond, & B. E.Compas (Eds.), *Primary prevention and promotion in the schools* (pp. 255–296). Newbury Park, CA: Sage. https://doi.org/10.1037/0022-006X.65.1.93

Weisz, J., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: A meta-analysis. *Psychological Bulletin*, *132*, 132–149. https://doi.org/10.1037/0033-2909.132.1.132

How to cite this article: Battal J, Pearrow MM, Kaye AJ. Implementing a comprehensive behavioral health model for social, emotional, and behavioral development in an urban district: An applied study. *Psychol Schs*. 2020;1–17. https://doi.org/10.1002/pits.22420