ABSTRACT

SOUTHERN, LOUISE BUCHHOLZ. Organizational and Behavioral Outcomes in PBIS Versus Non-PBIS Elementary Schools: A One State Analysis. (Under the direction of Dr. Ed Sabornie).

Positive Behavioral Interventions and Supports (PBIS) is a tiered model of integrated, evidence-based strategies designed to enhance an entire school's learning, teaching and social environment in ways that promote positive academic and socio-behavioral outcomes for all students. This study compared disaggregated office disciplinary referral, crime and violence, and suspension data, as well as annual teacher turnover rates and results from the 2016 North Carolina Teacher Working Conditions survey, between PBIS and non-PBIS public elementary schools across North Carolina within the 2015-16 school year. This study also evaluated these behavioral outcome data and indicators of organizational health across the state's PBIS implementation designations. Results indicated that there were no significant differences in organizational health and behavioral outcomes between schools implementing Positive Behavioral Interventions and Supports (PBIS) and non-implementing schools. Level of PBIS implementation was significantly associated with teacher satisfaction but not annual teacher turnover rates. Level of PBIS implementation was not significantly influential on aggregate behavioral outcomes (i.e., short or long-term suspension rates, and crime and violence rates), nor when these outcomes were stratified across three ethno-racial categories. Furthermore, PBIS status and level of PBIS implementation were not significantly influential on office disciplinary referral (ODR) rates among students with disabilities versus non-disabled students.

Keywords: Positive Behavioral Interventions and Supports, PBIS, behavior, suspension, organizational health, elementary schools, disproportionality.

Organizational and Behavioral Outcomes in PBIS Versus Non-PBIS Elementary Schools: A One State Analysis

by
Louise Buchholz Southern

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Curriculum and Instruction

Raleigh, North Carolina

2021

APPROVED BY:

Dr. Ed Sabornie Chair of Advisory Committee	Dr. Cathy Crossland
Dr. Patricia Marshall	Dr. John Nietfeld

DEDICATION

This project is dedicated to my father, Ted Buchholz.

BIOGRAPHY

Louise Buchholz Southern was born in Alexandria, Virginia in 1977. She grew up in Columbia, South Carolina and is the oldest of six children. She attended Wofford College where she received a Bachelor of Science degree in psychology with a double major in English literature. Upon graduation, she worked as a technician in early autism intervention programs and served as an instructional assistant in a special education classroom. She relocated to Winston-Salem, North Carolina to attend Wake Forest University's graduate program in English Literature. Throughout this period, she continued to work in comprehensive intervention programs for young children with autism. She met her future husband in Winston-Salem and they got married in May, 2004. Upon relocation to Raleigh, North Carolina, she continued working in the field of autism intervention and obtained a Master of Education degree from North Carolina State University. She completed a graduate course sequence in behavior analysis through the University of North Texas and then obtained certification as a Behavior Analyst. Louise has worked in a variety of public education, clinical, research and community services settings supporting individuals with autism, intellectual disabilities and behavior disorders. She currently serves as the Associate Clinical Director for the Autism Society of North Carolina. Louise and her husband have three children.

ACKNOWLEDGMENTS

I am grateful for the unyielding support of my incredible advisor, Dr. Ed Sabornie. I always knew he believed in me, even when I questioned it myself. I so appreciate the brilliant insight and guidance offered by all of my committee members: Dr. Cathy Crossland, Dr. Patricia Marshall, and Dr. John Nietfeld. While I am relieved to be done, I am also sad, because I feel there is so much more to learn from all four of them. I am also thankful for my very patient statistics tutor, David Stokes.

I would not have been able to do this academic work without the support of my boss at the Autism Society of North Carolina, Dr. Aleck Myers. He gave me the time I needed to complete many of my graduate projects and offered gentle nudges to help keep me moving along. There have been so many family members who have cheered me on throughout this process. My grandparents, aunt and uncle, and siblings have always rooted for me as I worked towards any major goals. I can't properly express my gratitude to my mother, Julia Brennecke and to my father and step-mother, Ted and Monica Buchholz, who made so many things possible for me. They laid down the path to allow me to reach these milestones. Most importantly, I thank my husband, Michael, for supporting me in every way. He never (well, almost never) complained about all of the evenings and weekend days that were given over to these academic endeavors. Finally, I must acknowledge my three sons, Jacob, Jeremiah and Fields. They are my sun, moon and all of my stars.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1: Introduction

School safety has gained considerable attention in recent decades. High profile school shootings in the 1990's catalyzed efforts to make schools more secure (Skiba et al., 2006). In this era, problem behavior in school was often met with swift and certain consequences designed to regain control and assert adult and school authority (Skiba & Knesting, 2001). For example, most public schools in the United States typically responded to rule-breaking behaviors by increasing monitoring and surveillance methods, restating and reemphasizing rules and policies for problem behavior, and extending the continuum of punishment consequences (Sugai & Horner, 2002). Yet, calls for change to conventional discipline methods have grown in recent decades and are supported by a growing body of research. For instance, the 2001 report on the prevention of school violence published by the Office of the U.S. Surgeon General stated that "the United States is well past the point where anyone can claim that "nothing works" to prevent youth violence or to modify the destructive life courses of youths who are either engaged in or appear to be headed for lifestyles characterized by violence" (Department of Health and Human Services, 2001, p. 153).

In fact, there is no evidence that a system rooted in reactive responses promotes and *maintains* a safe, positive school environment where teaching time and learning opportunities are maximized (Sugai & Horner, 2002). Furthermore, data reveal no clear and sustained decreasing trends in school-based threats, injuries, and homicides in relation to the enactment of "zero-tolerance" policies (Mongan & Walker, 2012). Zero tolerance was originally defined as routinely enforced suspension and expulsion policies in response to weapons, drugs, and violent acts in the school setting. Over time, however, zero tolerance has come to refer to school or district-wide

policies that mandate predetermined, typically severe consequences or punishment for a vast array of rule-breaking behaviors (National Association of School Psychologists, 2001).

According to the U.S. Department of Education (2016b), of the 49 million students enrolled in public schools in 2011-12, 3.45 million (over 7% of students) were suspended out of school. However, suspension may not achieve intended results (i.e., sustained positive change in individual student behavior). Rather, the use of suspension is often associated with a range of negative outcomes, including increased levels of problem behavior over time, lower academic achievement, increased drop-out rates, and an overall less satisfactory school climate (Skiba et al., 2005). The drift in conceptual understanding and application of zero tolerance policies may blur delineations between major and minor infractions, thereby making students with disabilities (e.g., those with behavior disorders) and students from culturally and linguistically diverse backgrounds even more vulnerable to inequitable enforcement of such policies. In fact, data from the Department of Education Office for Civil Rights (2014) indicate that preschool through secondary level students from culturally and linguistically diverse backgrounds, and students with disabilities, are significantly more likely to be suspended or expelled from school than their White counterparts.

There are large-scale efforts designed to address many of these systemic problems. For example, in 2014, the U.S. Department of Justice and Department of Education led a national initiative on school discipline that yielded federal civil rights guidance designed to reduce the utilization of, and disparities in, disciplinary exclusion (Skiba et al., 2016). With the reauthorization of the Individuals with Disabilities Education Act (IDEA, 2004), the federal government further underscored the application of evidence-based positive behavior support strategies to prevent the exclusion of students with disabilities. Specifically, IDEA 2004 endorses

functional behavioral assessment procedures to address any behavior that results in a long-term removal. The IDEA 2004 also emphasizes the provision of incentives for evidence-based early literacy programs, positive behavioral interventions and supports, and early intervention services to reduce the need to label children as disabled.

Under IDEA, states are required to collect and evaluate on an annual basis statewide and local education agency (LEA) data to determine whether significant disproportionality based on race or ethnicity is occurring with respect to the identification of children as those with disabilities, including identification within particular disability categories; the placement of children in particular educational settings; and the incidence, duration, and type of disciplinary actions, including suspensions and expulsions (Department of Education, 2016a). As it pertains to the present study, it is also important to note that the North Carolina General Assembly passed a school discipline law stating that "removal of students from school, while sometimes necessary, can exacerbate behavioral problems, diminish academic achievement, and hasten school dropout" (N.C. Gen. Stat. § 115C-390.1, 2011, para. 1).

In addition, there is substantial emphasis at every level (federal, state, and local) on consistently and systematically employing comprehensive approaches that more effectively integrate and organize evidence-based practices to promote successful academic and behavioral student outcomes (e.g., Fixsen et al., 2009; Frank Porter Graham Children Development Institute, n.d.). Inherent within such comprehensive models is the premise that the provision of a guiding framework facilitates consistent and accountable delivery of evidence-based interventions.

Funded by the U.S. Department of Education's Office of Special Education Programs (OSEP), the Technical Assistance Center on Positive Behavioral Interventions and Supports

(PBIS) is intended to aid schools, districts, and states in improving behavioral and academic outcomes for all students, including students with disabilities and students from underrepresented groups. Over 26,000 U.S. preschools, elementary, middle, high schools, and alternative schools are currently implementing PBIS (Sugai & Horner, 2020). Included in the school teams that have been trained in the PBIS implementation framework are 3 states with more than 60% of schools involved in PBIS implementation, 9 states with more than 40%, and 16 states with more than 30% (Sugai & Simonsen, 2012).

An increasing number of schools in North Carolina have also adopted and implemented PBIS, from less than ten schools in 2002-03 (Algozzine et al., 2010) to over 600 schools in 2015-16, which represents 23% of all public and charter schools in the state (North Carolina Department of Public Instruction [NCDPI], 2017a). As of May 2017, PBIS training was provided in 95 of North Carolina's 100 counties (i.e., Local Education Agencies [LEAs]), plus 16 City LEAs and Charter Schools (NCDPI, 2017a). It is noteworthy that North Carolina was recently identified as one of nine states showing clear indicators of systemic implementation capacity: presence of state-wide PBIS coordinator, minimum of 500 schools implementing PBIS, and minimum of five years of implementation in collaboration with the National Technical Assistance Center for PBIS (Gage et al., 2014).

Broad PBIS Principles and Objectives

The fundamental tenet of PBIS is that the entire school culture and context can be changed and improved via systematic application of proactive and preventative strategies, adjustments to response procedures, and data-driven decision-making. Positive Behavioral Interventions and Supports is not a packaged intervention, a reinforcement procedure, a special education program, or a curriculum, but rather, it is a tiered "framework" of integrated, evidence-

based strategies designed to enhance an entire school's learning, teaching, and social environment in ways that promote positive academic and socio-behavioral outcomes for all students. Proponents of PBIS aim to move beyond a "piecemeal approach" to behavioral intervention. This framework of integrated tiered interventions lives and breathes on collective buy-in, effective and efficient systems of support, and formative evaluation.

Positive Behavioral Interventions and Supports emphasizes a systems approach to facilitate the adoption and durable implementation of evidence-based practices (OSEP Technical Assistance Center, 2021). The systems approach is also intended to enhance a collective sense of purpose and efficacy among staff, which, in turn, may help to transform a school's overall culture in ways that yield meaningful outcomes for all students. As Bandura (1993) described, "Teachers operate collectively within an interactive social system rather than as isolates. The belief systems of staffs create school cultures that can have vitalizing or demoralizing effects on how well schools function as a social system" (p. 141). Positive Behavioral Interventions and Supports is intended to produce positive effects on faculty and staff through such outcomes as improved classroom management, increased faculty retention, reduced faculty absenteeism, improved substitute performance and perception, and increased ratings of faculty effectiveness (NCDPI, 2017a).

The PBIS model emphasizes implementation of policies (at the state, district, and school level) that are intended to enhance equity in school discipline. Elements of effective policy pertaining to this objective include a specific and explicitly stated commitment to equity that is known by all stakeholders, family and community partnerships in policy development, proactive and positive behavior support practices, operationalized problem behavior definitions, clear and objective discipline procedures, removal or reduction of exclusionary practices (e.g., out of

school suspensions, office referrals), and data collection and review procedures that promote accountability for equitable student outcomes (Green et al., 2015). School officials are encouraged to use a full range of responses to violations of disciplinary rules (e.g., conferences, counseling, peer mediation, behavior contracts, instruction in conflict resolution and anger management) that do not remove a student from the classroom or school building (Department of Education, 2020).

Positive Behavioral Interventions and Supports encompasses the philosophies of person-centeredness and inclusion (Carr et al., 2002) as it addresses the "what" (i.e., what behaviors to change) and the "why"—targeting outcomes of social significance that will enhance quality of life for the student and relevant stakeholders. In addition, PBIS draws heavily from applied behavior analytic learning principles (e.g., motivation, generalization, stimulus control) and practices (e.g., shaping, reinforcement, prompting) as it addresses the "how" of behavior change. These research-based behavioral practices are organized along a continuum which is consistent with the response-to-intervention approach (Sugai & Horner, 2009). The intent is for all PBIS practices to be applied in tandem with ongoing progress monitoring, team-based decision making, training and coaching, and evaluations of implementation fidelity (Sugai & Simonsen, 2012). Thus, PBIS extends beyond traditional applied behavior analysis as it fully acknowledges and responds to the multifaceted, complexity of problem behavior (and its complicated context) not with one single intervention, but with a comprehensive structure in which interventions are organized and systematically applied (Carr et al., 2002).

Framework Features

Just as the behavioral strategies and practices in PBIS schools are organized along a continuum, so too are the support systems that would facilitate implementation and progress

monitoring. Specifically, these evidence-based behavioral practices and systems are organized into three tiers, that represent three levels of implementation intensity. Primary prevention (i.e., Tier I) is employed by *all* school staff for all students, and across all settings. In North Carolina, Tier I is described as the "differentiated core" within a multi-tiered system of support (MTSS; North Carolina Department of Public Instruction, 2017a). Key components of Tier I include 3-5 clearly articulated behavioral expectations that are taught to students via direct instruction, frequent delivery of positive reinforcement (e.g., acknowledgement, social praise, tangible reinforcement) for meeting school-wide expectations, and a continuum of logical consequences for problematic behavior.

At the classroom level, there are several key elements of PBIS Tier I: (a) effective design of the physical environment (e.g., thoughtful arrangement of desks and students, organized stations and materials, teacher accessibility to all students), (b) proactively taught classroom routines and procedures, (c) posted and regularly reviewed behavioral expectations, (d) high rates of varied response opportunities, (e) active supervision, and (f) frequent and specific acknowledgement of desirable behaviors (OSEP Technical Assistance Center, 2021). Systems components at the Tier I level of support include team organization, data to guide implementation, and PBIS training in annual staff development (Horner et al., 2010). Approximately 80-85% of the student population will respond to primary prevention strategies and systems and will not require more intensive levels of intervention (Sugai & Horner, 2006).

Secondary prevention (i.e., Tier II strategies and systems plus those encompassed within Tier I) is necessary for a smaller portion (approximately 15-20%) of students who exhibit at-risk behaviors (Sugai & Horner, 2006). These "supplemental support" (NCDPI, 2017a) elements at the secondary prevention level include direct instruction in such areas as social skills,

organizational strategies, self-management, behavioral contracts, and increased contact with and feedback from support personnel via such programs as Check In / Check Out (CICO; Campbell & Anderson, 2008; Todd et al., 2008), and increased collaboration between home and school. Systems features at this level include increased data collection to facilitate quick adjustments to strategies, a team dedicated to selecting and monitoring implementation of secondary prevention level strategies, and designated intervention coordinators (Horner et al., 2010). Positive Behavioral Interventions and Supports trainers emphasize that schools should not identify students as "Tier II kids" or "Tier III kids" (NCDPI, 2015a). Such conceptualizations tend to place students into fixed categories. Rather, PBIS proponents assert that schools should enact universal interventions (i.e., Tier I elements) for all students, and systematically employ more intensive and specialized interventions (i.e., Tier II, Tier III) in order for some students to be successful in school.

Tertiary prevention supports (i.e., Tier III) are for students (2-5% of the school population) whose high-risk problem behavior has not changed in response to the primary or secondary level interventions (Sugai & Horner, 2006). Tertiary prevention nearly always involves the implementation of functional behavioral assessment procedures. The results of such procedures guide development of an intensive, individually tailored behavior support plan. Such plans typically specify the instructional and reinforcement strategies that will be applied to teach specific alternative, prosocial behaviors to replace the targeted problem behavior. These plans also outline the strategies that will be enacted to prevent the problem behavior, and consequences to apply when the problem behavior occurs. Tier III supports (i.e., "intensive supports;" NCDPI 2017a) require personnel trained in functional behavioral assessment procedures, coordination of school and community-based stakeholders to achieve wraparound support, and frequent progress

monitoring to determine whether the behavioral interventions are implemented as intended and whether these interventions are producing desired effects (Horner et al., 2010).

Implementation of PBIS

As Sugai and Horner (2002) describe, there are a number of key steps to initial PBIS implementation. First, a school must establish a PBIS leadership team, which is typically comprised of the principal, several general and special education teachers, other administrators and specialty support staff, and parents and community stakeholders when feasible. The leadership team is responsible for reviewing school priorities, identifying outcomes and progress indicators, establishing staff and school improvement action plans, and organizing staff development activities. This team must secure broader staff agreement with the rationale behind PBIS, and some of the long- and short-term school-specific goals of the initiative. Of course, without staff buy-in, the school will not be able to effectively implement PBIS. In fact, Sugai and Horner assert that the leadership team should not attempt to implement action plans until they have secured agreement with at least 80% of the staff. The North Carolina Department of Public Instruction (2015a) emphasizes a range of strategies to obtain broader staff buy-in on PBIS implementation: (a) present existing data and trends, (b) identify benefits to teachers (e.g., conservation of time, increased instructional time), (c) employ staff surveys and use results to guide development, (d) provide incentives for participation, and (e) build consensus around common values.

Once staff buy-in is confirmed, data (e.g., suspension, expulsion, office disciplinary referral, tardiness, attendance, etc.) are collected and analyzed to determine which school practices need to be initiated, maintained, enhanced, or eliminated. Beyond these behavioral data, self-assessment inventories and surveys are also used to inform action plan objectives and

progress indicators. Teams evaluate existing practices and identify needed changes within the following four systems: school-wide systems, non-classroom systems (e.g., hallways, cafeteria, buses, recess), classroom systems, and individual student systems. The PBIS action plans encompass measurable outcomes, a 1-3 year timeline of events, specific activities that lead to measurable outcomes, staff development and training activities, and resource and support needs (Sugai & Horner, 2002).

After an action plan is developed, steps must be taken to assure that the plan can be implemented with fidelity. To enact and sustain PBIS, the school must have strong administrative leadership and support, staff buy-in, training and coaching to support staff in precise implementation of strategies, and ongoing reinforcement for staff implementation efforts and progress (Sugai & Horner, 2002). Finally, formative data-based monitoring is crucial in order to evaluate both implementation fidelity and progress toward~~s~~ stated objectives. Office disciplinary referrals (ODRs) provide one data source often used by schools to evaluate progress made toward~~s~~ behavioral outcomes. However, it is important that schools have established clear operational definitions of relevant rule-breaking behaviors to assure that the information obtained within ODRs is reliable and valid. In addition, schools must establish clear and efficient procedures for processing ODRs, entering data, and summarizing and analyzing data. The PBIS leadership team should schedule frequent meetings to review available data, make informed decisions, and subsequently share these findings with the broader school staff (Sugai & Horner, 2002).

Once PBIS is “launched” in a school, the work is far from done. Most schools progressively adopt features of PBIS, moving through “phases” of implementation (Bradshaw, Debnam, et al., 2009). A variety of tools are available to support schools in evaluating

implementation status and implementation fidelity. For example, the Benchmarks of Quality (BoQ; Kincaid et al., 2010) is a self-report tool organized around 10 critical domains that reflect the fundamental features of PBIS. The Team Implementation Checklist (TIC; Sugai, Todd, & Horner, 2001) is completed by the school's PBIS team members on a monthly or quarterly basis, and takes only 10-15 minutes to complete. The TIC supports team members in determining the status (e.g., achieved, in progress, or not started) of various PBIS components.

The School-wide Evaluation Tool (SET; Sugai, Lewis-Palmer, et al., 2001) is administered by a trained outside evaluator, and it supports schools in assessing critical elements of school-wide behavior support via evaluation of various permanent products, observations, interviews, and surveys. The SET includes seven subscales: behavior expectations defined, behavioral expectations taught, reward system, violation system, monitoring and evaluation, management, and district support. It is typically administered in the pre-implementation period, approximately six weeks after implementation, and then annually thereafter.

In North Carolina, schools also use the Tiered Fidelity Inventory (TFI; Algozzine et al., 2014), a self-assessment tool that supports school-based teams in evaluating the extent to which they are implementing core features of the school-wide PBIS model. Scores on each feature range from 0 (not implemented) to 2 (fully implemented). To some extent, the tool features items that assess efforts to obtain ongoing feedback from relevant stakeholders whose buy-in is critical (e.g., "Faculty are shown schoolwide data regularly and provide input on universal foundations [e.g., expectations, acknowledgements, definitions, consequences] at least every 12 months," TFI, p. 9).

In North Carolina, PBIS training is segmented into three sequential two-day modules (developed by the University of Missouri's Center for School-Wide PBIS) that are completed by

the school's PBIS leadership team as the school meets implementation criteria over several years (NCDPI, 2017a). The school team receives ongoing technical assistance, coaching, and professional development by the PBIS LEA Coordinator, the Regional Behavior Support Consultant, and the Department of Public Instruction's State Consultant.

According to the NCDPI (2017a), PBIS implementation requires an initial investment of time and effort from the school PBIS team and the entire school staff. Costs for the school PBIS team to attend training are typically limited to providing substitutes as needed for team members. Generally, PBIS teams attend training provided by the regional consultants within their district or region to minimize travel costs. Staff-wide training events often occur in school-based staff meetings or during school-based teacher workday sessions. Beyond training expenses, schools determine the amount needed to support implementation activities. The NCDPI reports that most schools also invest a few hundred dollars to post school rules and to accrue tangible items that would be included in a reinforcement system (e.g., treasure boxes, school or grade-level stores). In many cases, schools establish partnerships with local businesses, seek grant funding, request parent donations, and obtain the support of their parent-teacher associations to help finance PBIS efforts.

Horner et al. (2012) estimated that the initial two-year implementation costs of Tier I PBIS per school (in a cohort of 15 implementing schools within one mid-sized 30-50 school district) would be from approximately $5,000 (for personnel time, training fees, data system) to $10,000 (includes unforeseen costs such as staff turnover, changes in capacity). Horner et al. acknowledge that the cost of training and implementation support for Tier II and Tier III interventions such as function-based support, CICO, wraparound systems of support, and social skills interventions, can be considerable, and are more difficult to project. Costs vary for training

expertise, materials, staff time for training, and related assessment and data collection tools. However, Horner et al. note that these costs are largely dependent upon existing expertise within a school system, number of implementing schools, and the extent to which existing data systems are available to monitor fidelity and progress towards student-specific outcomes.

In many states, including North Carolina, the OSEP and IDEA serve as the primary source of funding for initial implementation of school-wide PBIS (Gage et al., 2014). Other sources of funding often include State Improvement Grants, federal grants, Title IV funds, and other sources contained within a state's budget.

Considerable federal, state and local resources are allocated to this multi-faceted initiative. There exists a fairly expansive body of research evaluating the effects of PBIS on various outcomes of social significance such as disciplinary exclusion rates (e.g., Bradshaw, Mitchell, & Leaf, 2010; Luiselli et al., 2005; Simonsen, et al., 2012; Vincent & Tobin, 2011); student attendance rates (e.g., Freeman et al., 2015; Freeman et al., 2016); and teachers' perceptions of the school environment (e.g., Bradshaw et al., 2008; Houchens et al., 2017). However, fewer studies have evaluated the effects of PBIS implementation on both disaggregated behavioral outcomes and organizational health outcomes. Within PBIS-implementing schools, it is necessary to further examine how implementation fidelity may differentially impact such outcomes. The ambitious aim of PBIS is to change an entire school culture in ways that enhance the teaching and learning environment for students and educators. Hence, it is important to further understand how implementation impacts critical stakeholder groups.

Purpose Statement

The purpose of the present study was to compare office disciplinary referral, crime and violence, and suspension data, as well as teacher turnover rates and results from the 2016 North Carolina Teacher Working Conditions survey, between PBIS and non-implementing public elementary schools across North Carolina within the 2015-16 school year. An additional objective was to evaluate these behavioral outcome data and measures of organizational health across the state's PBIS schools with Green Ribbon, Model, and Exemplar implementation designations during the 2015-16 school year.

Research Questions

The following research questions guided this study:

1. What is the relationship between PBIS implementation and organizational health outcome measures in elementary level schools during the 2015-16 school year?
    a. Are there significant differences in teacher satisfaction ratings in PBIS-implementing versus non-implementing schools, when other factors are also considered (e.g., school population size, concentration of economically disadvantaged students, suspension rates, crime and violence rates)?
    b. Are there significant differences in teacher turnover rates in PBIS-implementing versus non-implementing schools, when other factors are also considered (e.g., school population size, concentration of economically disadvantaged students, suspension rates, crime and violence rates)?
    c. Within PBIS-implementing schools, what is the relationship between higher implementation designation and teacher satisfaction ratings, when other factors

are also considered (e.g., school population size, concentration of economically disadvantaged students, suspension rates, crime and violence rates)?

d. Within PBIS-implementing schools, what is the relationship between higher implementation designation and teacher turnover rates, when other factors are also considered (e.g., school population size, concentration of economically disadvantaged students, suspension rates, crime and violence rates)?

2. What are the effects of PBIS implementation on behavioral outcomes in elementary level schools for the 15-16 school year?

   a. Are there significant differences in suspension and crime and violence rates in PBIS-implementing versus non-implementing schools?

   b. Are there significant differences in suspension and crime and violence rates across distinct ethno-racial categories in PBIS-implementing versus non-implementing schools?

   c. Are there significant differences in office disciplinary referral rates among students with disabilities in PBIS-implementing versus non-implementing schools?

   d. Do PBIS-implementing schools with higher implementation fidelity demonstrate significant differences in suspension and crime and violence rates when compared to schools with lower implementation fidelity levels?

   e. Do PBIS-implementing schools with higher implementation fidelity demonstrate significant differences in suspension and crime and violence rates across ethno-racial categories when compared to schools with lower implementation fidelity levels?

f. Do PBIS-implementing schools with higher implementation fidelity demonstrate significant differences in office disciplinary referral rates among students with disabilities when compared to schools with lower implementation fidelity levels?

Significance of the Study

Few studies have evaluated both disaggregated behavioral outcomes and indicators of organizational health (e.g., teacher turnover rates, teacher satisfaction survey results) in a large sample of PBIS versus non-implementing schools. This study also evaluated any significant differences in these behavioral outcomes and indicators of organizational health as PBIS was implemented with greater fidelity. Results of the present study can be used at the school, LEA, and state level to aid stakeholders in more fully understanding the impact that PBIS implementation, and level of fidelity, has on critical school outcome foci for both students and teachers. Results could guide development of new policies and practices at the local and state level that are intended to promote positive outcomes for all students. Results may support PBIS teams in designing school-based training and coaching models that aid teams in sustainable implementation.

Definition of Terms

For the purpose of the present study, the term *office disciplinary referral (ODR*) is defined as (a) an event in which a student engages in a problem behavior that violates a school rule, (b) the problem behavior was observed or identified by school staff, (c) the event resulted in a consequence issued by administrative staff, and (d) the administrator or other designee produced a permanent product (written or electronic) describing the entire event (Sugai, Sprague,

et al., 2000). In typical cases, the referral results in a student being sent to the office to meet with the administrator. The duration of such meetings and related out of classroom time is at the discretion of the administrator. The permanent product typically captures such information as student name, referring teacher or staff, time of day, location of behavior, and nature of problem behavior (Irvin et al., 2006). Within their reporting structure, many schools categorize each problem behavior episode as either a minor infraction or a major infraction (Gion et al., 2014). Examples of minor infractions might include brief and low-intensity defiance or non-compliance, or late arrival to class; examples of major infractions include fighting (mutual participation in a physically violent incident with another peer) and gang affiliation displays (Todd et al., 2010). The NCDPI (2017a) estimates that responding to and processing one office disciplinary referral event consumes approximately 40 minutes of a teacher's instructional time (refer to Appendix A for a detailed description of variables included within the study).

In 1993, the North Carolina General Assembly passed the Safe Schools Act requiring LEAs to report specified acts of crime and violence to the State Board of Education (NCDPI, 2017c). These specific acts include: (1) homicide, (2) assault resulting in serious bodily injury, (3) rape, (4) sexual offense, (5) sexual assault, (6) kidnapping, (7) robbery with a dangerous weapon, (8) taking indecent liberties with a minor, (9) bomb threat, (10) burning of a school building, (11) possession of alcoholic beverage, (12) possession of controlled substance in violation of law, (13) possession of a firearm or powerful explosive, (14) possession of a weapon, (15) assault involving the use of a weapon, and (16) assault on school personnel. The term *weapon* is defined as: "(1) any BB gun, (2) stun gun, (3) air rifle, (4) air pistol, (5) bowie knife, (6) dirk, (7) dagger, (8) slingshot, (9) leaded cane, (10) switchblade knife, (11) blackjack, (12) metallic knuckles, (13) razors and razor blades, (14) fireworks, or (15) any sharp-pointed or

edged instrument, except instructional supplies, unaltered nail files, clips, and tools used solely for preparation of food, instruction, maintenance" (NCDPI, 2018a). The term *assault on school personnel* is defined as an act or attempt by force or violence (not involving the use of a weapon) to do injury to a school official, employee, or volunteer that causes reasonable apprehension of immediate bodily harm while the school official, employee, or volunteer is performing their duties. This offense includes assaults on school personnel that do not result in apparent serious injury (NCDPI, 2018a).

The term *short-term suspension* is defined as the removal of a student from school for a period up to but not exceeding ten school days (NCDPI, 2017c). The school principal has the authority to impose short-term suspension on a student who "willfully engages in conduct that violates a provision of the Code of Student Conduct authorizing short-term suspension" (Department of Education, 2020, p. 17). Many local education agencies in North Carolina outline within their code of student conduct examples of behaviors that align with different violation levels (in terms of severity), and corresponding consequences. For instance, such student behaviors as "substantially disruptive behavior" (Wake County Board of Education, 2018, para. 34), "disrespect to a school staff" (Robeson County School System, 2012, p. 18), or "disruption of the learning environment" (New Hanover County Schools, 2018, p. 4), may result in short-term suspension. Principals usually determine the duration of suspensions, and whether to suspend a student in-school (in a designated suspension classroom) or short-term out-of-school. According to NCDPI (2017c), when a school does not have an in-school suspension program or when offenses are more serious or chronic, those offenses may be addressed through short-term, out-of-school suspensions.

Long-term suspensions last from eleven days up to the remainder of the school year. A principal can recommend to the superintendent the long-term suspension of any student who "willfully engages in conduct that violates a provision of the Code of Student Conduct that authorizes long term suspension" (Department of Education, 2020, p. 17). Local education agencies in North Carolina typically outline within their code of student conduct examples of behaviors that warrant long-term suspension. For example, such behaviors as "theft or vandalism that could result in a criminal charge if the student were an adult" (New Hanover County Schools, 2018, p. 21), or "assault on school personnel or other adult" (Wake County Board of Education, 2018, para. 60), may result in a long-term suspension. When a student is suspended long-term, the student may not return to his or her regular school for the duration of the suspension period. Some districts may permit long-term suspended students to attend an alternative learning program (ALP) or alternative school during their long-term suspensions. For reporting purposes, students are not considered suspended while attending an ALP or alternative school (NCDPI, 2017c).

For the purpose of the present study, students with disabilities are those evaluated and subsequently identified as having a disability and who, by reason of the disability, receive special education and related services under IDEA (NCDPI, 2020a). The fourteen recognized disability categories are as follows: (1) Developmental Delay, (2) Autism, (3) Deaf-blindness, (4) Deafness, (5) Emotional Disability, (6) Hearing Impairment, (7) Intellectual Disability, (8) Multiple Disabilities, (9) Orthopedic Impairment, (10) Other Health Impairment, (11) Specific Learning Disability, (12) Speech or Language Impairment, (13) Traumatic Brain Injury, and (14) Visual Impairment including Blindness.

In this study, non-disabled students are those who did not receive special education services under IDEA within the 2015-16 school year. Some students categorized as non-disabled may have disabilities, medical conditions or psychiatric diagnoses, but they are not currently receiving special education services under IDEA.

The term *economically disadvantaged student* refers to a student with a household income under 185 percent of the poverty threshold who is therefore eligible to receive free or reduced-price lunch (FRPL) under the National School Lunch Program (NSLP; Department of Education, 2015). Each year, the U.S. Department of Agriculture (2015) publishes annual adjustments to the income eligibility guidelines to be used in determining eligibility for free or reduced-price lunch or free milk. The percentage of students eligible for FRPL provides a proxy measure for the concentration of low-income students within a school (McFarland et al., 2018).

According to NCDPI (2015b), PBIS-implementing status is met when a school has a designated PBIS school-based coach, and when it has established an active PBIS team that meets at least nine times per year. In addition, a school must satisfy criteria including but not limited to the following: (1) achieve a score of 70% or better for Tier I on the TFI, (2) achieve a total score of 80 or better on the SET and (3) submit required data (i.e., ODR rates; TFI and SET or Brief School-wide Evaluation Tool [BSET]) to the DPI.

A non-implementing school is one that is not recognized by NCDPI as satisfying the implementation criteria as described above. Some schools receive PBIS training, but this does not by itself meet “implementing” status. Some non-implementing schools submit required data to be considered for implementing status, but the NCDPI denies the application for specific reasons (e.g., incomplete data; incomplete application).

The NCDPI (2016a) utilizes specific criteria to determine whether a PBIS-implementing school can achieve Exemplar, Model, or Green Ribbon implementation designations. In order for a school to be recognized as Exemplar, it must have completed all three training modules, and must achieve a SET score of 95% (or a BSET score of 85%), a score of 80% or better on Tiers I, II, and III of the TFI, and must be able to demonstrate an improvement trend across three consecutive years in behavioral and achievement indicators, as well as an improvement trend in at least one additional element (e.g., special education referral information; attendance; staff retention; NCDPI, 2017b).

In order for a school to be recognized as Model, it must have completed at least the first two training modules, and must achieve a SET score of 90% (or a BSET score of 80%), and a score of 75% or better on Tier I and 80% or better on Tier II of the TFI. In order for a school to be recognized as Green Ribbon, it must have completed at least the first training module, and must achieve the basic "implementing" criteria described above.

Chapter 2: Literature Review

This literature review focuses on the Positive Behavioral Interventions and Supports (PBIS) model applied within traditional kindergarten through 12th grade school contexts. The review begins with a synthesis of the evidence-based behavioral strategies that are fundamental within Tier I of the PBIS framework. To provide context for the evolution of PBIS, recent efforts to scale up and integrate traditionally isolated behavioral intervention strategies are briefly summarized. Studies evaluating the effects of school-wide PBIS on behavioral and organizational health outcomes of interest are explicated. Finally, factors that have been associated with PBIS adoption, implementation fidelity, and sustainability are addressed.

Evidence-Based Tier I Behavioral Interventions

Foundational to PBIS is the selection of empirically validated behavior support practices. As Horner et al. (2017) describe, "The core features of PBIS arose more as the distillation from a large body of research than as a unique, untested set of practices" (p. 27). In fields such as special education, considerable focus has been on producing evidence of efficacy and identifying criteria that these interventions must meet to be considered evidence-based (e.g., Council for Exceptional Children, 2014; Fixsen et al., 2013; Gersten et al., 2005; Horner et al., 2005; What Works Clearinghouse, n.d.). Evidence-based interventions are practices and programs shown by methodologically sound research conducted across a specified minimum number of independent studies to produce meaningful effects on student outcomes (Cook & Odom, 2013). Practices under consideration as evidence-based must feature replicable instructional or intervention procedures that are operationally defined (Cook & Cook, 2011).

A number of behavioral interventions have been endorsed as evidence-based by various research synthesis groups and professional organizations, (e.g., Council for Children with

Behavioral Disorders, 2020; National Professional Development Center on Autism Spectrum Disorders, n.d.) and subsequently by state education agencies and local education agencies (e.g., NCDPI, 2015c). In general, school-based behavioral interventions are designed to reduce internalizing and externalizing problem behaviors that interfere with academic and/or social performance, and to increase desirable alternative behaviors. As Landrum and Kauffman (2006) described, "Seeing how students' behavior is related to its context and the subjective aspects of experience, yet can be shaped by the astute application of behavior principles, has made a behavioral approach to classroom management more understandable and useful to educators" (p. 61).

Reinforcement

A central intervention within Tier I of the PBIS model is the intentional and frequent application of reinforcement to increase desirable student behaviors across the school context. Research suggests that when teachers encounter disruptive behavior, they react more negatively to the students with a history of problem behavior than those without that history (Nelson & Roberts, 2000). Furthermore, students with a history of problem behavior may encounter more social attention for negative behavior and less attention for positive behavior than their peers (Moore-Partin et al., 2010). Hence, one effort of PBIS is to alter these patterns by supporting teachers and other school staff in increasing the frequency and quality of positive interactions that they have with all students, including those who may have some history of problem behavior in the classroom.

Positive reinforcement is technically defined as the contingent presentation of a stimulus delivered upon a response that increases the future probability of that response (Alberto & Troutman, 2013). Reinforcement comes in many forms. Within the PBIS framework, school staff

deliver behavior-specific praise (e.g., "Great job *walking quietly to the gym.*") when they observe a student or group of students exhibiting an expected behavior. Praise is typically comprised of favorable verbal or non-verbal attention directed toward a behavior or characteristic exhibited by the targeted student(s) (Jenkins & Floress, 2015). In addition to behavior-specific praise, staff might also issue tokens or tickets that can be accumulated and eventually exchanged for items (e.g., snacks, toys, school supplies, or gift cards) at the school-wide or grade level rewards supply. Classrooms might also compete to accumulate the most points across a specified period of time; the winning class obtains some larger prize such as a pizza party or movie party. Some schools use web-based classroom management systems that allow immediate delivery of tokens or points to each student's account; students and parents/guardians can also access the account in real time to see daily and long-term point totals (Robacker et al., 2016).

Significant effects of positive reinforcement procedures on a wide array of behaviors have been documented in numerous studies within the fields of education and behavioral intervention (e.g., Alberto & Troutman, 2013; Landrum & Kauffman, 2006; Moore-Partin et al., 2010; Wong et al., 2015). Landrum and Kauffman assert that positive reinforcement offers the foundation of much of what is known about effective teaching and classroom management. Reinforcement can be applied in conjunction with other strategies and is integral to a number of empirically-based PBIS Tier II and Tier III interventions such as: social skills training (Council for Children with Behavioral Disorders, 2017; Gresham et al., 2006), Check In—Check Out (Wolfe et al., 2016), function-based behavior support plans (e.g., Council for Children with Behavioral Disorders, 2017; IDEA, 2004), behavior contracts (Council for Children with Behavioral Disorders, 2017), and peer-mediated strategies including peer tutoring, peer modeling, and peer management (Collins et al., 2018).

Research suggests that behavior-specific praise has a positive effect on on-task behavior of students in special education (e.g., Sutherland et al., 2000) and general education classrooms (e.g., Thompson et al., 2012). For example, Ferguson and Houghton (1992) found that general education teachers who received a specific training protocol for praise delivery increased their rate of praise, and the overwhelming majority of their students exhibited increases in levels of on-task behavior during the intervention period. Proponents suggest that a teacher's systematic application of praise and attention may be their most powerful motivational and classroom management strategy (Albert & Heward, 1997).

Interestingly, Brophy (1981) argued that teacher praise may certainly operate as a reinforcer when effectively applied (thereby producing desirable effects on targeted behavior), but it is often used non-strategically, infrequently, and issued non-contingently based on a teacher's perception of the student's needs (rather than contingent upon a behavior). While teacher praise has been studied for many decades, large gaps remain in the literature regarding typical rates of praise in the classroom, differential rates across grade levels, and differential effects of praise across grade levels (Jenkins et al., 2015). Furthermore, praise that specifies the observed behavior (e.g., "Good job raising your hand first.") is widely endorsed by scholars and practitioners over generic or vague praise (e.g., "Good job"). Yet Jenkins et al. assert that there has been limited empirical work that specifically evaluates the differential effects of these two forms of teacher praise on student behavior. In their recent review of the body of research on behavior-specific praise applied in K-12 contexts, Royer et al. (2018) found that the strategy could be categorized as a "potentially evidence-based practice" under the Council for Exceptional Children's (CEC) 2014 Standards for Evidence-Based Practices in Special Education.

Research has focused on evaluating teacher praise rates and in training teachers to increase their rate of behavior specific praise for alternative desirable behaviors in relation to the rate of corrective feedback issued for problem behavior (Jenkins & Floress, 2015; Reinke et al., 2013). When corrective feedback is required, research suggests that it should be brief, specific, and informative, and delivered in a non-emotional manner that redirects the student to a desired response (Gable et al., 2009). Proponents of PBIS draw from the body of research on teacher praise to endorse a minimal ratio of five praise statements for every one correction (Office of Special Education Programs (OSEP) Technical Assistance Center, 2021).

When a problem behavior is targeted for reduction via some form of punishment (e.g., temporary removal of social attention from the teacher; lost access to preferred school activities/items for a specified period), the differential reinforcement of an acceptable replacement behavior(s) is strongly encouraged (Alberto & Troutman, 2013; Landrum & Kauffman, 2006). A substantial body of research has demonstrated the effects of differential reinforcement procedures on a range of prosocial and academic behaviors (Martens & Lambert, 2014; O'Brien & Repp, 1990; Petscher et al., 2009; Wong et al., 2015). A practical behavior management approach is that educators should teach and reinforce desirable behaviors, while withholding reinforcement for problem behaviors.

Many educators apply group-oriented contingency systems (Litow & Pumroy, 1975) in effort to manage classroom behavior by offering reinforcement (and in some cases such forms of punishment as a response cost) to the group contingent on performance of specific behaviors. For example, interdependent group-oriented contingency systems such as the Good Behavior Game (Barrish et al., 1969) require that the collective group satisfy a specified criterion (e.g., everyone quietly lines up in all four group transitions) to access a reward, or to obtain specified points

towards a deferred reward. This strategy has been demonstrated to be effective in reducing targeted problem behavior and increasing desired behavior in a variety of educational contexts (Tingstrom et al., 2006), including high school (e.g., Flower et al., 2014; Mitchell et al., 2015) and post-secondary settings (Cheatham et al., 2017).

It is important to note that such strategies as the Good Behavior Game are typically applied during specific instructional periods and may not produce effects on behavior that are maintained outside of the "game time" or generalized to novel conditions or contexts (Donaldson et al., 2015). More broadly, researchers and educators who design and employ any specialized reinforcement procedures to address group or individual behavior must proactively incorporate strategies intended to maintain effects as the most intensive procedures fade (Alberto & Troutman, 2013). Of course, if behavior change is not lasting, questions justifiably arise regarding social validity. Practicality and sustainability of such procedures are vital considerations. It can be difficult to obtain educator buy-in for reinforcement procedures in part because many can enact efficient, simple punishment procedures that produce rapid (though often temporary) effects (Maag, 2002). Some public figures (e.g., Kohn, 1993) and educators wholly resist the use of reinforcement or deny its effect (Axelrod, 1996; Fielding et al., 2013; Landrum & Kauffman, 2006). Reinforcement is viewed by some as coercive (more so than punishment) and synonymous with bribery, thereby undermining student self-determination (Maag, 2002).

Some argue that the application of extrinsic rewards undermines intrinsic motivation (e.g., Deci et al., 1999) and may stifle students' experimentation and creativity (Tegano et al., 1991). There has been considerable debate (e.g., Cameron & Pierce, 1996; Deci et al., 2001; Lepper et al., 1996) regarding the use of extrinsic rewards (i.e., external and typically

programmed environmental consequences such as verbal praise, access to tangibles) and their potential impact on intrinsically motivated behavior (i.e., behavior not under external control where engagement in or completion of the activity itself provides the incentive). In their two primary meta-analyses, Deci et al. (1999) evaluated the effects of extrinsic rewards (verbal and tangible rewards) on intrinsic motivation, measured by free choice task persistence (time on task) and self-report of interest. They further analyzed the studies that included tangible rewards and calculated separate effect sizes for unexpected versus expected rewards. The studies including expected tangible rewards were separated into the following categories: task noncontingent rewards, engagement contingent rewards, completion contingent rewards, and performance quality contingent rewards. The authors concluded that intrinsic motivation was undermined by every reward category except unexpected rewards and verbal rewards.

In response to these findings, Cameron et al. (2001), conducted a meta-analysis evaluating the effects of extrinsic rewards on intrinsic motivation, measured by free choice task persistence and self-report of interest. They divided the included studies into the primary categories of high and low initial task interest. They further classified the studies according to reward type (verbal or tangible), reward expectancy (expected or unexpected), and reward contingency (7 types; e.g., engagement contingent rewards, completion contingent rewards, and performance quality contingent rewards). Their results suggested that in general, rewards were not harmful to motivation to perform a task. Specifically, they found that rewards given for low-interest tasks enhanced free-choice intrinsic motivation. On high-interest tasks, verbal rewards produced positive effects on free-choice motivation and self-reported task interest. Cameron et al. found negative effects on high-interest tasks when the rewards were tangible, expected, and loosely tied to level of performance. However, when rewards were linked to level of

performance, measures of intrinsic motivation increased or did not differ significantly from the nonrewarded control group. Overall, the pattern of results indicated that reward contingencies did not have negative effects on intrinsic motivation as previously suggested (cf. Deci et al., 1999).

Though debate on this issue persists, there are broad conclusions that may guide development of effective and appropriate reinforcement programs in schools (including those implementing PBIS). For example, rewards should not be presented for mere participation in a task without regard for completion or quality (Akin-Little et al., 2004). Offering tangible rewards for engagement in or completion of high interest tasks may have a negative effect on motivation (Cameron et al., 2001). In addition, as noted by Akin-Little et al., the student's baseline performance must be taken into account, because students who are already performing a task at or above targeted levels likely do not require programmed reinforcement as a consequence for completion of that task. Conversely, students whose baseline performance is far below expected levels will likely require systematic procedures with initially dense reinforcement, where the criterion changes and reinforcement are carefully thinned over time, as the behavior improves (Akin-Little et al., 2004). Effective reinforcement requires strategic planning of such aspects as nature and type of reinforer(s) used, frequency of delivery, and timing of delivery.

In essence, the indiscriminate use of rewards without consideration of the student(s), the targeted behaviors, and the overall context is ill-advised and may have negative effects on intrinsic motivation and subsequent performance. On the other hand, when thoughtfully and systematically designed, reinforcement has been shown to produce positive effects on student behavior.

Explicit Instruction of Expected Behaviors

In addition to reinforcement-based interventions, another fundamental strategy within Tier I of the PBIS model is the proactive identification, posting, and explicit teaching of 3-5 positively framed behavioral expectations that are generalizable across the entire school context (OSEP Technical Assistance Center, 2021). Research suggests that these behavioral expectations should be (a) age appropriate in terms of language and expectations; (b) specific and observable; (c) stated positively, describing what students should do rather than what they should not do; and (d) enforceable (Reinke et al., 2013).

Emmer and Evertson (1982) found that teachers who were rated on observational scales as more effective classroom managers explicitly taught rules to students and provided written copies of the stated rules more often than those teachers who were rated as less effective. Teaching these rules via modeling of examples and nonexamples of the expected behavior can also be effective for some students (Gable et al., 2009). In general, the combination of posting, teaching, and reviewing expectations, and then providing feedback and reinforcement, have been associated with decreases in off-task behavior and disruptive talking out (Nelson et al., 1996), and increases in academic engagement and conflict resolution (Simonsen et al., 2008).

Active Supervision

Active student supervision is another core PBIS Tier I support to be applied in conjunction with reinforcement procedures and explicit instruction of 3-5 specific behavioral expectations. Active supervision involves ongoing monitoring, movement and proximity of the teacher, and anticipatory prompting (i.e., precorrection) to promote expected behavior (OSEP Technical Assistance Center, 2021). Precorrection involves prompting the student(s) to engage in a prosocial or more appropriate behavior before the problem behavior occurs (Colvin et al., 1997;

Colvin et al., 1993). These prompts might come in the form of non-verbal cues (e.g., gesture or model), or verbal review of expected behaviors (e.g., “Please remember that you should remain quiet during this quiz.”). Precorrection may also include guided rehearsal in advance of expected performance. Ennis et al. (2017) applied the CEC’s 2014 Standards for Evidence Based Practices in Special Education to the body of research on precorrection. They found that the strategy met the criteria when an 80% weighted criterion was used (i.e., four included studies did not meet all methodological quality indicators but were rigorous and of high-quality; three included studies met all criteria).

Active supervision and precorrection have been shown to be effective in the classroom, during hallway transitions, and other school contexts (Simonsen et al., 2008). For example, DePry and Sugai (2002) found that active supervision, precorrection, and daily data review produced a decrease in the occurrence of minor behavioral incidents (e.g., eating in the classroom, passing notes, out of seat behavior) in a sixth-grade general education classroom. Research has also demonstrated reductions in problem behavior among elementary students during three key transition points: (a) coming into the school building at the start of the school day, (b) leaving class and going to the cafeteria, and (c) leaving class and exiting the building at the end of the school day (Colvin et al., 1997). Interestingly, results of the study by Colvin et al. indicated that the degree of active supervision behaviors (movement, monitoring, proximity) was more important than the number of supervisors. In particular, the number of supervisory staff (which varied from one to six) was unrelated to changes in student transition behaviors in their study.

Frequent and Varied Response Opportunities

Active student engagement is a widely adopted strategy that typically involves providing frequent opportunities for student response and varied response opportunities in the context of instruction (Simonsen et al., 2008). Two common methods used to increase the rate of responding in a classroom include choral responding and response cards (Simonsen et al., 2008). Proponents of PBIS outline three primary strategies for increasing student response rates and varying opportunities during instruction: (a) individual or small group questioning in which all students are given ample opportunities to respond; (b) choral responding; and (c) nonverbal responses including response cards, hands up, guided notes, and student response systems via clickers and other tools (OSEP Technical Assistance Center, 2021). Logically, when students are actively involved in instruction, it is less likely that they are simultaneously engaged in problem behavior (Reinke et al., 2013). Research suggests that an increase in response opportunity rates is associated with improved academic and social behaviors (Moore-Partin et al., 2010).

In a study conducted by Sutherland et al. (2003), students demonstrated fewer disruptions, more correct responses, and increased task engagement during the intervention phases (i.e., when response opportunities were increased), in comparison to the baseline and treatment withdrawal phases. Interestingly, the authors noted that a combination of factors, including but not limited to increased response opportunities, teacher-delivered praise, and students' correct responses, may have produced the positive effects on behavior. As students were given more opportunities to respond, rates of correct responding also increased, which in turn, influenced the rate of teacher-delivered praise for correct responses.

In summary, Tier I behavioral strategies are logically applied in tandem with one another, rather than as isolated teacher tools. Overall, there is substantial evidence supporting the efficacy

of these interventions (e.g., Simonsen et al., 2008.). When applied in an integrated manner, they represent a basic prevention package that has relevance across a range of school contexts. Effective instruction of three-five specific and observable behavioral expectations hinges in part on reinforcement procedures to increase those expected behaviors. Furthermore, active supervision and student engagement are necessary to occasion opportunities to effectively deliver reinforcement, to promote and sustain desirable behavior, and to proactively redirect problem behavior.

Scaling Up Interventions

One key premise is that PBIS is not something that educators should "do" at certain times or with certain students. Proponents of PBIS assert that the model cannot be viewed or treated as "one more thing" (Sailor et al., 2009, p. 634) to add on to an already overloaded toolbox of flashy and sometimes fleeting educational strategies or programs. Rather, the effort is for an entire school culture to be wholly and continuously engaged in a set of integrated practices and systems that comprise the multi-tiered framework. As the field of behavior analysis evolved, an increasing number of scholars (e.g., Axelrod, 1993; Baer et al., 1987; Critchfield, 2015) have posited how behavioral principles and strategies might fit (or need to be repositioned) in relation to a "bigger picture." Such aims are intended to promote and improve dissemination, adoption, and sustainability. Effective intervention in the "real world" often requires full acknowledgement of dynamic interaction contexts in which many multiply controlled problem behaviors occur. Thus, PBIS developed as an effort to take behavioral strategies and processes beyond the behavior of the individual and toward the collective behaviors and systems of the entire school (Sugai & Horner, 2002). In fact, as Horner et al. (2010) described, the conceptual model for PBIS integrates the (a) principles of applied behavior analysis, (b) multi-tiered prevention logic

from community health, (c) rigorous use of universal screening and progress monitoring, (d) integration of behavioral and education practices for improving behavior and learning, and (e) the implementation technology required to apply effective practices at large scales.

Until recent years, the majority of behavior intervention research has focused primarily on defining criteria for evidence-based practices (e.g., Council for Exceptional Children, 2014), and on empirically evaluating the effectiveness of a particular strategy (e.g., Thompson et al., 2012) or isolated intervention package (e.g., Flower et al., 2014) for individual students or particular sub-groups. Yet overall, the chasm between empirical research and implementation in applied settings persists (Cook & Cook, 2011; Cook & Odom, 2013). Many educators struggle to enact or to sustain implementation of evidence-based practices, including those that fall under the category of behavioral interventions (Odom, 2009). When these evidence-based practices are presented and implemented at particular sites (e.g., schools, classrooms), there is often inadequate focus on the range of systems variables and implementation tools required to facilitate adoption, reliable use, sustainability over time, and generalization across settings (Horner et al., 2017). In response to these issues, there has been increased emphasis on identifying and employing the elements and processes that are fundamental to sustainable implementation at large scales.

Scaling Up Defined

Scaling up is the process of moving from “exemplars” (i.e., demonstration districts, schools, or classrooms) supported by “ghost systems” (i.e., external, extraordinary supports and consultation) to the “typical” supported by the “host system” (i.e., districts and states) that is designed to make full, effective, and sustained use of innovations (Fixsen et al., 2009, p. 2). Fixsen et al. asserted that host systems reach the “tipping point” for scaling when at least 60% of

the schools within a system are implementing a practice with fidelity and with demonstrable benefit to the students (p. 2). Districts act as a critical leverage point for producing student outcomes, as the development of sustained organizational supports is central to scaling evidence-based interventions (Horner et al., 2017). Without this district-level support, implementation of a practice at the school and classroom level (regardless of its status as evidence-based) is vulnerable to major fidelity violations if not eventual abandonment (Cook & Odom, 2013; Durlak & DuPre, 2008; Forman et al., 2009; Han, & Weiss, 2005). Some educators justifiably express concerns about the feasibility of implementation given potentially impeding factors such as limited time, material resources, personnel, inconsistent administrative support, and non-existent or insufficient coaching (e.g., Boardman et al., 2005; Han, & Weiss, 2005; Klingner et al., 2003). When organizational support is absent or insufficient, adoption at the classroom and school level alone can be swiftly neutralized by a change in principal leadership, funding, or improvement priority (Horner et al., 2017).

Implementation Science

Many scholars and practitioners engaged in efforts to scale up interventions have drawn from the implementation science literature (e.g., Fixsen et al., 2005; Nordstrum et al., 2017; Rogers, 2003) to develop guidelines and support structures. Implementation is defined as a specified set of activities enacted to put into practice an activity or program of known dimensions (Fixsen et al., 2005). This science of implementation has developed from a recognition that (a) implementation is a process that evolves in stages including but not limited to adoption, initial and full implementation, and sustainability (Cook & Odom, 2013; Fixsen et al., 2005; Rogers, 2003), (b) there is a temporal order to the critical steps of quality implementation (Meyers et al., 2012), and (c) the beliefs, values, and behaviors of practitioners directly affect many aspects of

implementation including adoption, fidelity and sustainability (Cook & Odom, 2013; Durlak & DuPre, 2008; Forman et al., 2009; Han & Weiss, 2005).

In order for a practice to be adopted and sustained, it must be devised and structured with sufficient flexibility such that it can be adapted to changing conditions, and furthermore, educators must understand the program well enough so that they are able to modify it without forfeiting the core principles and central intervention techniques (Han & Weiss, 2005). In essence, educators are not required to follow a rigid implementation script, but they do require clear parameters to assure that the core functions of the evidence-based practice are implemented (Fixsen et al., 2009).

Researchers and practitioners invested in full and sustained implementation of evidence-based practices recognize that a range of interconnected variables influence the extent to which a practice is adopted, implemented as intended, and maintained long term. The literature suggests that implementation fidelity is influenced by a number of intervention features (e.g., complexity, adaptability and flexibility, and compatibility with intervention context), personal factors (e.g., characteristics, beliefs, attitudes of implementers), and organizational context factors including attitudes, beliefs, behaviors of administrators, as well as the organizational policies, structures, and procedures (Cook & Odom, 2013; Durlak & DuPre, 2008; Forman et al., 2009; Han, & Weiss, 2005).

Scaling Up PBIS

Developers of the PBIS initiative have drawn from the broader implementation literature base in order to develop a framework that identifies the implementation phases: (a) exploration, (b) installation, (c) initial implementation, (d) full implementation, and (e) sustainability and scaling up (Sugai, O'Keefe, et al., 2012). The PBIS implementation process is facilitated by a

leadership structure that establishes four drivers that assure adequate capacity of the host system at every phase: (a) coaching, (b) training or professional development, (c), evaluation, and (d) expertise in behavioral strategies. To sustain these drivers, the leadership team utilizes a dynamic, data-based action plan that considers implementation and outcome visibility, long-term funding, political supports, and policy. Beyond these capacity-building elements, there are additional features necessary to enhance PBIS large-scale adoption, implementation fidelity and sustainability. One primary consideration is to identify the valued outcomes that the initiative is intended to provide (i.e., reduce problem behavior in schools and increase appropriate social and academic outcomes for all students; Horner et al., 2017). In other words, it is important to begin with the end in mind to assist stakeholders (e.g., implementers, policy-makers, community members) in clearly linking the PBIS strategies and systems to socially significant objectives.

Furthermore, in order to be fully adopted and to produce desired outcomes for all, the PBIS practices and organizational systems must be contextually and culturally responsive, meaning that (a) educators hold high expectations for all students, (b) students' cultures and experiences are authentically and frequently reflected in all aspects of the learning experience, (c) all students have access to effective instruction and adequate resources for learning, and (d) diverse community members are meaningfully engaged in ways that inform the schools' identification of goals and enactment of strategies (Sugai, O'Keefe, et al., 2012). To promote a strong contextual fit, new practices should be combined with already established and effective programs in ways that support efficient, sustainable implementation (Horner et al., 2017).

A commitment to fidelity is another fundamental feature of PBIS implementation, with the understanding that full implementation is an iterative and stage-based process (Horner et al., 2017). Horner et al. also emphasized that effective, large-scale implementation requires ongoing

collection and use of data at the organizational, building, classroom, and individual student levels. These data support teams in determining whether and how well (a) core practice features, and (b) core organizational system features are being implemented, and (c) valued outcomes are being achieved.

In summary, the field of behavioral intervention has evolved in its analysis of the "what" of evidence-based practices with isolated studies that demonstrate efficacy (Horner et al., 2010) to encompass a fuller appreciation and explication of the "how" of integrated evidence-based practices in applied contexts. The developers of PBIS assert a commitment to providing schools and districts with the tools necessary to achieve and sustain implementation of integrated evidence-based practices and support structures at scales of social importance (OSEP Technical Assistance Center, 2021).

Effects of PBIS on Behavioral Outcomes

Documenting the evidence base for school-wide PBIS is no simple task, as it represents a complex constellation of systems and practices implemented at three tiers of intensity (Horner et al., 2010). A fundamental objective of PBIS is the reduction of problem behaviors that typically lead to office discipline referrals (ODRs), suspensions, and diminished academic engagement. Hence, a number of studies have evaluated the effects of Tier I elements on student behavioral outcomes such as ODRs and suspension rates (e.g., Bradshaw, Mitchell, & Leaf, 2010; Bradshaw et al., 2012; Caldarella et al., 2011; Luiselli et al., 2005; McCrary et al., 2012; Simonsen et al., 2012; Vincent & Tobin, 2011).

Bradshaw, Mitchell, and Leaf (2010) analyzed data from a 5-year longitudinal randomized controlled trial of school-wide PBIS conducted in 37 elementary schools on student

suspensions, office discipline referrals, and academic achievement. The schools were matched on select baseline demographics (e.g., percentage of students receiving free or reduced meals), and 21 schools were randomized to the intervention condition with 16 assigned to the comparison sample. The schools trained in PBIS reported a significant reduction in the overall ODR rate, as well as the percentage of students who received an ODR. Bradshaw et al. cautioned that interpretation of these findings is hindered by the lack of both baseline ODR data and ODR data from the nontrained schools. To account for the small effect size for ODRs, the authors posited that the schools trained in PBIS tended to have fewer major ODRs per 100 students per school day relative to the national data for elementary schools. Hence, this sample of schools had little chance for significant reduction in ODRs. The schools trained in PBIS demonstrated a significant reduction in the rates of suspensions, whereas the rate in the nontrained schools remained unchanged during the study period. Annual assessments of PBIS implementation were employed in all PBIS trained schools using the School-wide Evaluation Tool (SET; Sugai, Lewis-Palmer, et al., 2001) and staff self-reports; both measures indicated that all PBIS schools reached and maintained high-fidelity implementation by the end of the study.

In a related study, Bradshaw et al. (2012) used data from the aforementioned randomized controlled trial to examine the effects of PBIS on teacher ratings of student behavior problems and social-emotional adjustment at five points across a four-year period for 12,344 elementary students. Their multilevel analyses indicated a significant positive effect of PBIS intervention on disruptive behaviors compared with those in the control schools. A similar effect was observed on concentration problems, and students in PBIS schools had higher levels of positive behaviors and emotional regulation compared with those in the control sample. In addition, Bradshaw et al.

found that students were 33% less likely to receive an ODR in PBIS schools than those in the comparison schools.

Simonsen et al. (2012) analyzed the effects of implementing PBIS, with and without fidelity across time, on student behavior (ODRs and suspensions) for a sample of 428 Illinois schools (including 91 middle and 17 high schools) implementing PBIS. Results indicated that fidelity (as measured by the SET) increased throughout the study, from 36% to 78% of schools. By the end of the study, elementary schools implemented the PBIS framework with the highest rate of fidelity at 81%. Simonsen et al. found that regardless of fidelity, schools experienced decreases in ODRs over time, but time alone did not lead to overall decreases in suspensions. To account for this effect, Simonsen et al. asserted that student behaviors that resulted in an ODR may be sensitive to change and responsive to PBIS implementation—regardless of fidelity—while student behaviors that resulted in suspension were obdurate with lower PBIS fidelity and only appeared to decrease with higher levels of fidelity. A noteworthy limitation is that the study included available data from 2000 to 2008, but not all schools had data for all years, and for many schools in the sample, data were only available for 1 or 2 years. Hence, as Simonsen et al. acknowledged, one cannot assume the same effect of year and SET fidelity on outcome measures for every school, and it is feasible that schools changed at different rates over time, and benefited differently from fidelity to PBIS across years.

Interestingly, in a longitudinal study using four years of data from 1,122 elementary, middle, and high schools (Childs et al., 2016), implementation of PBIS was linked with an immediate decline in ODRs and corresponding suspension events across all school types, and the decline was sustained over time. Yet, ODRs in schools implementing with greater fidelity (as measured by the Benchmarks of Quality [BoQ]; Kincaid et al., 2010) did not decrease at a faster

rate than schools implementing with less fidelity. Specifically, while results revealed a difference between discipline incidents each year with high-fidelity schools consistently lower on average than low-fidelity schools, the slope of change was the same for both school types across four years. In light of these findings, Childs et al. posited that implementation fidelity may reliably produce the most significant decreases in discipline outcomes during early stages of implementation, but fidelity alone is not sufficient to produce changes in the slope or rate of change for PBIS schools.

Barrett et al. (2008) provided an analysis of outcomes for one year where 33% of schools within the state of Maryland were PBIS-implementing. They reported 43% fewer ODRs per 100 students per school day in all Maryland elementary schools, when compared with the national average. Maryland middle schools reported 33% fewer ODRs overall, and high schools reported 37% fewer ODRs. The authors asserted that these results suggest that schools in Maryland experienced lower than typical rates of ODRs across all grade levels, which was likely due to the high level of fidelity of PBIS in Maryland schools, (as indicated by SET scores). Barrett et al. also conducted a repeated measures analysis comparing preintervention and postintervention suspension rates across one year for the elementary (n = 31) and middle (n = 31) schools that were trained in PBIS. Their analysis indicated a significant reduction in suspension rates following training in PBIS. However, the Barrett et al. study included data from only one year and it is unclear how the ODR rates might have changed across years. Furthermore, they did not address whether differential fidelity implementation levels may differentially impact the outcomes.

Luiselli et al. (2005) employed a quasi-experimental longitudinal study evaluating the effects of PBIS on office disciplinary referral rates and suspension rates in one large, urban

elementary school. They found that the average office discipline referrals rates (per 100 students) were 1.3 per day in preintervention (4 school months), .73 in intervention (1.5 school years), and .54 in follow-up (school year 3). Suspension rates were .31 per day in preintervention, .25 in intervention, and .20 in follow-up. It is noteworthy that these effects were maintained in the follow-up phase, when external consultation (i.e., support from the researchers) was no longer provided. Despite overall decreases in the average rates of ODRs and suspensions, ODRs on average decreased with greater magnitude than suspensions, and both outcome variables demonstrated gradually increasing trends during the second school year. As Luiselli et al. acknowledged, this effect was particularly noteworthy for suspensions, which were issued more frequently each month, reaching the highest level in the final month of the school year. It is also important to note that the suspension and ODR data in the preintervention phase were variable, with ODR data suggesting a decreasing trend just prior to intervention. The authors obtained social validity measures, with teachers reporting that the intervention was effective and contributed to better learning in classrooms. Implementation fidelity measures, however, were not reported.

In their longitudinal, quasi-experimental (non-equivalent two-group, pretest-posttest) design with an untreated control group, Caldarella et al. (2011) evaluated the effects of PBIS implementation across four years on ODR rates, unexcused absences, and tardiness among middle school students. Results revealed a statistically significant downward trend for ODRs, tardiness, and unexcused absences in the treatment school, yet effect sizes tended to be small. The control school also showed a statistically significant downward trend in ODRs and tardiness, though the slope of the change was not as steep as in the treatment school, resulting in a statistically significant interaction effect. A statistically significant interaction effect was also

indicated for unexcused absences, suggesting that the treatment school showed improvements while unexcused absences tended to increase within the control school. It is noteworthy that while the two schools were similar in demographic make-up, the treatment school had lower ratings of school climate and higher levels of ODRs in the pre-intervention phase.

Kelm et al. (2014) conducted a case study of a small elementary school that was partially implementing PBIS in year one of the study and fully implementing in year two (as measured by the SET). Results indicated that ODRs were reduced by more than half across two years. The authors emphasized that this reduction represented the restitution of 3,990 minutes of staff time and 7,980 minutes of instructional time for students. In addition, out-of- school suspensions were reduced by half between across both years. These results suggest that increased implementation fidelity is associated with enhanced outcomes.

A number of studies (e.g., Kelm et al., 2014; Luiselli et al., 2005) have demonstrated reductions in ODR and suspension rates over time (and as implementation fidelity increases). However, Warren et al. (2003) observed that the number of ODRs and suspensions in a middle school decreased significantly during the first year of implementation when direct intervention was delivered by the researchers, but these rates rose to a level that even exceeded baseline during the second and third years when only consultative support was delivered by the researchers. Warren et al. concluded that sustainability cannot be achieved without a diverse leadership team committed to withstanding the inevitable changes in personnel and programs, strong connections to the community, integrated service provision in schools, substantial district support, and family engagement.

Effects in High Schools

While there is a paucity of research specifically examining the effects of PBIS on behavioral outcomes in high schools, available evidence indicates a positive relationship between PBIS initiatives implemented with fidelity and increases in attendance rates (Freeman et al., 2015; Freeman et al., 2016), and a reduction in ODR rates (Flannery et al., 2014; Freeman et al., 2016).

The quasi-experimental study conducted by Freeman et al. (2015) included 883 high schools across 37 states that were implementing PBIS. In particular, high schools implementing PBIS with greater fidelity (as measured by the BoQ and SET) demonstrated higher attendance rates than those implemented with lower fidelity. It is noteworthy that while PBIS had positive effects on attendance at the high school level (a proximal and statistically significant indicator of high school dropout risk), PBIS had no direct short-term effect on dropout rates in the Freeman et al. study. Freeman et al. concluded that statistically significant changes in dropout rates may take longer to appear than other outcomes such as school attendance.

Flannery et al. (2014) evaluated the effects of school-wide PBIS across 3 years by comparing ODR rates in 12 high schools (with non-random assignment to treatment versus comparison groups). Their results demonstrated statistically significant decreases in ODRs in PBIS schools, with increases in comparison schools, when controlling for enrollment and percent of students receiving free or reduced-price meals. In follow-up analyses, Flannery et al. found a statistically significant inverse relation between strength of PBIS implementation fidelity and student problem behavior. Flannery et al. noted that the unique characteristics of high schools (e.g., larger student body and faculty; faculty's expectations for adolescent behavior; careful

planning to identify age-appropriate reinforcers and teaching strategies) require special consideration across formal professional development efforts and ongoing technical assistance.

In their longitudinal case study evaluating the effects of PBIS within one large, urban high school, Bohannon et al. (2006) found that the average ODR rate per day decreased by 20% from year two of PBIS implementation to year three. Interestingly, acts of disobedience of authority (one ODR behavioral classification area) decreased from 1.64 per every 100 students per day, to 0.05 per every 100 students in year three. However, one primary consideration is that this school did not reach full schoolwide implementation status (i.e., meeting at least 80% on all SET domains) across the study. Like many engaged in the evaluation of PBIS implementation at the high school level (e.g., Flannery et al., 2013; Flannery et al., 2014; Swain-Bradway et al., 2015), Bohannon et al. also cautioned that a variety factors specific to the high school context must be considered when designing and implementing PBIS. In particular, it can be more challenging to secure broad implementation in a large high school context due to physical and logistical issues. Therefore, it may be more difficult to gain consensus among high school staff on school policies, behavioral expectations, and related procedures (Bohannon et al., 2006).

In their systematic literature review, Lane et al. (2006) evaluated 14 PBIS studies representing a total of 63 secondary schools. They examined school-wide PBIS in relation to school characteristics, intervention focus and elements, research design, reliability and validity issues, and intervention outcomes. The researchers identified a number of methodological limitations including limited demographic information, insufficient descriptions of the interventions that would facilitate study replication, primarily descriptive research designs, and lack of reliability and validity information on the outcome measures. Lane et al. concluded that while many of the included PBIS studies produced desirable outcomes, numerous

methodological limitations limit the ability to draw accurate conclusions about intervention outcomes.

In summary, the limited body of research specifically examining the effects of PBIS on behavioral outcomes in high schools suggests that PBIS can have positive effects on ODRs and suspension and attendance rates, but full implementation may be more difficult to achieve and may take longer. Furthermore, implementation fidelity may differentially affect these outcome areas. Broad findings from this particular literature base must be interpreted with caution given that very few studies (e.g., Flannery et al., 2014) included comparison schools. A number of concerns regarding methodology further complicate interpretation of findings (see Lane et al., 2006). Future research is clearly required to resolve these issues, and to further examine the contextual influences specific to the high school environment that may impact implementation fidelity and intended outcomes.

Summary of effects on aggregate behavioral outcomes in K-12 settings. Gage et al. (2018) conducted a meta-analysis of the group-based experimental research on the aggregate effects of PBIS implementation to reduce disciplinary exclusion, including ODRs and suspensions. Application of their inclusion criteria resulted in identification of four group experimental studies that evaluated the effects of PBIS on disciplinary exclusion rates: two randomized controlled trials (Algozzine et al., 2012; Bradshaw, Mitchell, & Leaf, 2010) and two quasi-experimental designs (Flannery et al., 2014; Nelson et al., 2002). A total of 90 elementary and high schools were included in the analysis. Although there was no treatment effect for ODRs, there was a substantial and statistically significant effect on school suspensions. For interpretation, Gage et al. converted the effect size to the What Works Clearinghouse (2016) improvement index and found that schools implementing PBIS decreased suspensions by 16

percentiles compared with those schools not implementing PBIS. In an effort to account for the non-significant effect for ODRs, Gage et al. asserted the possibility that ODR rates may have actually increased in newly implementing schools, as all educators are trained to use ODRs in a consistent and systematic manner. In summary, Gage et al. asserted that the most striking finding from their broader review of the literature is that few studies evaluating the effects of PBIS on disciplinary exclusion included a comparison group.

Chitiyo et al. (2012) applied the criteria outlined by Horner et al. (2010) to 10 experimental studies included in their analysis of PBIS efficacy. The criteria were as follows: (a) the practice and participants are defined with operational precision, (b) the research employs valid and reliable measures, (c) the research is grounded in rigorous designs, (d) the research documents experimental effects without iatrogenic outcomes, and (e) the research documents effects (Horner et al.). Results of their systematic review suggested that only two of the 10 studies (i.e. Bradshaw, Mitchell, & Leaf, 2010; Horner et al., 2009) met all five criteria for the evidence base for PBIS. Chitiyo et al. concluded that these two studies defined the practice and participants with operational precision to allow replication, employed valid and reliable measures of student outcomes and treatment fidelity, documented experimental effects without iatrogenic effects, demonstrated high implementation fidelity (at least 80%), sustained administrative support (i.e., administration involvement in data collection and decision-making, as well as social validity measures reporting administrative satisfaction), and were grounded in rigorous research designs. Chitiyo et al. noted that while few studies met all criteria, the majority of studies documented experimental outcomes without any negative effects, and employed valid and reliable measures of both the outcomes and treatment integrity.

Overall, PBIS has been linked with reductions in aggregate rates of ODRs, suspensions, and increases in attendance rates. In some studies, PBIS implementation was linked with an immediate reductive effect on ODR rates (e.g., Childs et al., 2016). In other studies, ODR reductions were not realized in the early phase of implementation (e.g., Flannery et al., 2014). While results of some studies suggested that high implementation fidelity is linked with better behavioral outcomes (e.g., Flannery et al., 2014; Freeman et al., 2015; Kelm et al., 2014), these findings were not fully supported in other research (e.g., Bohannon et al., 2006; Simonsen et al., 2012). Though PBIS is a promising approach, future research on the model should carefully consider the rigor of the research designs employed (Chitiyo et al., 2012; Gage et al, 2018), treatment integrity, and operational definitions of participants and interventions (Chitiyo et al., 2012). Such efforts may strengthen the evidence-base supporting PBIS at a large scale. Further analysis of the differential effects of implementation fidelity on outcomes of interest may provide additional insight into the malleability of these outcomes at different points, and the factors that impact fidelity at different points in time. Beyond these aforementioned issues, there are other important areas that require more extensive consideration and analysis.

Effects on Disciplinary Equity

In the PBIS literature, there is a growing emphasis on evaluating the effects of this intervention framework on disciplinary equity across ethno-racial categories. Unfortunately, a number of studies show that students of color, and most significantly Black students, continue to encounter disproportionately higher ODR and suspension rates in PBIS-implementing schools (e.g., Bradshaw, Mitchell, O'Brennan, et al., 2010; Martinez et al., 2016; Skiba et al., 2016; Vincent et al., 2009; Vincent & Tobin, 2011; Vincent et al., 2012). Such findings are

disheartening, as a fundamental objective of the PBIS initiative is to improve behavioral and academic outcomes for *all* students.

For instance, in their study which utilized data from an earlier randomized controlled trial (see Bradshaw et al., 2009), Bradshaw, Mitchell, O'Brennan, et al. (2010) found in 21 elementary schools implementing PBIS with high fidelity that even after controlling for the student's level of teacher-rated behavior problems, teacher ethnicity, and other classroom factors, Black students were significantly more likely than White students to receive ODRs. More specifically, their results indicated that if two students (one Black and one White) had identical ratings on the teacher-rated behavior checklist and all other measures, the Black student had a 24%–80% increase in the odds of receiving an ODR (depending on the type of referral) compared to his or her White counterpart. Results also indicated that an ethnic match between students and their teachers did not reduce the risk for referrals among Black students. In fact, Bradshaw et al. found that Black boys with Black teachers had the greatest odds of receiving any type of ODR. It is important to note that there were no significant interaction effects of student ethnicity on the odds of receiving a major ODR or an ODR for defiance. Student gender had a consistent effect across ODR categories, as boys were twice as likely to receive any ODR type than girls.

Skiba et al. (2011) used data from 436 elementary and middle PBIS schools (which were at varying durations and stages of implementation) to determine (a) the extent to which race or ethnicity status contributed to rates of ODR in elementary or middle schools, (b) any specific ODR categories in which racial or ethnic disparities existed, and (c) any categories of disciplinary consequence in which racial or ethnic disparities occurred. Their inquiry was guided

by their assertion that disproportionate representation in school discipline can occur at either the point of referral or administrative decision.

Skiba et al. (2011) found at the K–6 grade level, 35.3% of Black students received ODRs, relative to their proportion in the student population (25.8%). At the 6-9 grade level, 32% of Black students received ODRs, relative to the proportion in the population (21.9%). Overall, Black elementary students received 43% of all ODRs and Black middle school students received 41.7% of ODRs. In contrast with Bradshaw, Mitchell, O'Brennan, et al. (2010), Skiba et al. found that Black students were more likely to receive an ODR for defiance, a behavioral category which may be more vulnerable to subjective interpretation. These findings are consistent with the broader research base that has analyzed discipline patterns for subjective problem behaviors such as "defiance" and "disrespect" (e.g., Girvan et al., 2017; Smolkowski et al., 2016).

At both the K–6 and 6–9 levels, Skiba et al. (2011) found that Black students were significantly overrepresented in ODRs across all discipline categories. At the 6–9 level, Hispanic/ Latino students were also overrepresented relative to White students for all ODR types. Results also suggested that Black and Latino students were overrepresented in suspensions and expulsions relative to White students at both K–6 and 6–9 levels. Black students were underrepresented in the application of detention at the K–6 level, and underrepresented in all administrative consequences except suspension or expulsion at the 6–9 level. In light of these findings, Skiba et al. concluded that race and ethnicity status impact administrative decisions regarding disciplinary consequence independent of the type of infraction. In summary, these results suggested that both differential referral rates at the classroom level and differential

processing at the administrative level perpetuate the disproportionate representation of Black and Latino students in disciplinary actions.

In their analysis of disciplinary exclusion events within 77 elementary, middle, and high schools implementing PBIS, Vincent and Tobin (2011) found that Black students were excluded from school at a disproportionate rate compared to all other ethnicities, and their analysis of patterns of long-term exclusions also indicated that Black students were excluded longer than 10 days during an academic year at a disproportionate rate compared to all other ethnicities. Results also suggested that students with disabilities were excluded from the classroom longer than their nondisabled peers. Vincent and Tobin also evaluated the association between reductions in exclusions and the level of PBIS implementation within four main contexts (whole school, classroom, non-classroom settings, and individual student systems), as measured by the Effective Behavior Support Survey (Sugai et al., 2003). They found that among elementary schools, stronger implementation in classrooms was statistically significantly associated with reduction in suspensions. Yet in high schools, implementation in non-classroom settings was statistically associated with reductions in suspensions. In interpretation of their findings, Vincent and Tobin (2011) cautiously posited that the behavior of elementary-age students might be most directly impacted by what happens in the classroom, whereas the behavior of high school students might be more directly affected by encounters with peers in non-classroom settings. Interestingly, results also indicated a statistically significant association between increases in suspension and PBIS implementation in school-wide settings in elementary schools and in classroom settings in high schools. Overall, the outcomes of this study suggested that while PBIS implementation was associated with decreased exclusions in particular settings (e.g., classroom versus non-

classroom), implementation appeared to have limited effects on the disproportionate exclusion of Black students from school.

While the majority of research indicates that PBIS is not consistently associated with disciplinary equity, some studies suggest that PBIS may have some positive effects on disproportionate exclusion rates when compared to non-implementing schools (Vincent, Swain-Bradway, et al., 2011). Vincent et al. found that PBIS-implementing elementary schools (n = 72) demonstrated lower overall ODR rates per 100 students than non-implementing schools (n = 81) across all three years of the study. While Black students were still overrepresented among students with ODRs in PBIS implementing and non-implementing schools, this overrepresentation was statistically significantly smaller in the implementing group. Though overrepresentation remained approximately constant across the three years in the implementing group and widened in the non-implementing group, the steady increase across time in the non-implementer group was statistically non-significant. Under-representation of White and Latino students among those with ODRs occurred in both groups, but it was statistically significantly smaller in the implementing group. Vincent et al. concluded that PBIS might be linked with reductions in the discipline gap, but additional research is warranted. They also acknowledged that while all PBIS-implementing schools in their study met the minimum SET criterion of 80/80 and a Team Implementation Checklist (TIC; Sugai, Todd, & Horner, 2001) score of 80%, these scores do not represent a validated criterion of full PBIS implementation. Thus, one cannot determine whether full implementation and higher fidelity may be associated with greater effects on disciplinary equity across ethno-racial lines.

Culturally responsive PBIS. A central feature of PBIS is the selection of behavior support practices that are empirically validated yet match the social and cultural context of a

school (Horner et al., 2017). *Contextual fit*, a core principle of PBIS, demands cultural responsivity (Leverson et al., 2016). Scholars in the field assert that few implementation studies incorporate cultural considerations within PBIS models (Bal et al., 2012; Swain-Bradway et al., 2014). In response to these problems, there have been clear calls to assure that PBIS initiatives are culturally relevant (Sugai, O'Keefe, et al., 2012) and systematically and consistently (versus superficially or initially) responsive to cultural and linguistic differences in students (Banks & Obiakor, 2015; Fallon & Mueller, 2017; Swain-Bradway et al., 2014).

As Gay (2002) suggested, changing what educators know and do requires that they are critically conscious of their cultural socialization, and how that affects their attitudes and practices directed toward individuals from ethno-racial backgrounds that are different from their own. Bal et al. (2012) asserted that educators must shift from viewing culture as a static student variable to culture in schools as contextual mediators. Via this lens, the effort to expose and deconstruct inequitable discipline practices that are deeply rooted in each school's structures and patterns is the starting point around which the PBIS framework should be enacted within empirical research and in practice. This requires that interventionists and other stakeholders collaboratively analyze data that allow for critical dialogue about cultural patterns in their school that reflect oppressive and marginalizing institutional practices (Bal et al., 2012). Informed by scholarly work in critical race theory (e.g., Dixson & Rousseau, 2005; Ladson-Billings, 1998), and critical multiculturalism (e.g., May & Sleeter, 2010), this involves interrogating and dismantling the insidious and powerful patterns that shape what behaviors are valued and targeted, how they are taught, how they are interpreted, and who makes (i.e., owns) these decisions.

Banks and Okiabor (2015) argued that culturally responsive intervention elements cannot be treated as add-ons to PBIS frameworks; rather, there are ways to effectively fuse these elements with traditional PBIS features. Within this integrated model, culturally responsive PBIS initiatives systemically promote staff members' cultural knowledge and self-awareness, which includes understanding differences in collectivistic versus individualistic orientations, expressiveness, communication styles, interactions between generations, the role of status and authority, and language (Vincent, Randall, et al., 2011). Within this vein, it is also important for all educators (particularly White educators) to recognize that they are racialized, and that Whiteness is normalized and positioned as central within schools and society (Flintoff et al., 2015; Howard, 2006; Marshall, 2002; Matias et al., 2014).

Furthermore, as Vincent, Randall, et al. (2011) described, schools must assure commitment to culturally relevant and validating student support practices, and reject the false premise of colorblindness which perpetuates notions that the historically privileged culture is neutral and normative. In addition, schools should employ culturally valid decision-making to enhance culturally equitable student outcomes. Rather than emphasizing sameness across all students, commitment to cultural equity emphasizes schools' ability to respond effectively to the differing needs of students from different backgrounds (Vincent, Randall, et al., 2011).

To enhance disciplinary equity, the OSEP National Technical Assistance Center on PBIS has provided a number of resources, including a PBIS Disproportionality Data Guide (McIntosh, Barnes, et al., 2014). This guide was designed to support teams in effective problem-solving via data-based decision making. The tool includes guidance intended to promote identification of particular decision points (i.e., vulnerable decision points) that may be susceptible to subjective interpretation and implicit bias (i.e., perceptions that automatically and unconsciously influence

judgments; McIntosh et al., 2018). As McIntosh et al. described, there are two elements of such decision points that make them vulnerable: (a) the situation itself (e.g., the inherent subjectivity in classifying a behavior or set of behaviors as defiance, disrespect, disruption), and (b) the teacher's decision state in that moment (e.g., fatigued, frustrated).

In their case study of one K-8 school, McIntosh et al. (2018) found that a common school-level vulnerable decision point was physical aggression on the playground (and most specifically, on the basketball court), where the Black - White ODR risk ratio was 4.5 compared to 2.67 overall. Following implementation of the problem-solving model, the ODRs for physical aggression on the playground decreased to one incident, making a risk ratio impossible to calculate. Additionally, over 3 years, ODR rates for Black students dropped to 0.2, below the national median for schools (though a disparity still existed, when compared to an ODR rate of approximately .07 for White students).

Restorative discipline and PBIS. Skiba et al. (2016) endorsed the integration of PBIS and restorative discipline to address persistent disproportionality in disciplinary exclusion rates across schools. Restorative discipline emphasizes positive student-teacher and peer relationships (via such strategies as active listening and reframing), disciplinary fairness and procedural justice, and communal problem-solving, as a set of practices to infuse within the PBIS framework (Skiba et al., 2016; Vincent et al., 2016). Such models involve proactively developed community-building opportunities in the classroom, responsive student and staff circles to promote accountability and to affirm positive norms when problems occur, and restorative conferences following an incident (Vincent et al., 2016).

In an analysis of the application of School-Wide Positive and Restorative Discipline (SWPRD) within a high school setting, implementation was linked to improvements in students'

perceptions of procedural justice in their school, reductions in ODRs overall, and reductions in racial disparities of ODR rates (Vincent et al., 2016). Interestingly, Vincent et al. found that at post-intervention, White students were most likely to rate their school environment as less fair than at pre-intervention. In light of these findings, Vincent et al. suggested that White students may have become more cognizant of inequities within their school environment.

Despite calls for improvement, and despite promising models that fuse culturally responsive and restorative approaches with PBIS, it is very difficult to locate any overt inclusion of culturally responsive features within the descriptions of interventions (i.e., PBIS professional development topics or operationalized elements) in the broader PBIS empirical base (Bal et al., 2012; Swain-Bradway et al., 2014). As it pertains to the present study, it is important to note that there are explicit references to "cultural relevance," "cultural equity," "culturally valid decision-making," and "culturally knowledgeable staff" in one slide within Module 2 of the NCDPI's three-part professional development package that drives training and implementation in North Carolina (NCDPI, 2016b). However, one cannot discern whether these training sections include, for instance, a critical analysis of deeply systemic power structures that perpetuate inequity, or the normalization and centralization of Whiteness around which all other "cultures" traditionally pivot (e.g., Dixson & Rousseau, 2005; Matias et al., 2014; May & Sleeter, 2010). There are no clear references within the modules (NCDPI, 2015a; 2016b) to unconscious bias and vulnerable decision points (McIntosh et al., 2018) as starting points for reflection upon and analysis of discipline patterns in one's classroom and school. Hence, while results of many PBIS studies suggest that the model is linked with reductions in aggregate rates of disciplinary exclusion, it is unclear whether the full emphasis on *all* students has been realized in PBIS research and practice.

PBIS and Organizational Health Outcomes

Positive Behavioral Interventions and Supports programming is intended to enhance the school's overall climate in ways that promote positive student outcomes. Yet it seems logical that this systematic change is difficult to achieve unless those responsible for implementing the model are also experiencing positive effects. Many teachers experience high levels of stress and burnout (Chang, 2009; Reinke et al., 2013), and in some cases their attrition is influenced by student discipline problems (Southern Poverty Law Center, 2008) and poor administrative support (Ingersoll & Smith, 2003). Thus, simply giving teachers isolated strategies to affect change at the classroom level alone may not suffice in addressing these problems (Ross et al., 2012). Rather, PBIS offers a coordinated and systems-driven approach to behavioral intervention that has been associated with increased satisfaction with working conditions and environment (Bradshaw et al., 2008; Bradshaw, Koth, et al., 2009; Caldarella et al., 2011; Houtchens et al., 2017), higher faculty attendance and retention rates (NCDPI, 2017a), enhanced perceptions of safety (Horner et al., 2009), and improvements in educators' perceptions of their own efficacy (Kelm & McIntosh, 2012; Ross & Horner, 2007; Ross et al., 2012). Such outcomes are integral to and indicative of a school's overall organizational health.

Organizational Health Defined

In the PBIS literature base, descriptions of school organizational health typically include the following key features: resource support, collegial relationships and morale among staff, shared commitment to students, academic emphasis, supportive and effective administrative leadership, consistent discipline policies, attention to safety issues, and institutional integrity (Bradshaw et al., 2008; Bradshaw, Koth, et al., 2009; Houtchens et al., 2017). Organizational health, school climate, and school culture are closely connected. School climate not only

encompasses the physical and more tangible aspects of the school environment (e.g., facilities, resources, programs, infrastructure), but also represents the social composition and social cohesion within a school (Caldarella et al., 2011; Van Houtte & Van Maele, 2011). Shared assumptions, beliefs, or values constitute a school's culture, which is encompassed within the broader "climate" (Van Houtte & Van Maele, 2011). Hence, school climate is a complex multidimensional construct that reflects the shared beliefs, values, and attitudes that shape interactions between students, teachers, and administrators and set the parameters of acceptable behavior and norms for the school (Gage et al., 2016).

Effects on Perceptions of Organizational Health

Bradshaw et al. (2008) employed three-level longitudinal analyses to examine the impact of PBIS on staff perceptions of school organizational health, as measured by the Organizational Health Inventory for Schools (OHI; Hoy & Feldman, 1987). Bradshaw et al. used data collected annually across a three-year period from a randomized controlled trial of PBIS involving 37 elementary schools. Bradshaw et al. matched PBIS and comparison schools on particular baseline demographics such as percentage of students receiving free or reduced meals, school enrollment, and percentage of students suspended. They found that staff in PBIS schools reported significantly higher levels of overall organizational health. In addition, staff in PBIS schools reported higher levels of resource influence (i.e., principal's ability to lobby for resources for the school and positively influence the allocation of district resources) and staff affiliation (i.e., warm and friendly interactions, commitment to students, trust and confidence among the staff, and sense of accomplishment). Bradshaw et al. identified a marginally significant effect for academic emphasis, but there were no significant intervention effects on collegial leadership (i.e., principal's behavior is supportive and egalitarian) or institutional

integrity (i.e., the school's ability to cope successfully with destructive outside forces; teachers are insulated from unreasonable community and parental demands).

In their interpretation of findings, Bradshaw et al. (2008) asserted that it is unclear exactly how PBIS impacts these various aspects of school climate. However, they postulated that the PBIS schools' connections with the district may have contributed to the increased opportunity to secure resources for the school and that training in PBIS may have produced a more positive and collaborative work environment for staff. Although not statistically significant, Bradshaw et al. asserted that positive perceptions of academic emphasis in PBIS schools might be a result of enhanced behavior management strategies. They also suggested that while there was no significant effect of PBIS status on collegial leadership, it is feasible that the OHI did not detect differences in attributes of effective leadership that are directly related to the implementation of PBIS or other school-based interventions.

The Bradshaw et al. (2008) study is one of few within the broader PBIS research base to include demographics of the PBIS implementers and educators (e.g., race/ethnicity; gender; age group; specific occupation) within their analysis. Interestingly, Bradshaw et al. reported that staff who were in the minority within the school (i.e., male or "non-White" as they were described) tended to report lower levels of organizational health. In their interpretation of findings, Bradshaw et al. suggested that this might be related to reduced job satisfaction, efficacy, and performance, which administrators and school psychologists might address by providing additional supports to these individuals to improve their perceptions of the school environment. It is unclear from their summary what these additional supports might entail.

In their follow-up analysis using data from the 37 elementary schools included in the original randomized controlled trial, Bradshaw, Koth, et al. (2009) evaluated the extent to which

duration of implementation (i.e., at year one, versus year two, etc.), speed of implementation (i.e., early versus late PBIS implementer status), and achieving 80% or greater on the SET (Sugai, Lewis-Palmer, et al., 2001) were associated with differential organizational health outcomes. Bradshaw et al. found that adjusted mean scores for PBIS versus comparison schools differed significantly at year three on overall OHI, collegial leadership, and staff affiliation. No significant differences between PBIS and comparison schools on resource influence and academic emphasis scores were detected until year four. Furthermore, by the end of year four, there were still no significant differences between PBIS and comparison schools in the adjusted mean scores on intuitional integrity. To partially account for these findings, Bradshaw et al. suggested that the impact of PBIS on school climate may peak around the third year, and level off or possibly decline slightly thereafter.

Results of the Bradshaw, Koth, et al. (2009) analysis also revealed that early implementers (i.e., PBIS schools meeting 80% or greater SET criteria by the end of year one) tended to be more organizationally healthy prior to PBIS training, but tended to improve less than the late implementing schools. In particular, the dichotomous SET score (i.e., at or above 80% or below 80%) was negatively associated with growth in overall OHI and collegial leadership, yet meeting the implementation criteria by the end of year one was positively associated with growth in academic emphasis. Thus, Bradshaw et al. cautiously suggested that the effects of PBIS on school climate may be strongest in schools which are less organizationally healthy.

Analysis of SET scores among PBIS and comparison schools suggested that while there were no significant differences in scores at baseline, all of PBIS schools met SET criteria (i.e., implemented PBIS with high fidelity) within the four years of post-training data collection

(Bradshaw, Koth, et al., 2009). In fact, two-thirds of PBIS schools met criteria within the first year and at least sustained this level across the study duration. In contrast, five of 16 comparison schools never met SET criteria across the duration of the study. Interestingly, when PBIS versus comparison status was removed from the model, results suggested a significant effect of ever meeting SET implementation criteria on the slope of overall OHI, academic emphasis, and staff affiliation. Also noteworthy is that when Bradshaw et al. controlled for the (randomized) PBIS training, there was no significant effect of ever making the SET fidelity criteria on growth in any measured facet of OHI. Overall, results suggested that baseline SET scores were not predictive of the growth in organizational health, nor were they associated with the intercept of organizational health.

Using results from the Teaching, Empowering, Leading, and Learning (TELL) Kentucky survey, Houchens et al. (2017) evaluated teacher perceptions of working conditions in PBIS and non-implementing schools. In addition, Houchens et al. analyzed the extent to which implementation fidelity (as measured by the BoQ; Kincaid et al., 2010) was associated with positive perceptions of working conditions. The study sample included 151 PBIS-implementing elementary, middle and high schools, and 151 non-implementing schools with matched school demographic variables (e.g., total enrollment, dollars spent per student, percentages of White students, percentages of male students, and percentages of students receiving free and reduced lunch rates). The TELL Kentucky survey contains Likert-type items related to eight constructs: Time, Facilities and Resources, Community Support and Involvement, Managing Student Conduct, Teacher Leadership, School Leadership, Professional Development, and Instructional Practices and Supports.

Results demonstrated significant differences between PBIS and non-implementing schools on two constructs: Managing Student Conduct, and School Leadership. Follow-up ANOVAs on Managing Student Conduct showed significant differences on two items. Specifically, teachers in PBIS schools reported higher levels of student and faculty understanding of expectations, and policies and procedures regarding student conduct. Follow-up ANOVAs on School Leadership indicated significant differences in two areas as well, with teachers in PBIS schools reporting more clearly defined expectations and more clearly defined missions and visions in their school. Interestingly, teachers in PBIS schools reported more concern about the use of time in school than those in non-implementing schools.

Results also suggested that the implementation fidelity level (high vs. medium vs. low) significantly influenced outcomes in the areas of Managing Student Conduct, Community Support and Involvement, and Teacher Leadership. Follow-up ANOVAs and other comparisons identified significant differences between any pair of fidelity levels (high vs. medium, medium vs. low, high vs. low) on all items within Managing Student Conduct (e.g., perceived safety of school environment, consistent enforcement of rules). Additional follow-up analyses also revealed significant differences between high- and low-fidelity levels on all items within the Community Support and Involvement construct. Teachers in high- and medium-fidelity PBIS schools reported more positive perceptions of parent involvement, parent–teacher communication, and community support than teachers in the low-fidelity PBIS schools. Regarding results within the Teacher Leadership domain, analyses indicated that teachers in the high-fidelity PBIS schools had more positive perceptions of the leadership opportunities and teacher roles than those in the medium- and low-fidelity PBIS schools.

While findings from the Houchens et al. (2017) study are promising, it did not include any baseline measures of perceived conditions. As Houchens et al. acknowledged, schools with more positive working conditions may have been more predisposed to adopt or to implement PBIS with greater fidelity. Hence, it is unclear whether PBIS had an effect on the actual working conditions as reported by teachers in those schools.

In their quasi-experimental design, Caldarella et al. (2011) used two measures of school climate issued annually across four years to evaluate staff perceptions in two middle schools (one PBIS-implementing, one non-equivalent comparison school). The PBS-Supplemental Questionnaire (PBS-SQ) was developed by Caldarella et al. to specifically measure features of school climate that are most closely aligned to PBIS. The PBS-SQ contains 18 items, each with a five-point Likert scale ranging from one (strongly disagree) to five (strongly agree). Items pertain to students' use of appropriate social skills, equity, community involvement, and teacher praise and encouragement. The Indicators of School Quality (ISQ; Taylor et al., 2006) contains 30 items on a 5-point Likert scale, which are organized into seven categories that comprise school climate: Teacher Excellence, Student Commitment, Parent Involvement, School Leadership, Instructional Quality, Resource Management, and School Safety.

Results of linear trend analyses revealed statistically significant upward trends in the PBIS-implementing school over the four years of the intervention in all of the ISQ categories, with the exception of School Safety. The control school tended to stay the same on measures of school climate, with no statistically significant trends for any of the ISQ categories. Furthermore, the PBIS-implementing school showed increases in the categories of Teacher Excellence, School Leadership, Instructional Quality, and Resource Management, while the control school tended to stay the same or worsen on each. Analyses also revealed statistically significant trends for all

three factors (Prosocial Behavior, School Collaboration/Communication, and Educational Assistance) of the PBS-SQ for the PBIS-implementing school. However, the control school did not show statistically significant changes in these factors over the four years, with the exception of student Prosocial Behavior, which significantly decreased. Effect sizes in the PBIS-implementing school were medium to large and in a positive direction, with the largest effect size being student Prosocial Behavior; effect sizes in the control school were small to medium and in a negative direction.

Overall, this study indicated an association between PBIS implementation and improvement in perceptions of school climate. Nearly all categories of the ISQ (except for School Safety) and all factors of the PBS-SQ showed statistically significant improvements in the PBIS-implementing school compared to the control school. However, one notable limitation was that the schools were not randomly assigned and matched, and the PBIS-implementing school actually indicated worse school climate at baseline than the comparison group. This is important because it might suggest that the effects of PBIS on school climate may be strongest in schools which are less organizationally healthy to start with, as Bradshaw, Koth, et al. (2009) also posited. Also noteworthy is that Caldarella et al. (2011) did not include measures of implementation fidelity across years of implementation within the PBIS school. Thus, it is impossible to know how fidelity changed over time, and how perceptions of school climate might have differed as fidelity levels changed.

While results of two studies (i.e., Caldarella et al., 2011; Sprague et al., 2002) did not link PBIS implementation and improved perceptions of school safety, Horner et al. (2009) found that enhanced implementation of PBIS was associated with improvements in the perceived safety of the school setting. In their randomized controlled trial including 63 elementary schools from two

states (33 randomly assigned to treatment, and 30 assigned to control with one-year delay of PBIS adoption), Horner et al. evaluated results from the School Safety Survey (SSS; Sprague et al., 1996) completed by at least five representatives within each participating school. The SSS produces a risk factor score (with higher scores indicating higher level of behavioral risk) which is based on 17 questions examining such aspects as design of space, crowding, perceived sensitivity to cultural differences, the quality of student–adult interactions, perceived fairness of school rules, and level of adult supervision. The protective factor score (with higher scores indicating higher protection from behavioral risk factors) is based on 16 questions evaluating such elements as clarity of behavioral expectations, student participation, opportunities for student skill acquisition, and formal and predictable systems for conflict resolution. Interestingly, Horner et al. found that the protective factor scores for schools within the sample were consistently high and precluded the option for assessing change. However, there was a statistically significant difference in risk factor scores between the treatment and the control groups after one year, although the decline in risk factor scores after year one in treatment schools did not reach statistical significance. Risk factor scores in control schools actually increased after one year. Finally, results suggested a statistically significant decrease in risk immediately after training for both the treatment and the control (one-year delayed training and implementation) group.

Kelm et al. (2014) conducted a descriptive study within a small elementary school to examine the effects of fidelity of implementation of PBIS on students' feelings of safety, understanding of school expectations, and reports of bullying. Kelm et al. evaluated results from the Satisfaction Survey (which is administered to all students in British Columbia), and found a 15% increase from the previous year in fourth grade students' perceptions of safety, compared

with a 1% change for fourth grade students in the district. Fourth grade students at the participating school also reported a 36% increase in their understanding of what is expected of them, compared to a 1% change for fourth grade students in the district. Finally, fourth grade students reported a 69% decrease in feelings of being bullied, compared with a 15% decrease reported by fourth grade students in the district. Similarly, responses from seventh grade students in the PBIS school also indicated considerable increases in perceptions of safety and decreases in feelings of bullying when compared with other seventh grade students in the district. Notably, seventh grade students in the school and the district reported a similar degree of change in their perceptions of understanding of what is expected of them at school. Overall, results of this study indicated that enhanced implementation fidelity (from year one to year two) was associated with improvements in student perceptions of school safety. As Kelm et al. acknowledged, however, findings of this non-experimental case study involving one small school must be interpreted with caution.

Effects on Teacher Self-Efficacy

Teacher self-efficacy represents a teacher's belief in his or her own ability to affect change in students' academic and behavioral outcomes (Bruning et al., 2011; Tschannen-Moran & Hoy, 2001; Woolfolk & Hoy, 1990). More specifically, self-efficacy involves a teacher's analysis of the teaching task and its context, and an assessment of personal teaching competence (Tschannen-Moran et al., 1998). In general, if a teacher believes that they have the knowledge and skills to enact a particular strategy and that it will lead to desired outcomes, they may be more motivated to initiate and sustain actions that are intended to achieve those objectives. Efficacy beliefs are linked to teacher persistence and resilience (Tschannen-Moran & Hoy, 2001), factors that would seem to have strong relevance to implementation of school-wide

approaches to behavioral intervention. While there is limited research specifically examining the relationship between teacher self-efficacy and PBIS implementation, several studies suggest positive effects.

Kelm and McIntosh (2012) conducted a study that included 22 middle school teachers from two schools implementing PBIS, and 40 teachers from three comparison schools. There were no statistically significant differences in demographics between the PBIS and comparison schools. The two PBIS schools were implementing the model with fidelity, as measured by the SET (Sugai, Lewis-Palmer, et al., 2001) in the year prior to the study, and as measured by the BoQ (Kincaid et al., 2010) during the study. In contrast, no SET or BoQ scores above 50% were obtained among the comparison schools. Participating teachers from the treatment and comparison schools completed the Teachers' Sense of Efficacy Scale (TSES; Tschannen-Moran & Hoy, 2001), which measures teachers' self-assessment of efficacy (on a nine-point Likert scale) in classroom management skills, student engagement, and diverse strategy use. Results from multi-level modeling suggested that teachers at PBIS-implementing schools reported higher levels of teacher self-efficacy than teachers in non-implementing schools, and the large effect size indicated that these differences were meaningful.

In their interpretation of findings, Kelm and McIntosh (2012) suggested that higher self-efficacy ratings are likely related to organizational health, as there may be a stronger sense of shared purpose, stronger administrative support and engagement, and improved perceptions of teaching conditions within the PBIS schools (Bradshaw et al., 2008; Houchens et al., 2017). Kelm and McIntosh also suggested that teachers in PBIS-implementing schools are inclined to use the behavior support strategies that are endorsed and modeled by fellow teachers and administrators, which may increase their self-efficacy in implementing these interventions.

Noteworthy limitations included the small sample obtained via convenience, the unequal sample sizes of participants in PBIS and comparison schools, and the use of only one instrument to measure self-efficacy. Kelm and McIntosh also cautioned that the mean ratings for teachers at both PBIS-implementing and comparison schools were above the mid-range rating, indicating that the majority of teachers felt that they had at least some influence on student outcomes.

Ross et al. (2012) analyzed results from 184 surveys of teachers across 40 elementary schools, and found that teachers in schools implementing PBIS with fidelity (80% or above on the SET) experienced significantly higher levels of efficacy (as measured by the TSES; Tschannen-Moran & Hoy, 2001) when compared to schools that were characterized as low-scoring (with scores below 80% on the SET). Results also suggested that teachers in schools implementing PBIS with fidelity experienced significantly lower levels of burnout and stress, as measured by the Maslach Burnout Inventory–Educators Survey (MBI-ES; Maslach & Jackson, 1981). The MBI-ES evaluates burnout across three factors: Emotional Exhaustion, Depersonalization, and Personal Accomplishment. In their analysis of predictors of teacher efficacy or teacher burnout (including environmental factors and school-level practices), Ross et al. found that none of the teacher or school-level variables involving education, years of experience, gender, ethnicity, or specific PBIS practices (e.g., behavioral expectations taught; acknowledgement procedures) were effective predictors. However, school SET scores and school socioeconomic status were significantly related to all of the burnout and efficacy outcome measures. Ross et al. found that teachers in schools with higher free or reduced-price lunch (FRPL) concentrations may benefit the most from PBIS implementation, as PBIS appeared to significantly increase their feelings of efficacy and the emotional resources that they can deploy.

Summary of Organizational Health Outcome Studies

While there is a growing body of research that links PBIS implementation and improvements in organizational health, additional work is necessary to further examine this relationship. For example, it is important to better understand how indicators of a school's organizational health might differ as implementation fidelity levels change. In this vein, additional study is necessary to identify particular constructs of organizational health that may be more or less responsive to PBIS implementation (as found by Houtchens et al., 2017). To expand upon the findings of Ross et al. (2012), future studies should examine individual and school-level predictors of self-efficacy and teacher burnout, in conjunction with analyses of student behavioral outcomes (e.g., ODR rates, suspension rates, acts of crime and violence). Other potential indicators of organizational health such as teacher turnover should also be evaluated in PBIS and non-implementing schools.

Furthermore, research is necessary to determine whether the positive effects of PBIS on school health are also reported by outside observers who are unaware of the schools' implementation condition (Bradshaw et al., 2008), as one cannot know if measurement bias influences perceptions of organizational health and self-efficacy (Kelm & McIntosh, 2012). Analyses of students' and community members' (e.g., parents and caregivers') perceptions of organizational health in high, medium, and low fidelity, and non-implementing schools, would also provide valuable insight. Studies evaluating the relationship between culturally responsive PBIS and various indicators of organizational health would contribute to the research base and broader applications. Additionally, just as it is crucial for additional studies to analyze disaggregated student outcome data in conjunction with organizational health outcome measures, it is also important for future studies to examine potential differences in teachers' perceptions of

organizational health and self-efficacy when teacher data are stratified across ethno-racial, gender, age, and role (e.g., general educator, special educator; administrator, lead teacher, assistant) categories. Few studies evaluating the relationship between indicators of organizational health and PBIS implementation (e.g., Bradshaw et al., 2008; Ross et al., 2012) have included such factors.

As researchers have acknowledged (e.g., Kelm & McIntosh, 2012), additional randomized controlled trials with large sample sizes should be conducted, in which indicators of organizational health and teacher self-efficacy are measured both before and after implementation of PBIS. It is important to further examine whether the effects of PBIS on school climate may be strongest in schools that are less organizationally healthy at baseline, and how gains made within and across organizational health domains might plateau over time (as suggested by Bradshaw, Koth, et al., 2009). In summary, the various facets of a school's organizational health are likely influenced by PBIS through multiple pathways (Bradshaw et al., 2008), thus it is clear that additional research is warranted to more fully understand this complex and bidirectional interplay.

Factors Associated with PBIS Implementation

Despite promising results in particular areas, many scholars (e.g., McIntosh et al., 2009; Yeung et al., 2016) caution that the broader relevance of PBIS hinges largely on implementation fidelity and sustainability. Some researchers suggest that organizational health is potentially both an outcome of PBIS (Bradshaw, Koth, et al., 2009; Houtchens et al., 2017) and a predictor of implementation quality (Bradshaw et al.; Pas et al., 2014). Results of a longitudinal randomized controlled trial conducted by Pas et al. involving 37 elementary schools, suggested that a number of school-level contextual factors (e.g., school size, behavioral disruptions, baseline SET scores,

student-to-teacher ratio) were associated with variability in the implementation of PBIS at the classroom level. Interestingly, teachers in PBIS schools with higher suspension rates showed greater growth in the implementation of positive behavior strategies over time. In addition, higher student-to-teacher ratios, higher percentage of Black students, and higher SET baseline scores were associated with higher growth over time on the Effective Behavior Support Survey (EBS; Sugai et al., 2003). To account for these results, Pas et al. posited that teachers in these schools may have been more motivated to implement behavior supports. While these factors were associated with positive main effects, analyses of interaction effects revealed that teachers in PBIS schools with higher student-to-teacher ratios and higher percentages of Black students demonstrated the least growth in their classroom implementation of positive behavioral supports over time.

Two teacher level factors (i.e., perceptions of school organizational health, and grade level taught) were associated with variability in the implementation of PBIS at the classroom level. Teachers serving students in lower elementary grades provided a higher level of positive behavioral supports, as measured by the EBS (Sugai et al., 2003). Interestingly, results indicated that teachers with more favorable baseline perceptions of the school organizational health (as measured by the OHI; Hoy & Feldman, 1987) had markedly higher baseline scores on the EBS. However, these teachers also demonstrated less growth over time, as perhaps they had less room for improvement (Pas et al., 2014). This finding is similar to previous research (Bradshaw, Koth, et al., 2009) which suggested that schools with higher baseline OHI scores exhibited less growth over time.

Using statewide evaluation data from 807 elementary schools, Bradshaw and Pas (2011) identified several school-level and district-level factors associated with PBIS training, adoption,

and implementation quality. In particular, schools with higher suspension and student mobility rates were more likely to engage in PBIS training; furthermore, schools with higher student mobility rates were more likely to reach adoption status (i.e., receive PBIS training and submit implementation data). Interestingly, suspensions, truancy, enrollment, students per teacher, special education rates, and percent of teachers certified were not significantly associated with PBIS adoption. Schools with higher reading achievement scores were less likely to be PBIS trained or to achieve adoption status. Schools in larger districts were less likely to be PBIS trained, but schools in districts with higher percentages of implementation data submission were more likely to be PBIS trained and to achieve implementing status. Overall, district-level factors were more associated with training and adoption than with implementation quality.

Regarding implementation quality, PBIS schools with higher concentrations of certified teachers demonstrated higher scores on the Implementation Phases Inventory (Bradshaw, Koth, et al., 2009). Indicators of school disorder such as suspension, mobility, and truancy rates were not significantly associated with implementation quality, a finding generally supported by additional research (Bradshaw et al., 2015). Bradshaw and Pas (2011) also found that the number of years since PBIS training was positively associated with implementation quality. Finally, it is noteworthy that while FRPL percentages were not included in this analysis, proxies for family income (e.g., district-level education expenditures and the percent of schools receiving Title I funding) were not significantly associated with differences in PBIS training, adoption, or implementation quality.

Analyses of data (Molloy et al., 2013) from over 27,000 students and 166 public elementary and secondary schools across seven states indicated that a number of factors, including percentage of FRPL, school size, and school level were associated with variability in

implementation quality in three key areas: expectations taught, rewards system and violations system (as measured by the SET; Sugai, Lewis-Palmer, et al., 2001). In particular, Molloy et al. found that school level (elementary versus secondary) significantly predicted implementation quality of expectations taught, and marginally significantly predicted quality of schools' reward systems, with elementary schools demonstrating higher levels of implementation quality in both areas. Smaller schools achieved higher implementation quality in both the areas of reward systems and expectations taught. To account for this effect, Molloy et al. asserted that the smaller student body may allow staff to more closely monitor and more frequently engage students. Finally, percent of students with FRPL negatively predicted expectations taught, suggesting that higher socioeconomic status schools achieved better implementation quality than lower socioeconomic status schools. Molloy et al. suggested that schools with more resources may be able to secure what is needed to employ critical PBIS elements, and to prioritize the initiative over other programs.

In their analysis of SET data from 890 rural, suburban, and urban schools across 20 states, Frank et al. (2009) did not find an association between school socioeconomic status (measured by percentage of students eligible for FRPL) and implementation quality. Specifically, very low socioeconomic status schools (more than 75% of students eligible for FRPL) were almost equally as likely to achieve 80% implementation status on the SET (Sugai, Lewis-Palmer, et al., 2001) as their very high socioeconomic status counterparts (less than 10% students eligible for FRPL). Interestingly, results also indicated that implementation was higher in schools with moderate racial diversity (25% - 50% students of color) than for schools with high or low diversity. While these findings are compelling, the authors acknowledged that selection bias, potential for SET data collection errors, and non-equal demographic groups were

notable limitations. In addition, as data were only analyzed for one year, it is impossible to know how changes in implementation status and fidelity over time might be associated with the demographic factors examined therein.

In order to identify factors associated with sustained implementation, McIntosh et al. (2015) administered the School-wide Universal Behavior Sustainability Index: School Teams (SUBSIST; McIntosh et al., 2011) to staff at 860 elementary and secondary schools at varying stages of PBIS implementation. The SUBSIST is comprised of two school level factors (school priority and team use of data) and two district-level factors (district priority and capacity building). In addition, staff responded to three Likert-type questions pertaining to school-team actions: frequency of PBIS, frequency of sharing data with the whole school staff, and hours of PBIS coaching received. School demographic factors included grade levels served (elementary, middle, or high school), enrollment, urbanicity (on a scale of one to four), percent of students of color, and percent of students receiving FRPL. Results suggested that grade levels served and years implementing were both significant predictors of team use of data and school priority (i.e., school-level factors). Specifically, high schools demonstrated consistently lower scores on these school-level sustainability factors. Strikingly, high schools exhibited significantly higher district-level sustainability factor scores than elementary and middle schools in the study. McIntosh et al. posited that high schools may be the last schools to adopt PBIS within a district that has already established strong systems of support, along with clear policies on the initiative.

McIntosh et al. (2015) found that other school demographics (e.g., percent of students receiving FRPL; percent of students of color) were not significantly related to sustainability. Rather, the frequency of team meetings and hours of PBIS coaching were each related to one sustainability factor, and the frequency of teams sharing data with all school staff was

significantly related to all four sustainability factors. While promising, McIntosh et al. conceded that these findings must be interpreted with caution, as the study included data from only one year. Hence, it is unclear how school demographic and other factors might have been associated with sustained implementation over time.

McIntosh, Mercer, Nese, and Ghemraoui (2016) employed a longitudinal study that included over 5,000 elementary and secondary schools to identify school and district-level predictors of membership within the following four implementation pattern groups: sustainers, slow starters, late abandoners, and rapid abandoners. Implementation measures were obtained via available scores on the SET, BoQ, and the PBIS Self-Assessment Survey (SAS; Sugai, Horner, & Todd, 2000), but only the measure with the strongest external evaluation component was included (i.e., first the SET, then BoQ, then the SAS). Significant predictors of group membership included grade levels served, enrollment, proportion of schools implementing PBIS within the district, and size of the implementation cohort. Elementary schools, larger schools, schools in districts with more schools already implementing PBIS, and those starting within a larger initial district cohort were more likely to be in the sustaining groups. As McIntosh et al. described, this larger community of practice may be more effective in advocating for and securing resources from the district, which may cultivate a greater sense of accountability to and membership within the larger group.

Middle and high schools were more likely to be in the abandoners (rapid and slow) group, a finding that generally aligns with other studies which point to unique implementation challenges at the secondary level (e.g., Bohannon et al., 2006; Feuerborn et al., 2016; Flannery et al., 2014; McIntosh et al., 2015; McIntosh, Mercer, Nese, Strickland-Cohen, & Hoselton, 2016; Molloy et al., 2013). In contrast with some previous research (Bradshaw & Pas, 2011; Molloy et

al., 2013), after controlling for grade levels served and district size, schools with higher enrollment were more likely to be in the sustaining groups. McIntosh, Mercer, Nese, and Ghemraoui (2016) suggested that larger schools may have more resources to apply toward such initiatives, and larger schools may be less vulnerable to the negative effects of teacher turnover because there are more staff. On a related note, Andreou et al. (2015) found in their qualitative study that participants identified staff turnover as one potential barrier to sustained implementation of PBIS.

The results of a longitudinal study conducted by McIntosh, Mercer, Nese, Strickland-Cohen, and Hoselton (2016) suggested that grade level and speed of implementation were significant predictors of sustained implementation of PBIS. Specifically, in their analysis of SET (Sugai, Lewis-Palmer, et al., 2001), BoQ data (Kincaid et al., 2010) and PBIS Self-Assessment Survey data (SAS; Sugai, Horner, & Todd, 2000) obtained from years one, three and five cohorts, they found that middle and high schools were less likely to meet PBIS fidelity criteria in both years three and five than elementary schools. Furthermore, results indicated that schools with lower rates of students receiving FRPL were more likely to sustain at high criterion, a finding which is generally consistent with other PBIS studies (e.g., Molloy et al., 2013) but not supported by other studies (e.g., Frank et al., 2009; McIntosh et al., 2015). Inconsistent with the findings of some studies (e.g., McIntosh et al., 2015; McIntosh, Mercer, Nese, & Ghemraoui, 2016), urbanicity and high enrollment were not found to be predictors of sustained implementation in this case. Additionally, schools that met criterion for implementation in year one (as measured by the TIC; Sugai, Todd, & Horner, 2001) were more likely to be sustaining in year five.

Beyond identification of school characteristics that predict sustained implementation, McIntosh, Mercer, Nese, Strickland-Cohen, and Hoselton (2016) also examined the percentage of variance at the school, district, and state levels at years one, three, and five. They found the largest differences in fidelity between the states, and therefore concluded that states played a more significant role in initial and sustained implementation. In light of the increasing proportion of variance at the district level, McIntosh et al. asserted that district systems play a smaller role in initial implementation speed, but an increasing role as schools progress towards five years. Additional PBIS research (e.g., Coffey & Horner, 2012; George & Kincaid, 2008; Horner et al., 2014; McIntosh, Mercer, Nese, & Ghemraoui, 2016) also points to the significant impact that state-level and district-level support (e.g., funding, coaching and training capacity, established cohorts or communities of practice) have on sustained and scaled-up PBIS implementation.

At the school level, PBIS research suggests that such factors as administrator support (Coffey & Horner, 2012; McIntosh, Predy, et al., 2014) and effective school teams (McIntosh, Predy, et al, 2014.; McIntosh et al., 2015) are necessary for initial implementation and sustainability. Interestingly, McIntosh, Predy, et al. (2014) also found in their quantitative and qualitative analyses that staff support, integration into typical practice, and parent involvement were rated as significantly more important to sustainability than initial implementation. Staff support (i.e., buy in) likely exists on a continuum and is influenced by a variety of multi-level factors (e.g., personal beliefs, time, resources) that mutually interact to affect the degree to which an educator enacts and sustains PBIS (Bambara et al., 2012; Filter et al., 2016).

Summary of Implementation Studies

There is increasing focus on identifying factors (at the classroom, school, district, and state level) that influence the extent to which PBIS is implemented and sustained over time.

Some studies have linked variability in implementation to such school-based characteristics as school size (McIntosh, Mercer, Nese, & Ghemraoui, 2016; Molloy et al., 2013; Pas et al., 2014), grade level (McIntosh et al., 2015; McIntosh, Mercer, Nese, Strickland-Cohen, & Hoselton, 2016; Molloy et al., 2013.; Pas et al., 2014), concentration of certified teachers (Bradshaw & Pas, 2011), and percentage of students eligible for FRPL (McIntosh, Mercer, Nese, Strickland-Cohen, & Hoselton, 2016; Molloy et al., 2013). Research suggests that particular factors such as school size, established communities of practice within the district, and grade level, may influence speed of implementation as well as implementation trajectories over time (McIntosh, Mercer, Nese, & Ghemraoui, 2016). Additional research emphasizes district-level factors as strongly associated with implementation (e.g., Bradshaw & Pas; Coffey & Horner, 2012; McIntosh, Mercer, Nese, & Ghemraoui, 2016).

Finally, some research (i.e., Pas et al., 2014) suggests that higher baseline levels of school disorder (e.g., higher suspension rates) may be associated with greater improvements in implementation, as one might presume that such schools are more motivated to address these issues. Yet, other studies (e.g., Bradshaw & Pas, 2011) do not find an association between indicators of school disorder, and PBIS adoption and implementation quality. In light of the existing research, it is difficult to draw clear conclusions about which specific factors (or specific interactions) most influence PBIS implementation and sustainability. However, it seems plausible that while some school characteristics play a marked role in sustained implementation, other factors (e.g., district-level support) might mitigate these risks (McIntosh, Mercer, Nese, Strickland-Cohen, & Hoselton, 2016). Additional research is certainly necessary to further identify these factors, to understand the potentially complex interplay involved, and to evaluate the differential effects of these factors on PBIS adoption and implementation fidelity over time.

Summary

The PBIS model strives to promote and sustain socially significant behavior change via the application of evidence-based practices and effective support systems that are organized within a fluid framework and employed at scales of social importance. Within the ever-expanding body of research on PBIS, there is strong interest in evaluating the multi-faceted effects of this model on outcomes that encompass and extend beyond student behavior. Overall, PBIS has been linked with reductions in aggregate rates of ODRs, suspensions, and increases in attendance rates (e.g., Childs et al., 2016; Flannery et al., 2014; Kelm et al., 2014). The PBIS initiative has also been associated with educators' increased satisfaction with working conditions and environment (e.g., Horner et al, 2009; Houtchens et al., 2017), and improvements in educators' perceptions of their own efficacy (e.g., Kelm & McIntosh, 2012). Yet, to this point, PBIS has not been consistently connected with disciplinary equity (Skiba et al., 2016). Additional research is necessary to further examine the effects of PBIS (implemented at different levels of fidelity) on socially significant student and staff outcomes. It is also essential for the field to better understand how particular contextual factors (of a school, district, etc.) might be associated with such effects.

Chapter 3: Method

This chapter provides a description of the research design and methodology of the study. It includes a description of the sample included within the study, and the variables of interest. Data compilation and analysis procedures are also addressed.

Research Design

This study used a descriptive quantitative design. This method was employed in order to evaluate the relationship between Positive Behavioral Interventions and Supports (PBIS) implementation and particular organizational health and behavioral outcome measures within the sample.

Sample

The original data sets included all kindergarten through 12th grade public schools within North Carolina for the 2015-16 school year. In this study, however, all charter schools (n =159), public alternative schools (n = 74), middle schools (n = 479), and schools serving ninth grade through early college (n = 588) were excluded. Elementary schools serving sixth grade and intermediate schools (n = 136) were also eliminated. Only those schools serving grade combinations of kindergarten through fifth grade were considered, thereby yielding an initial total of 1,230 schools, and 92% Local Education Agencies (LEAs) across the state (including city and county units). Forty-eight percent of these LEAs had no PBIS-implementing elementary schools. Schools serving middle and high school students were excluded because as the original data set conveys, the average statewide rates of crime and violence are substantially higher in these contexts (0.2% of elementary students versus 1% of high school students). Secondary schools were also excluded in light of the fact that the overwhelming majority of PBIS-

implementing schools are elementary and middle school level (Freeman et al., 2016), and among PBIS schools, different school levels (e.g., elementary versus high school) adopt and sustain PBIS to differential degrees (McIntosh, Mercer, Nese, & Ghemraoui, 2016; Molloy et al., 2013). The original data set supports these findings, as high schools represent only 8% of the PBIS-implementing cohort in North Carolina within the 2015-16 school year.

It is noteworthy that 66% of the schools also provided pre-kindergarten services, and 28% of included schools served all grades kindergarten through fifth (see Table 1 for a summary of all grade combinations represented within the sample). Six percent of the elementary schools operated on a year-round school calendar and the remaining schools operated on a traditional calendar.

Of the 1,230 elementary schools, only those schools contained within LEAs which had at least one PBIS-implementing school and at least one non-implementing school were included in analyses. These criteria were applied given the interest in comparing organizational health and behavioral outcome measures in PBIS schools versus non-implementing schools, nested within their LEAs. As a result, a total of 835 schools and 46 LEAs were included in the study. A total (i.e., average daily membership—ADM) of 445,951 students attended these elementary schools. The ADM of individual schools ranged from 36 to 1,172 students. Eighty percent of schools had ADM totals between 250 and 750 students (see Table 2 for a summary of ADM totals represented).

The elementary student population was approximately 45% White, 27% Black, 18% Hispanic, 3% Asian, 1% Indigenous (the DPI uses the term “American Indian”), 0.2% Pacific Islander, and 4% multiracial. The remaining two percent of students were not identified within any ethno-racial category. Fourteen percent of students in this sample were identified as having a

disability. In 59% of the schools, over one-half of all students were identified as economically disadvantaged (EDS) and thus eligible for free or reduced-price lunch (see Table 3 for a summary of EDS concentrations represented).

In terms of LEA representation, the largest LEA included 107 elementary schools and the smallest LEA included three elementary schools. Forty-eight percent of LEAs within this study were represented by ten or fewer elementary schools (see Table 4 for a full summary of LEA representation). The North Carolina Department of Public Instruction (NCDPI) divides the state into eight geographic regions (2020b). All regions were represented in this study, with 46% of schools located in the western regions, 37% of schools in the central regions and 17% of schools in the eastern regions.

Of those elementary schools included in this study, a total of 323 schools (39%) met PBIS-implementing status criteria for the 2015-16 year. In terms of PBIS implementation recognition, 18% of those schools within this sample received Exemplar status, 40% received Model status, and 42% received Green Ribbon designations. Of those LEAs included in the study, 33% achieved only one implementation level across schools within the LEA, 37% achieved two implementation levels across schools, and 31% achieved all three implementation levels across schools. Within the LEAs, PBIS implementation concentration ranged from 5% to 93% of schools. Thirty percent of LEAs had between 5% and 25% concentration, 33% had between 26% and 50% concentration, 20% had between 51% to 75% concentration, and 17% had over 75% concentration.

Data Source

The overwhelming majority of data analyzed in this study were obtained from the North Carolina Department of Public Instruction (DPI) via a collaborative research project involving faculty from a large state university and the DPI. However, the annual teacher turnover data were accessed directly from the DPI's public website. The university's institutional review board approved the research project in January, 2017. The data analyzed in this study were originally collected by district-level staff and subsequently submitted to the DPI. The DPI staff examined submitted data and consulted with school district personnel to address missing data or possible coding errors.

Data Collection Instrument

Each data set provided by the DPI was subsequently merged and organized within a Microsoft Excel spreadsheet. Each school was identified by name, and by three-digit local LEA code followed by three-digit school code. This spreadsheet included the following information for each included school within each LEA or district:

- Average daily membership (ADM)
- PBIS implementing or non-implementing status
- PBIS implementation level, if applicable: Green Ribbon, Model, or Exemplar
- Percentage of students eligible to receive free or reduced price lunch (FRPL)
- Concentration of economically disadvantaged students by category: low, mid-low, mid-high, or high
- Overall composite mean rate of agreement on Teacher Working Conditions Survey
- Annual teacher turnover rates (per school, LEA, and statewide)

- Aggregated acts of crime and violence per 100 students (per school, LEA, and statewide)
- Short-term suspensions per 100 students (per school, LEA, and statewide)
- Long-term suspensions per 100 students (per school, LEA, and statewide)
- Number of crime and violence acts within each school, disaggregated by category
- Number of students with disabilities
- Number of office disciplinary referrals issued to students with disabilities
- Number of office disciplinary referrals issued to nondisabled students
- Number of students disaggregated by ethno-racial category
- Number of short-term suspensions disaggregated by ethno-racial category
- Number of long-term suspensions disaggregated by ethno-racial category
- Number of crime and violence acts disaggregated by ethno-racial category

(Refer to Appendix for a description of variables included within the study).

Once all data were merged, the primary investigator examined the information to ensure that it was complete. The database was uploaded and analyzed using R (Version 3.5.3; 2019). Upon discovery of several extreme values within particular variables (i.e., school suspension rates per 100 students), faculty involved in the research project confirmed with the DPI staff that there were no coding errors. Data are considered missing completely at random (MCAR; Tabachnick & Fidell, 2007). Hence, complete case analyses were appropriate and missing values were dropped from the models.

Data Analysis

To prepare for analysis, all data were converted into functional forms. Several nominal variables were dichotomous and thus converted to a “1” or “0.” These included a school’s status

as PBIS-implementing or non-implementing. Ordinal variables (Green Ribbon, Model, and Exemplar implementation levels) were converted to a “1,” “2,” or “3,” respectively. To support meaningful interpretation and to align with categories applied by the U.S. Department of Education (2015), the percentage of students eligible to receive free or reduced-price lunch within each school was converted to the concentration of economically disadvantaged students by category: low, mid-low, mid-high, or high.

Analyses were employed to address the following research questions:

1. Are there significant differences in organizational health outcome measures (i.e., teacher satisfaction ratings and one-year teacher turnover rates) among PBIS-implementing versus non-implementing schools when other variables are also considered (e.g., LEA, ADM, concentration of economically disadvantaged students, crime and violence rates, and suspension rates)?

2. Within PBIS-implementing schools, are there significant differences in organizational health outcome measures (i.e., teacher satisfaction ratings and teacher turnover rates) across different implementation levels when other variables are also considered (e.g., LEA, ADM, concentration of economically disadvantaged students, crime and violence rates, and suspension rates)?

3. Are there significant differences in behavioral outcome measures (i.e., short-term suspension, long-term suspension, and crime and violence rates) among PBIS-implementing versus non-implementing schools?

4. Are there significant differences in suspension and crime and violence rates across distinct ethno-racial categories in PBIS-implementing versus non-implementing schools?

5. Are there significant differences in office disciplinary referral rates among students with disabilities and non-disabled students in PBIS-implementing versus non-implementing schools?

6. Do PBIS-implementing schools with higher implementation fidelity demonstrate significant differences in suspension and crime and violence rates when compared to schools with lower implementation fidelity levels?

7. Do PBIS-implementing schools with higher implementation fidelity demonstrate significant differences in suspension and crime and violence rates across ethno-racial categories when compared to schools with lower implementation fidelity levels?

8. Do PBIS-implementing schools with higher implementation fidelity demonstrate significant differences in office disciplinary referral rates among students with disabilities when compared to schools with lower implementation fidelity levels?

To address each question, the investigator initially applied linear mixed effects regression analyses to account for the random effect of LEA. The rationale for the mixed effects procedures is that a LEA may produce significant variation between schools and may have a significant effect on the school's structure as it pertains to these organizational health and behavioral outcomes areas. Poisson regression analyses were initially applied to model count data. A Bonferroni critical value procedure was applied to adjust for the number of models included and to control for error rates (Harris, 1998).

Chapter 4: Results

Chapter 4 provides a summary of the results of the present study on organizational health and behavioral outcome measures in elementary schools implementing Positive Behavioral Interventions and Supports (PBIS), compared with non-implementing elementary schools. The chapter is organized to align with the primary research questions: (a) Are there significant differences in organizational health outcome measures in PBIS-implementing versus non-implementing schools when a variety of factors are also considered (i.e., district, school population size, concentration of economically disadvantaged students, suspension rates, and crime and violence rates)?; (b) Within PBIS schools, are there significant differences in organizational health outcome measures as PBIS implementation levels change when the aforementioned factors are also considered?; (c) Are there significant differences in behavioral outcome measures (i.e., suspension rates and crime and violence rates) in PBIS-implementing versus non-implementing schools?; (d) Are there significant differences in behavioral outcome measures stratified across different ethno-racial categories in PBIS-implementing versus non-implementing schools?; (e) Are there significant differences in office disciplinary referral rates for students with disabilities and nondisabled students in PBIS-implementing versus non-implementing schools?; (f) Within PBIS schools, are there significant differences in behavioral outcome measures across different ethno-racial categories as PBIS implementation levels change?; and (g) Within PBIS schools, are there significant differences in office disciplinary referral rates for students with disabilities and nondisabled students as PBIS implementation levels change?

Simultaneous linear mixed effects regression analyses were originally conducted to address the majority of these questions. Poisson regression (i.e., log-linear modeling) analyses

were conducted to model count data (i.e., behavioral outcomes disaggregated across ethno-racial categories; office disciplinary referrals among students with disabilities versus nondisabled students). For any models which met primary linear regression assumptions, the marginal-$R^2$and conditional-R^2 were computed (Nakagawa & Schielzeth, 2013) to account for the proportion of variance explained by the fixed effects, and to account for the proportion of variance explained by the fixed and random effects combined.

Across all models, normality assumptions were checked by plotting the random effects, and by plotting the residuals (via quantile-quantile plots). Model linearity and homoscedasticity were evaluated by plotting the residuals versus the predictors (Bruce et al., 2020). Diagnostics suggested that the majority of models contained violations of linearity, and normally distributed residuals signaled by skewness (deviation from the normal line) and kurtosis (increased probability in the tails). These models were fit via a generalized estimating equation (GEE), which is a semi-parametric extension of regression analysis that does not assume normally distributed residuals (Ballinger, 2004). Generalized estimating equations estimate population averaged effects (the marginal effects). In other words, a GEE provides an estimate of how much the average response would change for every one-unit change in the predictor variable, whereas the linear mixed model estimates the effects of covariates when holding the random effects constant. In mixed models, it is assumed that the random effects are normally distributed around the population parameters, however, this assumption is not made within a GEE. Rather, in the GEE models, the local education agency (LEA) served as a grouping factor in the correlation matrix because LEA is a significant source of variation and closer to the true correlation structure. In each GEE model, the correlation structure was specified *a priori* as exchangeable, meaning that the correlation between any two measures within the same cluster (i.e., LEA) is the

same as the correlation between any two measures in a different cluster (Halekoh et al., 2006). Practically, this decision was applied because it is likely that the LEA is influential on its schools, and will likely have a comparable degree of influence on all schools within its cluster. Furthermore, this structure is appropriate given that these clustered data were not collected over time (Ballinger, 2004). The GEE model is robust to errors in specification of the correlation structure because estimates of the regression parameters remain consistent; in other words, the estimated values approach the true values with probabilistic certainty as the sample size increases (Ballinger, 2004). Given that marginal mean and variance is correctly specified, the standard errors are considered robust. While the estimate of the variance produced under the GEE model could be highly biased when there are fewer than 20 clusters (Ballinger, 2004), the models within this study included a larger number of clusters (i.e., LEAs). The GEE models in this study produced the Wald statistic, which is mathematically related to the *z*-statistic, and is calculated by squaring the value obtained when the beta coefficient is divided by its standard error (Wilcox, 2017).

Prior to diagnostics of the Poisson models, raw residuals were transformed to Pearson's residuals in order to meaningfully compare observed to expected counts (Bruce et al., 2020). In addition to the aforementioned diagnostics, the principal investigator also assessed for overdispersion, which occurs when data have greater conditional variance than the conditional mean in the Poisson model (Payne et al., 2018). The principal investigator estimated dispersion for each model by calculating the ratio of model deviance to its corresponding degrees of freedom; when the ratio is equivalent to one, the Poisson assumption is met (Payne et al., 2018). However, a number of these models suggested overdispersion, which could lead to an underestimation of parameter standard errors. To address overdispersion, a negative binomial

model was applied. The negative binomial model has a more flexible variance structure than the Poisson, accounting for a variance greater than the mean by adding an additional variance parameter to the model (Payne et al., 2018). Overdispersion was still evident in most negative binomial models, and thus the Poisson models were fit using the GEE approach as well.

Organizational Health Outcome Measures in PBIS and Non-Implementing Schools

The analyses comparing organizational health and behavioral outcomes in PBIS versus non-implementing schools included 835 schools (within 46 LEAs). The cases represented schools where there was at least one instance of PBIS implementation and at least one instance of non-implementation within its LEA, due to the interest in PBIS effects while considering schools nested in LEAs. The choice of at least one instance was made to maintain a sense of variability due to PBIS implementation and to retain as many representative LEAs as possible.

Simultaneous linear mixed effects regression analyses were initially conducted to identify any differences in the following organizational health outcome measures: annual teacher turnover rates and overall teacher satisfaction (represented by the composite mean rate of agreement score on the Teacher Working Conditions Survey for 2015-16) for PBIS versus non-implementing schools. These analyses accounted for the assumption that each school's LEA is a significant source of variation (i.e., the random effect in the model) between schools and is influential within each school as well. Thus, within the model, each school had a random intercept depending on LEA. Subsequent to diagnostics, these models were fitted via a GEE. The coefficient estimates produced by the GEE describe changes in the population mean given changes in the predictors, while still accounting for within-LEA non-independence of observations.

Teacher Turnover and PBIS Status

Other fixed factors beyond PBIS implementation status (dichotomously coded as "1" or "0," non-implementing) were included in the GEE model to account for their potential effects on teacher turnover. These included the following: school population (i.e., average daily membership, standardized to account for the wide variation across schools), concentration of economically disadvantaged students (categorically represented as low, mid-low, mid-high, or high), long-term suspension rates, short-term suspension rates, crime and violence rates, and the overall composite mean rate of agreement score on the North Carolina Teacher Working Conditions Survey.

The effect of PBIS status on teacher turnover rates was not significant when the following factors were considered: school population size, concentration of economically disadvantaged students, short and long-term suspension rates, crime and violence rates, and overall composite teacher satisfaction scores. However, concentration of economically disadvantaged students ($p < .001$) and crime and violence rates ($p < .001$) were significantly associated with teacher turnover rates. In particular, when holding other factors constant, a move from a high concentration of economically disadvantaged students to low concentration produced an estimated 4% decrease in teacher turnover. Interestingly, for each additional act of crime and violence (per 100 students), teacher turnover decreased by approximately 0.38%. In the original GEE model, short and long-term suspension rates were significantly positively associated with teacher turnover. However, following Bonferroni critical value procedures (i.e., accounting for all of the models run), suspensions no longer met the significance criterion. Other factors included in the model (e.g., school population, teacher satisfaction) were not significant. Parameter estimates are presented in Table 5.

Teacher Satisfaction and PBIS Status

Additional fixed factors beyond PBIS implementation status were included in the GEE model to account for their potential effects on teacher satisfaction. These included the following: school population (i.e., average daily membership via standardization, given the wide variation across schools), concentration of economically disadvantaged students, long-term suspension rates, short-term suspension rates, crime and violence rates, and teacher turnover rates. The effect of PBIS status on overall teacher satisfaction was not significant when these other factors were considered. However, concentration of economically disadvantaged students ($p < .001$), short-term suspension rates ($p < .001$), long-term suspension rates ($p < .001$), and school population size ($p < .001$) were significantly associated with teacher satisfaction. Specifically, moving from a high concentration of economically disadvantaged students to a low concentration (holding other factors constant) produced an estimated 9% increase in teacher satisfaction. For each additional short-term suspension (per 100 students), teacher satisfaction decreased by approximately 0.15%; every 10 additional suspensions (per 100 students) yielded an estimated 1.5% decrease in teacher satisfaction. Each additional long-term suspension (per 100 students) was associated with an estimated 4% decrease in teacher satisfaction. In addition, as the school population increased by one standard deviation, overall teacher satisfaction decreased by an estimated 1.8%. Holding other variables constant, larger schools were associated with a decrease in teacher satisfaction. Other factors included in the model were not significantly associated with teacher satisfaction (see Table 6).

Organizational Health Outcome Measures within PBIS Schools

The analyses comparing organizational health outcomes across different PBIS implementation levels included 285 PBIS-implementing schools (within 31 LEAs). Given the interest in evaluating the effects of PBIS implementation fidelity while considering schools nested in LEAs, these cases represented schools where there were at least two implementation levels achieved within its LEA. Any implementation level combination (e.g., Green Ribbon and Exemplar; Green Ribbon and Model) was acceptable.

Teacher Turnover and PBIS Implementation Level

Initially, a linear mixed effects regression analysis was conducted to identify any differences in annual teacher turnover rates across the three PBIS implementation levels (Green Ribbon, Model, and Exemplar). The school's LEA accounted for the random effect in the model. The following fixed factors were included in the mixed model to account for their potential effects on teacher turnover: school population (i.e., average daily membership; a standardized variable), concentration of economically disadvantaged students, short-term suspension rates, long-term suspension rates, crime and violence rates, and the overall composite mean rate of agreement score on the Teacher Working Conditions Survey. Model diagnostics revealed non-normality in the residuals, and thus a GEE approach was applied.

The effect of PBIS implementation level on teacher turnover rates was not significant when these other factors were considered. In the original GEE model, a school's concentration of economically disadvantaged students had a small but statistically significant effect on teacher turnover rates. However, following Bonferroni critical value procedures, this factor no longer met the significance criterion. No other factors were significantly influential (see Table 7).

It is interesting that when PBIS implementation level was included in the GEE model (versus PBIS implementation status), the concentration of economically disadvantaged students within a school was not significantly influential on teacher turnover rates. In terms of comparison, however, it must also be noted that this model included only schools that were PBIS implementing (and nested within LEAs with at least two implementation levels represented).

Teacher Satisfaction and PBIS Implementation Level

A linear mixed effects regression analysis was conducted to identify any differences in teacher satisfaction across the three PBIS implementation levels. The LEA was modeled as the random effect. The following fixed factors were included in the mixed model to account for their potential effects on teacher satisfaction: school population (i.e., average daily membership), concentration of economically disadvantaged students, short-term suspension rates, long-term suspension rates, crime and violence rates, and annual teacher turnover rates. This model met linearity and normality assumptions.

In this model, the effects of PBIS implementation level on teacher satisfaction were significant even when these other factors were included, F (2, 271.42) = 10.83, $p < .001$. Specifically, as shown in Table 8, moving from Green Ribbon to Model implementation level produced an estimated 4% increase in teacher satisfaction when all other variables were held constant; moving from Model to Exemplar produced an estimated 5% increase in teacher satisfaction. In addition to PBIS implementation level, the concentration of economically disadvantaged students, F (3, 276.90) = 8.33, $p < .001$, and school population, F (1, 266.19) = 12.05, $p < .001$, were also significantly influential factors. Parameter estimates suggest that teacher satisfaction increased by approximately 10% in PBIS-implementing schools with a low concentration of economically disadvantaged students in comparison to schools with high

concentrations of economically disadvantaged students (when holding other factors constant in the model). In addition, as the school population increased by one standard deviation, overall teacher satisfaction decreased by an estimated 1.8%. Other variables such as long-term and short-term suspension rates were not significantly influential. The fixed factors in the model accounted for 18% of the total variance in teacher satisfaction, marginal-R^2= .18. The fixed and random effects combined accounted for 26% of the total variance, conditional-R^2= .26. It is noteworthy that when PBIS implementation level was included in the model (versus PBIS implementation status), suspension rates no longer had a significant effect on teacher satisfaction.

Behavioral Outcome Measures in PBIS and Non-Implementing Schools

Following diagnostics which revealed non-normality in the residuals, these models were fit via the GEE approach. The analyses comparing behavioral outcomes in PBIS versus non-implementing schools included 835 schools within 46 LEAs (i.e., clusters). The analyses accounted for the assumption that each school's LEA is a significant source of variation (i.e., the random effect in the model) between schools and is influential within each school, as well. After Bonferroni correction procedures, results indicated no significant differences in short-term suspension rates for PBIS-implementing versus non-implementing schools. While not significant, coefficient estimates suggest that PBIS-implementing schools had approximately 1.5 more short-term suspensions (per 100 students) than non-implementing schools (see Table 9).

Results of the GEE approach suggested that there were no significant differences in long-term suspension rates among PBIS implementing versus non-implementing schools (see Table 9). It is important to note that there were 42 long-term suspensions, and 97% of elementary

schools included in this analysis reported zero long-term suspension events for the 2015-16 year. The long-term suspension events occurred within 28% of the LEAs in this sample.

Implementation status also did not significantly influence crime and violence rates, as shown in Table 9. Fifty-five percent of schools in the analysis reported zero instances of crime and violence for the 2015-16 year. Of those acts reported, the most common offense types were possession of a weapon (46% of total acts), assault on school personnel (44%), and possession of a controlled substance in violation of the law (6%). Thirty-two percent of elementary schools in this analysis reported at least one weapon possession offense, and 18% of the elementary schools reported at least one assault on school personnel.

Behavioral Outcomes across Ethno-Racial Categories

Initially, Poisson regression analyses were conducted to evaluate the differences in behavioral outcome measures (i.e., suspension rates; crime and violence rates) across ethno-racial categories in PBIS versus non-implementing schools because the dependent variables in these models (distributed across eight ethno-racial categories) were originally presented in frequency counts by race. The total student population values for each ethno-racial category within each school were included as offsets in the Poisson regression analyses, where raw suspension counts and raw crime and violence counts depend on the ethno-racial populations within each school. As with other analyses, the local education agency was included as a random effect in the models. In addition, school was also included as a random effect in the models, to account for the significant source of variation within each school as it pertains to representation across ethno-racial categories.

The original intent was to evaluate differences in behavioral outcomes across all eight ethno-racial categories. However, the principal investigator encountered model convergence problems which led to the decision to include only the following race categories: Black, Hispanic and White students. In total, these students represented 90% of the students originally included in the sample. While it may have been feasible to combine the remaining race categories (i.e., Asian, Indigenous, Pacific Islander, multiracial, and not identified) into one group, this would ignore potentially significant differences in discipline patterns among these groups, as supported by recent research evaluating these outcomes (e.g., Nguyen et al., 2019).

Negative binomial modeling (Payne et al., 2018) did not correct overdispersion in the Poisson models, and hence the models were fit using the GEE approach. Raw coefficient values (in terms of the response scales in log counts of short-term suspensions, long-term suspensions, and log crime and violence acts, per school population) were exponentiated to support meaningful interpretation.

Short-Term Suspensions. Results of the GEE approach revealed that there were no significant differences in short-term suspension rates across White, Hispanic and Black students in PBIS-implementing versus non-implementing schools (see Table 10). However, it is noteworthy that there were significant differences in suspension rates between race categories. Coefficient estimates suggested that the suspension rate for White students in non-implementing schools was approximately 21% of the rate for Black students; the suspension rate for Hispanic students in non-implementing schools was approximately 18% of the rate for Black students. The suspension rate for White students in PBIS-implementing schools was approximately 25% of the rate for Black students; the suspension rate for Hispanic students was approximately 20% of the rate for Black students. Relative risk ratios were calculated by dividing the exponentiated

estimate for Black students by the exponentiated estimate of the compared student group. Within non-implementing schools, relative to White students, Black students were approximately 4.7 times more likely to be suspended. Relative to Hispanic students, Black students were approximately 5.4 times more likely to be suspended. In PBIS-implementing schools, Black students were approximately four times more likely to be suspended than White students, and approximately five times more likely to be suspended than Hispanic students.

Long-Term Suspensions. The GEE method was attempted (with only Black, Hispanic and White students included in the model), however, the principal investigator encountered model convergence problems possibly due to the high number of zero counts included. Of the 3% of elementary schools that reported a long-term suspension for 2015-16, approximately 64% of those suspensions were issued to Black students, 24% were issued to White students, 5% were issued to Hispanic students, and 5% were issued to students who are Indigenous.

Crime and Violence. Results indicated that there were no significant differences in crime and violence rates across White, Hispanic and Black students in PBIS-implementing versus non-implementing schools (see Table 11). However, there were significant differences in reported crime and violence rates across ethno-racial categories. Coefficient estimates suggested that the crime and violence rate in non-implementing schools for White students was approximately 39% of the rate for Black students; the crime and violence rate for Hispanic students was approximately 35% of the rate for Black students. In PBIS-implementing schools, the crime and violence rate for White students was approximately 51% of the rate for Black students; the crime and violence rate for Hispanic students was approximately 35% of the rate for Black students. Relative risk ratios were calculated by dividing the exponentiated estimate for Black students by the estimates of the compared student group. In non-implementing schools,

relative to White students, Black students were approximately 2.6 times more likely to be associated with a crime and violence act. Relative to Hispanic students, Black students were approximately 3.1 times more likely to be associated with a crime and violence act. In PBIS-implementing schools, relative to White students, Black students were approximately two times more likely to be associated with a crime and violence act. Relative to Hispanic students, Black students were approximately 2.9 times more likely to be associated with a crime and violence act.

Office Disciplinary Referrals across Disability Status

Initially, Poisson regression analysis was conducted because the dependent variable in this model (office disciplinary referrals across students with disabilities and nondisabled students) was originally presented in frequency counts. The total number of nondisabled students and number of students with disabilities within each school were included as offsets in the Poisson regression analyses where raw counts depended on the total number of nondisabled students and students with disabilities within each school. As with other Poisson models, the local education agency and school were included as stochastic effects. Negative binomial modeling (Payne et al., 2018) did not correct overdispersion in this Poisson model, and hence the GEE approach was applied. Raw coefficient values (in terms of the response scale in log counts of disciplinary referrals per school population) were exponentiated to support meaningful interpretation.

Results indicated that there were no significant differences in office disciplinary referral rates among students with disabilities and nondisabled students in PBIS-implementing versus non-implementing schools (see Table 12). However, overall, there were significant differences between referral rates for students with disabilities versus their nondisabled peers. It is notable

that in non-implementing schools, the office disciplinary referral rate among nondisabled students was approximately 5% of the rate for students with disabilities. In PBIS-implementing schools, however, the office disciplinary referral rate among nondisabled students was approximately 6% of the rate for students with disabilities. Relative risk ratios were calculated by dividing the exponentiated estimate for students with disabilities by the exponentiated estimate for nondisabled students. Within non-implementing schools, the office disciplinary referral rate was approximately 20 times greater for students with disabilities relative to nondisabled students. Within PBIS schools, the office disciplinary referral rate was approximately 16 times greater for students with disabilities relative to nondisabled students.

Behavioral Outcome Measures within PBIS Schools

The analyses comparing behavioral outcomes across different PBIS implementation levels included 285 PBIS-implementing schools (within 31 LEAs). Following diagnostics which revealed non-normality in the residuals, these models were fit via the GEE approach. Results of these analyses indicated no significant differences in short-term suspension rates (see Table 13), long-term suspension rates (see Table 14), or crime and violence rates (see Table 15), across the three PBIS implementation levels.

Outcomes across Ethno-Racial Categories

Poisson regression analysis was originally conducted to evaluate the effects of PBIS implementation level on short suspensions across three ethno-racial categories (i.e., Black, Hispanic, and White). As with other Poisson models, negative binomial modeling (Payne et al., 2018) did not correct overdispersion, and hence the GEE approach was applied. Raw coefficient values (in terms of the response scale in log counts of short-term suspensions per school

population) were exponentiated to support meaningful interpretation. As shown in Table 16, results indicated that there were no significant differences in short-term suspension rates among White, Hispanic, and Black students across PBIS-implementation levels. Yet there were significant differences in suspension rates between race categories. For instance, within Green Ribbon schools, coefficient estimates suggested that the suspension rate for White students was approximately 28% of the rate for Black students; the suspension rate for Hispanic students was approximately 22% of the rate for Black students. Relative risk ratios were calculated by dividing the exponentiated estimate for Black students by the exponentiated estimate of the compared student group. Within Green Ribbon schools, relative to White students, Black students were approximately 3.6 times more likely to be suspended. Relative to Hispanic students, Black students were approximately 4.5 times more likely to be suspended.

The GEE method was attempted to evaluate differences in long-term suspensions across ethno-racial categories (with only Black, Hispanic and White students included in the model) across PBIS-implementation levels. However, the principal investigator encountered model convergence problems possibly due to the high number of zero counts included. Within this sample, there were 17 long-term suspension events reported by seven elementary schools (3% of this sample), within five LEAs. Of those suspension events, approximately 82% were issued to Black students, 12% were issued to Indigenous students, and 6% were issued to White students.

The GEE method was also attempted to evaluate differences in crime and violence rates across ethno-racial categories (i.e., Black, Hispanic and White students) across PBIS-implementation levels. However, the principal investigator encountered model convergence problems. Within this sample, there were 387 acts of crime and violence, reported by 87% of the LEAs. Of those reported acts, approximately 47% were issued to Black students, 33% were

issued to White students, 10% to Hispanic students, 6% to Multi-Racial students, 3% to Indigenous students, and 0.02% to Asian students.

Outcomes across Disability Status

Originally, Poisson regression analysis was conducted because the dependent variable in this model (office disciplinary referrals across students with disabilities and nondisabled students in PBIS schools) was originally presented in frequency counts. The total number of nondisabled students and number of students with disabilities within each school were included as offsets in the Poisson regression analyses where raw counts depended on the total number of nondisabled students and students with disabilities within each school. As with other Poisson models, overdispersion was not corrected by negative binomial modeling, and so the GEE approach was applied. Raw coefficient values (in terms of the response scale in log counts of disciplinary referrals per school population) were exponentiated to support meaningful interpretation.

Results indicated that there were no significant differences in office disciplinary referral rates among students with disabilities and nondisabled students across PBIS implementation levels (see Table 17). However, it is notable that in schools with Green Ribbon implementation designations, the office disciplinary referral rate among nondisabled students was approximately 7% of the rate for students with disabilities. Relative risk ratios were calculated by dividing the exponentiated estimate for students with disabilities by the exponentiated estimate of the compared student group. Within Green Ribbon schools, the office disciplinary referral rate was approximately 12 times greater for students with disabilities relative to nondisabled students.

Summary of Organizational Health and Behavioral Outcomes

The effect of PBIS status on teacher turnover rates and teacher satisfaction was not significant when other factors (i.e., school population size, concentration of economically disadvantaged students, short-term suspension rates, long-term suspension rates, crime and violence rates) were considered. While PBIS status did not have a significant effect on organizational health outcome measures in this study, the school's concentration of economically disadvantaged students significantly influenced teacher satisfaction and teacher turnover. Several differential effects were also notable. In particular, school population size and suspension rates were significantly associated with teacher satisfaction, but not significantly associated with teacher turnover.

Level of PBIS implementation was significantly associated with teacher satisfaction, but not teacher turnover. As PBIS implementation levels increased (e.g., Green Ribbon to Model), composite teacher satisfaction scores increased. The concentration of economically disadvantaged students and school population were also significantly associated with teacher satisfaction within PBIS-implementing schools.

Analyses revealed no significant differences in short-term and long-term suspension rates, and crime and violence rates, among PBIS-implementing versus non-implementing schools. When stratified across three primary ethno-racial categories, there were no significant differences in short-term suspension or crime and violence rates between PBIS versus non-implementing implementing schools. However, there were notable differences in these behavioral outcomes overall across race categories, with Black students at substantially greater risk for suspension than their White or Hispanic counterparts in PBIS and non-implementing schools. Results indicated that short-term suspension rates among White, Hispanic, and Black

students did not change significantly across PBIS-implementation levels. Yet when compared to their White and Hispanic peers, Black students overall were suspended at substantially higher rates within these PBIS-implementing schools.

There were no significant differences in office disciplinary referral rates of students with and without disabilities in PBIS versus non-implementing schools. Yet students with disabilities within this sample were at greater risk for referrals than their nondisabled peers. Similarly, referral rates among students with disabilities versus non-disabled students did not significantly change across implementation levels, though students with disabilities were referred at substantially higher rates than their nondisabled peers.

Chapter 5: Discussion

The intent of this study was to evaluate whether there were significant differences in organizational health and behavioral outcomes between schools implementing Positive Behavioral Interventions and Supports (PBIS) and non-implementing elementary schools within one state in the 2015-16 school year. Additionally, the aim was to determine whether there were significant differences in these outcomes across implementation designations (i.e., Green Ribbon, Model, and Exemplar). Additional factors beyond PBIS status and implementation level (e.g., school population size, suspension rates, crime and violence rates) were included in the generalized estimating equation (GEE) models and the multiple regression model of organizational health outcomes. Generalized estimating equation analyses for Poisson regression were completed to determine the effects of PBIS implementation on suspension and crime and violence rates across ethno-racial categories, and office disciplinary referral (ODR) rates among students with disabilities versus nondisabled students. Implications and limitations of the present study are addressed below, along with suggestions for future research and implementation of tiered support models to improve organizational health and behavioral outcomes within schools.

Factors Associated with Organizational Health Outcomes

The effect of PBIS status on annual teacher turnover rates and overall teacher satisfaction was not significant when other factors (i.e., school population size, concentration of economically disadvantaged students, short-term suspension rates, long-term suspension rates, crime and violence rates) were considered. This finding is generally inconsistent with research that links PBIS implementation to increased teacher satisfaction with working conditions and environment (cf. Bradshaw et al., 2008; Bradshaw, Koth, et al., 2009; Caldarella et al., 2011;

Houtchens et al., 2017). Houchens et al., for example, evaluated teacher perceptions of working conditions in PBIS and nonPBIS schools via results from the Teaching, Empowering, Leading, and Learning (TELL) Kentucky survey. The NC Teacher Working Conditions survey (from which overall composite mean rate of agreement scores in the present study were analyzed) is based on the TELL survey. Houchens et al. found significant differences between PBIS and nonPBIS schools on two constructs: Managing Student Conduct, and School Leadership. Within the present study, however, teacher satisfaction was represented by the overall composite mean rate of agreement score per school, and teacher responses were not disaggregated across the eight survey constructs. For further context, Bradshaw et al. (2008) found a significant impact of PBIS on the overall Organizational Health Inventory (OHI) score, along with significant effects on several specific aspects (resource influence, staff affiliation, and academic emphasis).

While PBIS status was not significantly associated with differences in organizational health outcomes, the school's concentration of economically disadvantaged students significantly influenced teacher satisfaction and teacher turnover. As schools moved from a lower concentration of economically disadvantaged students to a higher concentration, annual teacher turnover rates increased and overall teacher satisfaction scores decreased. For general comparison, Bradshaw et al. (2008) found a small but significant association between the percentage of students receiving free or reduced-price lunch (FRPL), and the staff affiliation aspect of the OHI. Results from studies outside of the PBIS domain suggest that teacher turnover is influenced by individual factors and organizational factors including the school's concentration of economically disadvantaged students (e.g., Qin, 2019) and more broadly, poor working conditions (e.g., limited resources, limited administrative support, and exclusionary discipline patterns) sometimes associated with higher poverty schools (Simon & Johnson, 2015).

While PBIS status had no significant effect on teacher turnover rates, it is noteworthy that for each additional act of crime and violence (per 100 students), teacher turnover actually decreased by approximately 0.38%. These findings are surprising in light of research which links severe problem behavior (Regan & Michaud, 2011) and crime and violence within schools (Curran et al., 2019; Marinell & Coca, 2013), to teacher turnover. Research also suggests that teachers' perceptions of working conditions (i.e., teacher satisfaction) impact teacher turnover rates (Borman & Dowling, 2008; Ladd, 2011). However, results from the four analyses modeling organizational health outcomes within the present study did not demonstrate a significant relationship between overall teacher satisfaction and teacher turnover.

Within the present study, school population size and short and long-term suspension rates were significantly associated with teacher satisfaction, but not significantly associated with teacher turnover. As short and long-term suspensions increased, teacher satisfaction scores decreased. These findings are generally consistent with research identifying a strong negative correlation between disciplinary exclusion events and teachers' perceptions of school climate (Bear et al., 2014). In the present study, larger schools were associated with lower teacher satisfaction scores, a finding generally supported by other studies evaluating the effects of school size on various student and organizational outcomes (e.g., Leithwood & Jantzi, 2009).

Organizational Health and PBIS Implementation Level

Level of PBIS implementation was significantly associated with teacher satisfaction, but not teacher turnover. As PBIS implementation levels increased (from Green Ribbon to Model, and from Model to Exemplar), composite teacher satisfaction scores increased. For comparison, Houchens et al. (2017) found that the implementation fidelity level (high vs. medium vs. low) significantly influenced outcomes in the areas of Managing Student Conduct, Community

Support and Involvement, and Teacher Leadership. However, it is plausible that schools with more positive working conditions may be predisposed to implement PBIS with greater fidelity. Hence, in the present study (and in Houchens et al.), it is unclear whether PBIS had an effect on teacher satisfaction in those schools.

In the present study, the concentration of economically disadvantaged students and school population size were also significantly associated with teacher satisfaction within PBIS implementing schools. Teacher satisfaction increased in PBIS implementing schools with a low concentration of economically disadvantaged students in comparison to schools with high concentrations of economically disadvantaged students. As the school population increased, overall teacher satisfaction decreased. These findings are consistent with the original model which included PBIS and non-implementing schools.

According to the North Carolina Department of Public Instruction (NCDPI, 2017b), schools achieving Exemplar status for 2015-16 must have completed all three training modules, and must achieve a SET score of 95% on the School-wide Evaluation Tool (SET; Sugai, Lewis-Palmer, et al., 2001) or a Brief SET score of 85%, as well as a score of 80% or better on Tiers I, II, and III of the Tiered Fidelity Inventory (TFI; Algozzine et al., 2014). Schools must also demonstrate an improvement trend across three consecutive years in behavioral and achievement indicators, as well as an improvement trend in at least one additional element (e.g., special education referral information, attendance, and staff retention; NCDPI, 2017b). Therefore, within the elementary schools in the present study, it is possible that schools reached Exemplar implementation status without demonstrating improvement trends in teacher turnover rates (or perhaps with worsening trends in turnover).

Within PBIS implementing schools, no factors (i.e., PBIS implementation level, school population size, concentration of economically disadvantaged students, overall teacher satisfaction, short-term suspension rates, long-term suspension rates, crime and violence rates) were significantly associated with annual teacher turnover rates. This is interesting given that in the GEE model evaluating the effects of PBIS status on teacher turnover (which included all PBIS and non-implementing schools; $n = 835$), the concentration of economically disadvantaged students, and crime and violence rates, were significantly influential on teacher turnover. In comparison, this sample of PBIS implementing schools was smaller ($n = 285$) and represented schools where there were at least two implementation levels achieved within its local education agency (LEA). In terms of overall concentration of economically disadvantaged students within PBIS implementing schools, representation across concentration categories was generally consistent with the original sample containing all PBIS and non-implementing schools (see Table 3). Thirteen percent of PBIS-implementing schools in this model had a high concentration of economically disadvantaged students, 55% represented the mid-high category, 26% represented the mid-low category, and 6% represented the low category. In summary, within the PBIS-implementing schools included in this model, teacher turnover rates were influenced by factors beyond those examined herein.

Factors Associated with Behavioral Outcomes

Analyses revealed no significant differences in short-term and long-term suspension rates, and crime and violence rates, among PBIS-implementing versus non-implementing schools. While the body of comparison studies remains relatively small, these findings are generally inconsistent with studies evaluating aggregate suspension rates between PBIS and non-implementing schools (e.g., Barrett et al., 2008; Childs et al., 2016; Gage et al., 2018;

Notlemeyer et al., 2019). According to NCDPI (2015b), in order to achieve "implementing" status (the Green Ribbon level), a school must satisfy criteria including but not limited to the following: (1) achieve a score of 70% or better for Tier I on the TFI, (2) achieve a total score of 80 or better on the SET, and (3) submit required data (i.e., ODR rates; TFI and SET, or Brief School-wide Evaluation Tool [BSET]) to the DPI. Yet, a school is not required to demonstrate a decreasing trend in suspension rates to be identified as "implementing." Even at the Exemplar implementation designation, schools are not required to demonstrate a decreasing trend in suspension rates (NCDPI, 2017b). Though schools seeking PBIS implementation status in North Carolina undergo training modules that include discussion and analysis of problematic discipline trends (NCDPI, 2015a), it is not surprising that there are no significant differences in suspension outcomes between PBIS and nonPBIS schools in the present study.

It is noteworthy that the crime and violence rates and long-term suspension rates within the elementary level sample in the present study were substantially lower than short-term suspension rates. Approximately 97% of schools issued at least one short-term suspension; 3% of schools issued a long-term suspension; and 45% of schools reported at least one crime and violence act. This is consistent with statewide data on elementary school suspensions and crime and violence rates presented within the original data set that the DPI provided to the university faculty research group.

As demonstrated in the present study, suspension rates and crime and violence rates were not significantly different across PBIS implementation levels (from Green Ribbon, to Model, to Exemplar status). Other studies evaluating implementation effects over time (e.g., Bradshaw, Mitchell, & Leaf, 2010) found that increased fidelity was significantly associated with reductions in short-term suspension and referral rates within elementary schools. The DPI's (2017b) criteria

used to distinguish Exemplar schools indicate that schools must demonstrate a reduction in disciplinary events (e.g., ODRs or suspensions) across at least three years. Thus, it is plausible that the Exemplar schools in the present study could demonstrate a decreasing trend in aggregate ODRs over time while not demonstrating a decreasing trend in aggregate suspensions or crime and violence acts.

Outcomes across Race Categories

When stratified across three primary ethno-racial categories (i.e., Black, Hispanic, and White), there were no significant differences in short-term suspension or crime and violence rates between PBIS versus non-implementing schools. Results also indicated that short-term suspension rates among White, Hispanic, and Black students did not change significantly across PBIS implementation levels. Unfortunately, there were notable differences in these behavioral outcomes overall across race categories, with Black students at substantially greater risk for short-term suspension than their White or Hispanic counterparts in PBIS and nonPBIS implementing schools. Of the long-term suspension events in PBIS and nonPBIS schools in this study, approximately 64% were issued to Black students. However, Black students made up approximately 27% of students within this sample. These findings are generally consistent with studies showing that students of color, and most significantly Black students, continue to encounter disproportionately higher suspension rates in PBIS-implementing schools than their peers (e.g., Skiba et al., 2011; Vincent & Tobin, 2011). However, Gage et al. (2019) found that schools implementing PBIS with fidelity reported statistically significantly fewer suspensions of Black students than non-implementing schools.

As demonstrated in the present study, even when schools implement PBIS with greater fidelity (toward Model or Exemplar levels), Black students continue to be disproportionately

excluded via short-term suspensions. While addressed to some extent within the DPI's training modules (2015a; 2016b), schools are not required to demonstrate reductions in disproportionate disciplinary exclusion patterns to achieve implementation status (at any level of fidelity) for 2015-16.

Outcomes across Disability Status

There were no significant differences in office disciplinary referral rates of students with and without disabilities in PBIS versus non-implementing schools; nor did referral rates among students with disabilities versus nondisabled students significantly change across PBIS implementation levels. Yet, in general, students with disabilities within this sample were at significantly greater risk for referrals than their nondisabled peers. To date, no peer-reviewed, published studies are available that specifically evaluate the effect of PBIS implementation status or fidelity level on ODR rates for students with disabilities versus nondisabled students, yet these findings are consistent with a large body of research identifying the disproportionate disciplinary exclusion of students with disabilities (e.g., McIntosh et al., 2018). In one promising study, Gage et al. (2019) found that schools implementing PBIS with fidelity reported statistically significantly fewer suspensions of students with disabilities than non-implementing schools.

According to the DPI (2015b), in order to achieve implementing status (the Green Ribbon level), a school must submit required data (i.e., ODR rates; TFI and SET, or Brief School-wide Evaluation Tool [BSET]) to the DPI. However, data disaggregation (e.g., across ethno-racial categories and disability status) is not required for a school to meet any implementation designation (Green Ribbon, Model or Exemplar) in 2015-16.

Implications

Findings from the present study and other PBIS studies (e.g., Martinez et al., 2016; Vincent et al., 2012) evaluating the effects of implementation on disciplinary inequity, point to the need for culturally responsive (e.g., Bal et al., 2012) and restorative approaches (e.g., Skiba et al., 2016; Vincent et al., 2016) to serve as the main drivers (rather than the add-ons) of a tiered support framework. There are explicit references to "cultural relevance," "cultural equity," and "culturally valid decision-making," in one slide of the DPI's three-part professional development package on PBIS implementation in North Carolina (DPI, 2016b). Yet the training materials make no reference to guiding staff through a critical and sustained analysis and dismantling of deeply systemic power structures that perpetuate inequity within schools, or unconscious bias and vulnerable decision points (McIntosh et al., 2018). Despite many calls to address systemic issues and to enact culturally responsive PBIS initiatives (e.g., Bal et al., 2012; Swain-Bradway et al., 2014), there is no requirement within the 2015-16 implementation criteria set by DPI (2015b) for schools to demonstrate reductions in disproportionate exclusionary discipline practices. It is important to note that the DPI recently adjusted its criteria (2018b), and schools may still achieve PBIS implementing status (at any level) without demonstrating improvement in these disaggregated trends.

Relative risk ratios calculated in the present study suggest that compared to White students, Black students were at least four times more likely to receive a short-term suspension. These findings are generally consistent with research identifying that the relative risk for suspension varies across North Carolina counties; Black students (in 2015-16) were anywhere from two to over five times more likely to be suspended than their White counterparts (cf. Riddle & Sinclair, 2019). Within some LEAs in North Carolina (e.g., Charlotte-Mecklenburg; Wake) in

2015-16, the relative risk of suspension for Black students was even higher (i.e., over seven times more likely; Southern Coalition for Social Justice, 2017). Notably, in the analysis which included PBIS and non-implementing schools, the ODR rate was approximately 20 times greater for students with disabilities relative to nondisabled students in non-implementing schools; in PBIS-implementing schools, the rate was approximately 16 times greater for students with disabilities. Future research is necessary to evaluate how these ODR rates differ across distinct disability categories and by infraction type. It is also important to assess the extent to which these ODRs among students with disabilities result in out of school suspensions, and the effects of infraction type, race, gender and disability category, on manifestation determination decisions.

As scholars (e.g., Harn et al., 2015; Hawken et al., 2008) and leaders within some states (e.g., NCDPI, 2021) continue to emphasize a shift toward a Multi-Tiered System of Support (MTSS; a tiered model that integrates the academic support structure of Response to Intervention with the behavioral support structure of PBIS), future research should prioritize evaluation of MTSS and PBIS frameworks that merge behavioral science, restorative discipline (e.g., Skiba et al., 2016) and critical multiculturalism (e.g., May & Sleeter, 2010). The primary tools used by schools to assess critical elements of school-wide positive behavior support (e.g., TFI; Algozzine et al., 2014; SET; Sugai, Lewis-Palmer, et al., 2001) or a school's organizational health (e.g., The NC Teacher Working Conditions Survey) do not explicitly assess the extent to which a school prioritizes multiple and diverse perspectives, equitable practices and patterns, or antiracism, for instance. Proponents of culturally responsive PBIS have developed resources to be used in conjunction with these primary assessment tools (see Levenson et al., 2019), but by nature, supplemental materials may be vulnerable to minimization. Hence, additional research is needed to design and validate assessment tools that directly integrate these features. Furthermore,

implementation teams at the school level may need additional support from their district leadership in collecting and disaggregating discipline data (e.g., across race, disability status, disability category) and other critical indicators (e.g., special education referrals, least restrictive environment designations, restraint and seclusion incidents, etc.) to yield formative analysis of existing patterns.

Given the complexity of any school's climate and culture, many PBIS researchers who evaluate the relationship between implementation and organizational health measures (e.g., Bradshaw et al., 2008) acknowledge that there are multiple factors within and beyond the PBIS framework that impact a school's organizational health, and these factors are bidirectional and multifaceted in their influence. Findings from the present study support the assertion that PBIS adoption or enhanced implementation alone is not likely responsible for improvements in a school's organizational health. Future research should continue to evaluate these factors in conjunction with a validated measure of organizational health that encompasses overt indicators of a school's commitment to equity.

Limitations

Existing research suggests that the overwhelming majority of PBIS-implementing schools are elementary and middle school level (Freeman et al., 2016), and that among PBIS schools, different school levels (e.g., elementary versus high school) may adopt and sustain PBIS to differential degrees (McIntosh, Mercer, Nese, & Ghemraoui 2016; Molloy et al., 2013). However, the completed study included only public, traditional elementary schools that served some combination of students within grades pre-k through 5th grade. In addition, this study did not evaluate whether there were differential outcomes between elementary schools serving different grade combinations (e.g., schools serving pre-kindergarten through second grade versus

schools serving third through fifth grades). Furthermore, no alternative elementary schools, private, or charter schools were included in the analyses. Results may not generalize to the entire state of interest within this study (and to other states), given that only 46 LEAs (out of 100 county LEAs and 15 city LEAs) met inclusion criteria. The definitions of high, medium, or low fidelity operationalized in this study (i.e., the DPI's criteria for the three implementation designations) may not wholly align with that of other studies, thereby complicating comparisons and further limiting generalizability.

This study examined data within one school year (2015-16), and thus did not evaluate how behavioral outcomes (e.g., office disciplinary referral rates, suspension rates) and indicators of organizational health (e.g., teacher turnover rates; overall teacher satisfaction with their working environment) might change over time in PBIS versus nonPBIS schools. Some research in this field suggests that implementation across years, and increased implementation fidelity over time, may result in differentially better behavioral outcomes (e.g., Flannery et al., 2014) and organizational health outcomes (e.g., Bradshaw et al., 2008). However, Warren et al. (2003) found that the number of ODRs and school suspensions decreased significantly during the first year of implementation but increased to a level that even exceeded baseline during the second and third years. Given that these data were collected within the 2015-16 year, it is impossible to know whether schools that were identified as non-implementing in this study were PBIS-implementing in any year prior (i.e., and then abandoned implementation toward other school or district initiatives). In addition, according to NCDPI (2016a), schools that received training in PBIS only meet implementation status when they submitted all required data and when those data meet minimum criteria. Hence, it is plausible that some schools identified as non-implementing may contain a number of "ingredients" from the PBIS framework (e.g., training in

and application of evidence-based practices in behavioral intervention; clearly defined schoolwide expectations for behavior; emphasis on reducing exclusionary discipline practices). Furthermore, the schools within this study may have very different baseline levels of school disorder (e.g., higher suspension rates; higher ODR rates; lower teacher satisfaction), which further complicates comparison and generalization. While certainly important, longitudinal evaluation of the effects of implementation on behavioral and organizational health outcomes was beyond the scope of this study.

Though the completed study included such demographic factors as a school's population size and its concentration of economically disadvantaged students, the study did not consider other potentially relevant demographic factors such as geographic positioning of school (e.g., urban versus rural), overall percentage of culturally or linguistically diverse students, or size of the LEA (number of schools). Within the PBIS literature, there is increasing emphasis on identifying systems-level factors that promote and sustain PBIS implementation over time. For example, McIntosh et al. (2016) found that schools in districts with more schools already implementing PBIS, and those starting within a larger initial district cohort, were more likely to adopt and sustain PBIS. The present study did not examine whether there were significant differences in organizational health outcome measures and behavioral outcomes measures between PBIS implementing schools nested within LEAs with a higher concentration of PBIS implementing schools versus a lower concentration. Within the LEAs included in this study, PBIS implementation concentration ranged from 5% to 93% of schools. Given the assertion that the minimum threshold for scaling up occurs when at least 60% of schools within a district are implementing with fidelity (Fixsen et al., 2009), it is notable that only 17% of LEAs in the present study had over 75% concentration.

Evaluation of the relationship between teacher demographics as a percentage of licensed teachers and average years of experience, and PBIS buy-in and implementation, has been addressed in research (e.g., Filter et al., 2016), but was also beyond the scope of the present study. While the completed study evaluated whether there were significant differences in overall composite teacher satisfaction ratings in PBIS-implementing schools versus nonPBIS schools, it does not include data from any student surveys. Of course, students' perceptions of school safety, and their overall satisfaction with the school climate and culture, are highly relevant. Few PBIS studies (e.g., Kelm et al., 2014) have addressed this critical outcome area.

Within the present study, acts of crime and violence were not disaggregated by type (e.g., possession of a weapon, assault resulting in serious injury, assault on school personnel), and so it is unclear whether PBIS implementation status had differential effects on particular crime and violence types. Moreover, this study did not examine whether there were any significant associations between particular types of crime and violence and the organizational health outcomes evaluated within this study (e.g., annual teacher turnover rates; overall teacher satisfaction ratings).

This study focused on the effects of PBIS on such student outcomes as office disciplinary referrals, suspensions, and crime and violence rates, yet other important student outcomes were not evaluated. For example, while PBIS is intended to reduce special education referral rates via its tiered intervention approach (National Education Association, 2014), aggregate referral rates per school (for the 2015-16 year) were not immediately available for analysis in the present study. In this same vein of inquiry, referral rates disaggregated by student ethno-racial category and referred disability category in PBIS and nonPBIS schools were similarly not available. Other student outcomes targeted by the PBIS initiative such as attendance rates (e.g., Freeman et al.,

2015; Freeman et al., 2016), and tardiness (Caldarella et al., 2011), were not evaluated in the present study. The analysis of disciplinary exclusion rates across gender and grade level is an important area of focus (e.g., Martinez et al., 2016), but was also beyond the scope of this study.

This study did not evaluate differences in aggregate ODR rates between PBIS and non-implementing schools (and across implementation levels), nor did it analyze ODR rates disaggregated by race categories, because these data were not available at the point of primary analyses. While several studies reveal positive effects of PBIS implementation on aggregate ODR rates (Barrett et al., 2008; Bradshaw et al., 2012), a number of studies demonstrate that students of color, and most significantly Black students, continue to encounter disproportionately higher ODR rates in PBIS-implementing schools (Bradshaw, Mitchell, O'Brennan, & Leaf, 2010; Martinez et al., 2016; Skiba et al., 2016).

This study evaluated differences in office disciplinary referral rates among students with disabilities versus nondisabled students in PBIS-implementing versus nonPBIS schools. However, results were not disaggregated across specific disability categories. While there is limited research evaluating the effects of PBIS on disciplinary exclusion rates across specific disability categories, the broader base of research on discipline disproportionality suggests that particular groups (e.g., students with emotional disability) are more susceptible to exclusion (Sullivan et al., 2014). This study also did not evaluate behavioral outcomes when stratified across intersecting student factors such as race and disability status. This is an important area of focus, as research suggests that students with disabilities have the highest rates of suspension after Black students and, when combined, the risk increases substantially (Losen et al., 2015).

In research of behavioral outcomes among PBIS implementing versus non-implementing schools (and across PBIS implementation levels), analytical models did not include other factors

(e.g., school population size, teacher turnover rates, concentration of economically disadvantaged students) which may be influential on suspension and crime and violence rates. Several other PBIS studies have evaluated the effects of implementation on behavioral outcomes while controlling for such school characteristics as FRPL percentages (e.g., Flannery et al., 2014; Notlemeyer et al., 2019) or while treating FRPL percentages as status risk factors ultimately shown to be statistically significant predictors of student behavioral outcome variables (Freeman et al., 2015).

Horner et al. (2010) acknowledged that the application of behavioral supports within the PBIS framework does not directly result in improved mathematics, reading, or writing skills. Rather, they suggest that a predictable, consistent, positive, and safe school culture will improve the behavioral engagement of students in learning. Further, if time in class is increased (via reductions in ODRs and suspensions), so too is the opportunity to access instruction. To this point, the majority of research evaluating the impact of PBIS implementation on academic achievement has demonstrated little to no significant effects (Freeman et al., 2016; Gage et al., 2015; Houchens et al., 2017). Results from other studies provide preliminary indications that schools implementing PBIS with fidelity were associated with increased academic performance (Gage et al., 2017; Horner, et al., 2009). While important, an evaluation of academic outcomes (e.g., percentage proficient data; end of grade average scores; Read to Achieve measures) in PBIS versus nonPBIS schools was outside the scope of this study.

Conclusion

While findings from many studies evaluating the effect of PBIS implementation on organizational health indicators and aggregate behavioral outcomes are promising (e.g., Flannery et al.; Freeman et al., 2015; Houchens et al., 2017), there is certainly more work to be done.

Results from the present study suggest that PBIS has not achieved some of its most fundamental objectives in statewide elementary schools in one school year. There were no significant differences in organizational health and behavioral outcomes between schools implementing PBIS versus non-implementing schools, and level of PBIS implementation was not significantly influential on behavioral outcomes. To move toward the *all* that PBIS aspires to impact, continued inquiry is essential to further understand factors within and beyond the PBIS initiative that influence critical outcomes in schools.

REFERENCES

Akin-Little, K. A., Eckert, T. L., Lovett, B. J., & Little, S. G. (2004). Extrinsic reinforcement in the classroom: Bribery or best practice. *School Psychology Review*, *33*(3), 344–362. https://doi.org/10.1080/02796015.2004.12086253

Alber, S. R,, & Heward, W. L. (1997). Recruit it or lose it! Training students to recruit positive teacher attention. *Intervention in School and Clinic, 32*, 275-282. https://doi.org/10.1177/105345129703200504

Alberto, P. A., & Troutman, A. C. (2013). *Applied behavior analysis for teachers*. (9th ed.). Pearson.

Algozzine, B., Horner, R. H., Sugai, G., Barrett, S., Dickey, S. R., Eber, L., Kincaid, D., Lewis, T., & Tobin, T. (2010). *Evaluation blueprint for school-wide positive behavior support.* OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. https://www.pbis.org/resource-type/blueprints

Algozzine, B., Wang, C., White, R., Cooke, N., Marr, M. B., Algozzine, K., Helf, S. S., & Zamora Duran, G. (2012). Effects of multi-tier academic and behavior instruction on difficult-to-teach students. *Exceptional Children, 79,* 45–64. https://doi.org/10.1177/001440291207900103

Algozzine, B., Barrett, S., Eber, L., George, H., Horner, R., Lewis, T., Putnam, B., Swain-Bradway, J., McIntosh, K., & Sugai, G. (2014). *School-wide PBIS Tiered Fidelity Inventory*. OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. https://pbis.org/resource/tfi

American Psychological Association. (2020). Publication manual of the American Psychological Association. (7th ed.). https://doi.org/10.1037/0000165-000

Andreou, T. E., McIntosh, K., Ross, S. W., & Kahn, J. D. (2015). Critical incidents in sustaining school-wide positive behavioral interventions and supports. *The Journal of Special Education, 49*(3), 157-167. https://doi.org/10.1177/0022466914554298

Axelrod, S. (1993). Integrating behavioral technology into public schools. *School Psychology Review*, *8*(1), 1-9. https://doi.org/10.1037/h0088827

Axelrod, S. (1996). What's wrong with behavior analysis? *Journal of Behavioral Education, 6*(3), 247-256. https://doi.org/10.1007/BF02110126

Baer, D. M., Wolf, M. M., & Risley, T. R. (1987). Some still current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 20*, 313-327. https://doi.org/10.1901/jaba.1987.20-313

Bal, A., Thorius, K. K., & Kozleski, E. (2012). Culturally responsive positive behavioral support matters. *Tempe AZ: The Equity Alliance.* http://www.equityallianceatasu.org/sites/default/files/CRPBIS_Matters.pdf

Ballinger, G. (2004). Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods, 7*(2), 127-150. https://doi.org/10.1177/1094428104263672

Bambara, L. M., Goh, A., Kern, L., & Caskie, G. (2012). Perceived barriers and enablers to implementing individualized positive behavior interventions and supports in school settings. *Journal of Positive Behavior Interventions, 14*(4), 228-240. https://doi.org/10.1177/109830071243721

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, *28*, 117-148. https://doi.org/10.1207/s15326985ep2802_3

Banks, T., & Obiakor, F. E. (2015). Culturally responsive positive behavior supports: Considerations for practice. *Journal of Education and Training Studies, 3*(2), 83-90. https://doi.org/10.11114/jets.v3i2.636

Barrett, S. B., Bradshaw, C. P., & Lewis-Palmer, T. (2008). Maryland statewide PBIS initiative: Systems, evaluation, and next steps. *Journal of Positive Behavior Interventions*, *10*(2), 105–114. https://doi.org/10.1177/1098300707312541

Barrish, H. H., Saunders, M., & Wolf, M. W. (1969). Good Behavior Game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis, 2*, 119–124. https://doi.org/10.1901/jaba.1969.2-119

Bear, G. G., Yang, C., Pell, M., & Gaskins, C. (2014). Validation of a brief measure of teachers' perceptions of school climate: Relations to achievement and student suspensions. *Learning Environments Research, 17*, 339-354. https://doi.org/10.1007/s10984-014-9162-1

Blonigen, B., Harbaugh, W., Singell, L., Horner, R. H., Irvin, L. K., & Smolkowski, K. (2008). Application of economic analysis to school-wide positive behavior support (SWPBS) programs. *Journal of Positive Behavior Interventions, 10*(1), 5-19. https://doi.org/10.1177/1098300707311366

Boardman, A. G., Arguelles, M. E., Vaughn, S., Hughes, M. T., & Klingner, J. (2005). Special education teachers' views of research-based practices. *Journal of Special Education*, *39*, 168–180. https://doi.org/10.1177/00224669050390030401

Bohanon, H., Fenning, P., Carney, K., Minnis, M., Anderson-Harriss, S., Moroz, Hicks, K. J., Casper, B. B., Culos, C., Sailor, W., & Piggott, T. (2006). School-wide application of positive behavior support in an urban high school: A case study. *Journal of Positive Behavior Interventions, 8*(3), 131-145. https://doi.org/10.1177/10983007060080030201

Borman, G., & Dowling, N. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research, 78*(3), 367–409. https://doi.org/10.3102/0034654308321455

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. https://doi.org/10.1037/a0012883

Bradshaw, C., Debnam, K., Koth, C.K., & Leaf, P. (2009). Preliminary validation of the implementation phases inventory for assessing fidelity of school-wide positive behavior supports. *Journal of Positive Behavior Interventions*, *11*(3), 145-160. https://doi.org/10.1177/1098300708319126

Bradshaw, C., Koth, C., Thornton, L., & Leaf, P. (2009). Altering school climate through school-wide positive behavioral interventions and supports: Findings from a group-randomized effectiveness trial. *Prevention Science*, *10*, 100–115. https://doi.org/10.1007/s11121-008-0114-9

Bradshaw, C. P., Mitchell, M., & Leaf, P. (2010). Examining the effects of school-wide positive behavioral interventions and supports on student outcomes: Results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions, 12*(3), 133-148. https://doi.org/10.1177/108300709334798

Bradshaw, C. P., Mitchell, M. M., O'Brennan, L. M., & Leaf, P. (2010). Multilevel exploration of factors contributing to the overrepresentation of Black students in office disciplinary referrals. *Journal of Educational Psychology*, *102*(2), 508-520. https://doi.org/10.1037/a0018450

Bradshaw, C. P., & Pas, E. T. (2011). A state-wide scale-up of Positive Behavioral Interventions and Supports (PBIS): A description of the development of systems of support and analysis of adoption and implementation. *School Psychology Review, 40*, 530-548. https://doi.org/10.1080/02796015.2011.12087528

Bradshaw, C. P., Waasdorp, T. E., & Leaf, P. J. (2012). Effects of school-wide positive behavioral interventions and supports on child behavior problems. *Pediatrics*, *130*(5), 1136–1145. https://doi.org/10.1542/peds.digest1305

Bradshaw, C. P., Pas, E. T., Debnam, K. J., & Lindstrom-Johnson, S. (2015). A focus on implementation of positive behavioral interventions and supports (PBIS) in high schools: Associations with bullying and other indicators of school disorder. *School Psychology Review, 44* (4), 480-498. https://doi.org/10.17105/spr-15-0105.1

Bradshaw, C. P., Waasdorp, T. E., & Leaf, P. J. (2015). Examining variation in the impact of school-wide positive behavioral interventions and supports: Findings from a randomized

controlled effectiveness trial. *Journal of Educational Psychology*, *107*(2), 546–557. https://doi.org/10.1037/a0037630

Brophy, J. (1981). Teacher praise: A functional analysis. *Review of Educational Research, 51*, 5–32. https://doi.org/10.3102/00346543051001005

Bruce, P., Bruce, A., & Gedeck, P. 2020. *Practical Statistics for Data Scientists*. (2nd ed.). O'Reilly Media.

Bruning, R., Schraw, G. J., & Norby, M. M. (2011). *Cognitive psychology and instruction*. (5th ed.). Pearson.

Caldarella, P., Shatzer, R. H., Gray, K. M., Young, K. R., & Young, E. L. (2011). The effects of school-wide positive behavior supports on middle school climate and student outcomes. *Research in Middle Level Education, 35*, 1–14. https://doi.org/10.1080/19404476.2011.11462087

Cameron, J., & Pierce, W. D. (1996). The debate about rewards and intrinsic motivation: Protests and accusations do not alter the results. *Review of Educational Research*, *66*(1), 39–51. https://doi.org/10.3102/00346543066001039

Cameron, J., Banko, K. M., & Pierce, W. D. (2001). Pervasive negative effects of rewards on intrinsic motivation: The myth continues. *The Behavior Analyst*, *24*(1), 1-44. https://doi.org/10.1007/BF03392017

Campbell, A., & Anderson, C. (2008). Enhancing effects of Check-in/ Check-out with function-based support. *Behavior Disorders, 33*(4), 233-245. https://doi.org/10.1177/019874290803300404

Carr, E. G., Dunlap, G., Horner, R. H., Koegel, R. L., Turnbull, A. P., & Sailor, W., Anderson, J., Albin, R.W., Koegel, L.K., & Fox, L. (2002). Positive behavior support: Evolution of an applied science. *Journal of Positive Behavior Interventions*, *4*, 4–16. https//doi.org/10.1177/109830070200400102

Chang, M. (2009). An appraisal perspective of teacher burnout: Examining the emotional work of teachers. *Educational Psychology Review, 21*, 193–218. https://doi.org/10.1007/s10648-009-9106-y

Cheatham, J. M., Ozga, J. E., St. Peter, C. C., Mesches, G. A., & Owsiany, J. M. (2017). Increasing class participation in college classrooms with the Good Behavior Game. *Journal of Behavioral Education*, *26*(3), 277–292. https://doi.org/10.1007/s10864-017-9266-7

Childs, K. E., Kinkaid, D., George, H. P., & Gage, N. A. (2016). The relationship between school-wide positive behavioral interventions and supports and student discipline outcomes. *Journal of Positive Behavior Interventions*, *18*(2), 89-99. https://doi.org/10.1177/1098300715590398

Chitiyo, M., May, M., & Chitiyo, G., (2012). An assessment of the evidence-base for school-wide positive behavior support. *Education and Treatment of Children, 35*(1) 1-24.

https://doi.org/10.1353/etc.2012.0000

Coffey, J. H., & Horner, R. H. (2012).The sustainability of positive behavior interventions and supports. *Exceptional Children*, *78*(4), 407-422. https://doi.org/10.1177/001440291207800402

Collins, T. A., Hawkins, R. O., & Flowers, E. M. (2018). Peer-mediated interventions: A practical guide to utilizing students as change agents. *Contemporary School Psychology, 22*(3), 213-219. https://doi.org/10.1007/s40688-017-0120-7

Colvin, G., Sugai, G., & Patching, W. (1993). Precorrection: An instructional approach for managing predictable problem behaviors. *Intervention in School and Clinic, 28*, 143–150. https://doi.org/10.1177/105345129302800304

Colvin, G., Sugai, G., Good III, R. H., & Lee, Y. Y. (1997). Using active supervision and precorrection to improve transition behaviors in an elementary school. *School Psychology Quarterly, 12*, 344. https://doi.org/10.1037/h0088967

Cook, B. G., & Cook, S. C. (2011). Unraveling evidence-based practices in special education. *Journal of Special Education*. *47*(2), 71-82. https://doi.org/10.1177/0022466911420877

Cook, B., & Odom, S. (2013). Evidence-based practices and implementation science in special education. *Exceptional Children*, *79*(2), 135– 144. https://doi.org/10.1177/001440291307900201

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied Behavior Analysis* (2nd ed.). Pearson.

Council for Children with Behavior Disorders. (2020). *Resources for teachers and parents*. https://community.cec.sped.org/ccbd/teacherresources

Council for Exceptional Children. (2014). Council for exceptional children: Standards for evidence-based practices in special education. *TEACHING Exceptional Children, 46*(6), 206-212. https://doi.org/10.1177/0040059914531389

Critchfield, T. (2015). PBIS may not qualify as classical applied behavior analysis. So what? *Behavior Analysis in Practice*, *8*, 99-100. https://doi.org/10.1007/s40617-015-0048-1

Curran, F. C., Viano, S. L., & Fisher, B. W. (2019). Teacher victimization, turnover, and contextual factors promoting resilience. *Journal of School Violence, 18*(1), 21–38. https://doi.org/10.1080/15388220.2017.1368394

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*(6), 627-668. https://doi.org/10.1037%2F0033-2909.125.6.627

Department of Agriculture. (2015). *Child nutrition programs income eligibility guidelines (2015-2016)*. https://www.fns.usda.gov/cn/fr-033115b

Department of Education Office for Civil Rights. (2014). *Data snapshot: School discipline*. (Issue Brief No. 1). https://www2.ed.gov/about/offices/list/ocr/docs/crdc-discipline-snapshot.pdf

Department of Education. (2015). *Digest of Education Statistics, 2015.* https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2016014

Department of Education. (2016a). Assistance to states for the education of children with disabilities; preschool grants for children with disabilities. *Federal Register, 81*(243). https://www.gpo.gov/fdsys/pkg/FR-2016-12-19/pdf/2016-30190.pdf

Department of Education. (2016b). *School climate and discipline: Know the data*. https://www2.ed.gov/policy/gen/guid/school-discipline/data.html

Department of Education. (2020). *North Carolina compilation of school discipline laws and regulations*. https://safesupportivelearning.ed.gov/sites/default/files/discipline-compendium/North%20Carolina%20School%20Discipline%20Laws%20and%20Regulations.pdf

Department of Health and Human Services (2001). *Youth violence: A report of the surgeon general.* Office of the Surgeon General, National Center for Injury Prevention and Control, National Institute of Mental Health, Center for Mental Health Service https://www.ncbi.nlm.nih.gov/books/NBK44294/s.

De Pry, R. L., & Sugai, G. (2002). The effect of active supervision and pre-correction on minor behavioral incidents in a sixth-grade general education classroom. *Journal of Behavioral Education, 11*, 255-264. https://doi.org/10.1023/A:1021162906622

Dixson, A. D., & Rousseau, C. K. (2005). And we are still not saved: Critical race theory in education ten years later. *Race, Ethnicity and Education*, *8*(1), 7-27.

https://doi.org/10.1080/1361332052000340971

Donaldson, J. M., Wiskow, K. M., & Soto, P. L. (2015). Immediate and distal effects of the Good Behavior Game. *Journal of Applied Behavior Analysis, 48*, 685-689.

https://doi.org/10.1002/jaba.229

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), 327-250. https://doi.org/10.1007/s10464-008-9165-0

Emmer, E. T. & Evertson, C. M (1982). Effective management at the beginning of the school year in junior high classes. *Journal of Educational Psychology, 74*, 485–498. https://doi.org/10.1086/461192

Ennis, R. P., Royer, D. J., Lane, K. L., & Griffith, C. E. (2017). A systematic review of precorrection in PK-12 settings. *Education & Treatment of Children*, *40*(4), 465-495. https://doi.org/10.1353/etc.2017.0021

Fallon, L. M, & Mueller, M. R. (2017). Culturally responsive wraparound supports: Collaborating with families to promote students' behavior regulation across settings. *Contemporary School Psychology*, *21*, 201-210. https://doi.org/10.1007/s40688-016-0116-8

Ferguson, E. & Houghton, S. (1992). The effects of contingent teacher praise, as specified by Canter's Assertive Discipline Programme, on children's on-task behaviour. *Educational Studies, 18*(1), 83-93. https://doi.org/10.1080/0305569920180108

Feuerborn, L. L., Wallace, C., & Tyre, A. D. (2016). A qualitative analysis of middle and high school teacher perceptions of schoolwide positive behavior supports. *Journal of Positive Behavior Interventions, 18*(4), 219-229. https://doi.org/10.1177/1098300716632591

Fielding, C., Lowdermilk, J., Lanier, L., Fannin, A. G., Schkade, J., Rose, C., & Simpson, C. G. (2013, Spring/Summer). Applied behavior analysis: Current myths in education. *Journal of the American Academy of Special Education Professionals*, 83-97.

Filter, K. J., McKenna, M. K., Benedict, E. A., Horner, R. H., Todd, A. W., & Watson, J. (2007). Check in/Check out: A post-hoc evaluation of an efficient, secondary level targeted intervention for reducing problem behaviors in schools. *Education and Treatment of Children*, *30*(1), 69–84. https://doi.org/10.1353.etc.2007.0000

Filter, K. J., Sytsma, M. R., & McIntosh, K. (2016). A brief measure of staff commitment to implement school-wide positive behavior interventions and supports. *Assessment for Effective Intervention, 42*(1), 18-31. https://doi.org/10.1177/1534508416642212

Fixsen, D. L., Naoom, S. F., Blasé, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature.* Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, National Implementation Research Network.

Fixsen, D. L., Blase, K. A., Horner, R., & Sugai, G. (2009). *Scaling up evidence-based practices in education.* Frank Porter Graham Child Development Institute, University of North Carolina at Chapel Hill. (Brief No. 1). https://nirn.fpg.unc.edu/resources/scaling-evidence-based-practices-education

Fixsen, D., Blase, K., Metz, A., & Van Dyke, M. (2013). Statewide implementation of evidence-based programs. *Exceptional Children*, *79*(2), 213-230.

https://doi.org/10.1177/001440291307900206

Flannery, K. B., Frank, J., Kato, M., Doren, B., & Fenning, P. (2013). Implementing school-wide positive behavior support in high school settings: Analysis of eight high schools. *The High School Journal, 96*, 267–282. https://doi.org/10.1353/hsj.2013.0015

Flannery, K. B., Fenning, P., Kato, M. M., & McIntosh, K. (2014). Effects of school-wide positive behavioral interventions and supports and fidelity of implementation on problem behavior in high schools. *School Psychology Quarterly*, *29*, 111–124. https://doi.org/10.1037/spq0000039

Flintoff, A., Dowling, F., & Fitzgerald, H. (2015). Working through whiteness, race and (anti) racism in physical education teacher education. *Physical Education and Sport Pedagogy*, *20*(5), 559-570. https://doi.org/10.1080/17408989.2014.962017

Flower, A., McKenna, J., Muething, C. S., Bryant, D. P., & Bryant, B. R. (2014). Effects of the Good Behavior Game on classwide off-task behavior in a high school basic algebra resource classroom. *Behavior Modification, 38,* 45–68. https://doi.org/10.1177/0145445513507574

Forman, S. G., Olin, S.S., Hoagwood, K. E., Crowe, M., & Saka, N. (2009). Evidence-based interventions in schools: Developers views of implementation barriers and facilitators. *School Mental Health*, *1*, 26-36. https://doi.org/10.1007/s12310-008-9002-5

Frank, J. L., Horner, R. H., & Anderson, C. M. (2009). *Influence of school level socioeconomic status and racial diversity on Schoolwide Positive Behavior Support implementation*. OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. https://www.pbis.org/resource-type/evaluation-briefs

Frank Porter Graham Child Development Institute. (n.d.) State implementation and scaling up of evidence-based practices (SISEP) assistance center. http://sisep.fpg.unc.edu/

Freeman, J., Simonsen, B., McCoach, D. B., Sugai, G., Lombardi, A., & Horner, R. (2015). An analysis of the relationship between school-wide positive behavior interventions and supports and high school dropout rates. *The High School Journal*, *98*(4), 290-315. https://doi.org/10.1353/hsj.2015.0009

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

Gable, R. A., Hester, P. H., Rock, M. L., & Hughes, K. G. (2009). Back to basics: Rules, praise, ignoring, and reprimands revisited. *Intervention in School and Clinic, 44*(4), 195–205. https://doi.org/10.1177/1053451208328831

Gage, N. A., Freeman, J., Horner, R. H., Sugai, G., Lewis, T. J., & Strickland-Cohen, K. (2014). Funding support for school-wide positive behavioral interventions and support implementation and scale-up: A descriptive study of nine states. *Journal of Disability Policy Studies, 25*(3), 164-174. https://doi.org/10.1177/1044207313488943

Gage, N. A., Sugai, G., Lewis, T. J., & Brzozowy, S. (2015). Academic achievement and school-wide positive behavior supports. *Journal of Disability Policy Studies*, *25*(4), 2015. https://doi.org/10.1177/1098300715580992

Gage, N. A., Larson, A., Sugai, G., & Chafouleas, S. M. (2016). Student perceptions of school climate as predictors of office discipline referrals. *American Educational Research Journal, 53*, 492–515. https://doi.org/10.3102/0002831216637349

Gage, N. A., Leite, W. L., Childs, K., & Kincaid, D. (2017). Average treatment effect of school-wide positive behavioral interventions and supports on school-level academic achievement in Florida. *Journal of Positive Behavior Interventions, 19*(3), 158-167. https://doi.org/10.1177/1098

Gage, N. A., Whitford, D. K., & Katsiyannis, A. (2018). A review of schoolwide positive behavior interventions and supports as a framework for reducing disciplinary exclusions.

Journal of Special Education, *52*(3), 142–151. https://doi.org/10.1177/0022466918767847

Gage, N. A., Grasley-Boy, N., Peshak George, H., Childs, K., & Kincaid, D. (2019). A quasi-experimental design analysis of the effects of school-wide positive behavior interventions and supports on discipline in Florida. *Journal of Positive Behavior Interventions*, *21*(1), 50–61. https://doi.org/10.1177/1098300718768208

Gay, G. (2002). Culturally responsive teaching in special education for ethnically diverse students: Setting the stage. *Qualitative Studies in Education, 15*(6), 613-629.

https://doi.org/10.1080/0951839022000014349

George, H. P., & Kincaid, D. K. (2008). Building district-level capacity for positive behavior support. *Journal of Positive Behavior Interventions, 10*, 20–32. https://doi.org/10.1177/1098300707311367

Gersten, R., Euchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, *71*, 149-164. https://doi.org./10.1177/001440290507100202

Gion, C. M., McIntosh, K., & Horner, R. (2014). *Patterns of minor office discipline referrals in schools using SWIS*. OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. https://www.pbis.org/resource-type/evaluation-briefs

Girvan, E. J., Gion, C., McIntosh, K., & Smolkowski, K. (2017). The relative contribution of subjective office referrals to racial disproportionality in school discipline. *School Psychology Quarterly, 32*, 392–404. https://doi.org/10.1037/spq0000178

Green, A., Nese, R., McIntosh, K., Nishioka, V., Eliason, B., & Delabra, A. C. (2015). *Key elements of discipline policy to address disproportionality: A guide for district and school teams*. OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. https://www.pbis.org/school/equity-pbis

Gresham, F. M., Van, M. B., & Cook, C. R. (2006). Social skills training for teaching replacement behaviors: Remediating acquisition deficits in at-risk students. *Behavioral Disorders*, *31*(4), 363–377. https://doi.org/10.1177/019874290603100402

Halekoh, S., Hojsgaard, U., & Yan, J. (2006). The R package geepack for generalized estimating equations. *Journal of Statistical Software, 15*(2), 1-11. https://doi.org/10.18637/jss.v015.i02

Han, S. S., & Weiss, B. (2005). Sustainability of teacher implementation of school-based mental health programs. *Journal of Abnormal Child Psychology*, *33*(6), 665-679. https://doi.org/10.1007/s10802-005-7646-2

Harn, B., Basaraba, D., & Chard, D. (2015). The impact of schoolwide prevention efforts: Lessons learned from implementing independent academic and behavioral support systems. *Learning Disabilities, 13*(1), 3-20.

Harris, M.B. (1998). *Basic statistics for behavioral science research*. (2nd ed.). Allyn & Bacon.

Hawken, L. S., Vincent, C. G., & Schumann, J. (2008). Response to intervention for social behavior: Challenges and opportunities. *Journal of Emotional and Behavioral Disorders*, *16*, 213-225. https://doi.org/10.1177/1063426608316018

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, *71*(2), 165-179. https://doi.org/10.1177/001440290507100203

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*, *11*(3), 133-144. https://doi.org/10.1177/1098300709332067

Horner, R. H., Sugai, G., & Anderson, C.M. (2010). Examining the evidence base for school-wide positive behavior support. *Focus on Exceptional Children*, *42*(8), 1-14. https://doi.org/10.1002/9781118660584.ese1902

Horner, R. H., Sugai, G., Kinaid, D., George, H., Lewis, T., Eber, L., Barrett, S., & Algozzine, B. (2012). *What does it cost to implement school-wide PBIS?* OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. https://www.pbis.org/resource-type/evaluation-briefs

Horner, R. H., Kincaid, D., Sugai, G., Lewis, T., Eber, L., Barrett, S., Rossetto Dickey, C., Richter, M., Sullivan, E., Boezio, C., Algozzine, B., Reynolds, H., & Johnson, N. (2014). Scaling up school-wide positive behavioral interventions and supports: Experiences of seven states with documented success. *Journal of Positive Behavior Interventions, 16*, 197–208. https://doi.org/10.1177/1098300713503685

Horner, R. H., Sugai, G., & Fixsen, D. L. (2017). Implementing effective educational practices at scales of social importance. *Clinical Child and Family Psychology Review*, *20*(1), 25-35. https://doi.org/10.1007/s10567-017-0224-7

Houchens, G. W., Zhang, J., Davis, K., Niu, C., Chon, K. H., & Miller, S. (2017). The impact of positive behavior interventions and supports on teacher perceptions of teaching conditions and student achievement. *Journal of Positive Behavior Interventions*, *19*(3), 168-179. https://doi.org/10.1177/1098300717696938

Howard, G. (2006). *We can't teach what we don't know: White teachers, multiracial schools.* Teachers College Press.

Hoy, W., & Feldman, J. (1987). Organizational health: The concept and its measure. *Journal of Research and Development in Education*, *20*, 30-37.

Individuals with Disabilities Education Improvement Act of 2004, Pub. L. No. 108-446; 20 U.S.C. §§ 1401 et seq. (2004). https://sites.ed.gov/idea/regs/

Ingersoll, R. M., & Smith, T. M. (2003). The wrong solution to the teacher shortage. *Educational Leadership, 60*(8), 30–33.

Irvin, L. K., Horner, R. H., Ingram, K., Todd, A. W., Sugai, G., Sampson, N. K., & Boland, J. B. (2006). Using office discipline referral data for decision making about student behavior in elementary and middle schools: An empirical evaluation study. *Journal of Positive Behavior Interventions, 8*,10–23. https://doi.org/10.1177/10983007060080010301

Jenkins, L. N., Floress, M. T., & Reinke, W. (2015). Rates and types of teacher praise: A review and future directions. *Psychology in the Schools*, *52*(5), 463–476. https://doi.org/10.1002/pits.21835

Kelm, J. L., & McIntosh, K. (2012). Effects of school-wide positive behavior support on teacher self-efficacy. *Psychology in The Schools*, *49*(2), 137-147. https://doi.org/10.1002/pits.20624

Kelm, J. L., McIntosh, K., & Cooley, S. (2014). Effects of implementing school-wide positive behavioural interventions and supports on problem behavior and academic achievement in a Canadian elementary school. *Canadian Journal of School Psychology, 29*(3), 195-212. https://doi.org./10.1177/0829573514540266

Kincaid, D., Childs, K., & George, H. P. (2010). *Tier 1 Benchmarks of Quality (Revised*). Department of Child and Family Studies, University of South Florida, Tampa. https://www.pbis.org/resource/boq

Klingner, J. K., Ahwee, S., & Pilonieta, P. (2003). Barriers and facilitators in scaling up research-based practices. *Exceptional Children*, *69*, 411-429. https://doi.org/10.1177/001440290306900402

Ladd, H. F. (2011). Teachers' perceptions of their working conditions. *Educational Evaluation and Policy Analysis, 33*(2), 235–261. https://doi.org/10.3102/0162373711398128

Ladson-Billings, G. (1998). Just what is critical race theory and what's it doing in a nice field like education? *Qualitative Studies in Education*, *11*(1), 7-24. https://doi.org/10.1080/095183998236863

Landrum, T. J., & Kauffman, J. M. (2006). Behavioral approaches to classroom management. In C.M. Evertson & C.S. Weinstein (Eds.), *Handbook of Classroom Management: Research, Practice and Contemporary Issues.* (pp. 47-71). Routledge.

Lane, K. L., Robertson, E. J., & Graham-Bailey, M. A. L. (2006). An examination of school-wide interventions with primary level efforts conducted in secondary schools: Methodological considerations, applications of research methodology. *Advances in*

Learning and Behavioral Disabilities, 19, 157-199. https://doi.org/10.1016/ S0735-004X(06)19007-2

Leithwood, K., & Jantzi, D. (2009). A review of empirical evidence about school size effects: A policy perspective. *Review of Educational Research*, *79*(1), 464–490. https://doi.org/10.3102/0034654308326158

Lepper, M. R., & Keavney, M., & Drake, M. (1996). Intrinsic motivation and extrinsic rewards: A commentary on Cameron and Pierce's meta-analysis. *Review of Educational Research*, *66*(1), 5–32. https://doi.org/10.3102/00346543066001005

Leverson, M., Smith, K., McIntosh, K., Rose, J., & Pinkelman, S. (2016). *PBIS cultural responsiveness field guide: Resources for trainings and coaches*. OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. https://www.pbis.org/resource/pbis-cultural-responsiveness-field-guide-resources-for-trainers-and-coaches

Litow, L., & Pumroy, D. K. (1975). A brief review of classroom group-oriented contingencies. *Journal of Applied Behavior Analysis, 8*, 341–347. https://doi.org/10.1901/jaba.1975.8-341

Losen, D. J., Ee, J., Hodson, C., & Martinez, T. E. (2015). Disturbing inequities: Exploring the relationship of discipline disparities for students with disabilities by race with gender with school outcomes. In D. J. Losen (Ed.), *Closing the school discipline gap: Equitable remedies for excessive exclusion* (pp. 89–106). Teacher's College Press.

Luiselli, J. K., Putnam, R. F., Handler, M. W., & Feinberg A. B. (2005). Whole-school Positive Behavior Support: Effects on student discipline problems and academic performance.

Educational Psychology, *25*(2-3), 183-198. https://doi.org/10.1080/0144341042000301265

Maag, J. W. (2002). Rewarded by punishment: Reflections on the disuse of positive reinforcement in schools. *Exceptional Children, 61*, 173–186. https://doi.org/10.1177/001440290106700203

Marinell, W. H., & Coca, V. M. (2013). *Who stays and who leaves? Findings from a three-part study of teacher turnover in NYC middle schools*. Research Alliance for NYC Schools. https://steinhardt.nyu.edu/research-alliance/research/publications/who-stays-and-who-leaves

Marshall, P. L. (2002). Racial identity and challenges of educating white youths for cultural diversity. *Multicultural Perspectives*, *4*(3), 9-14. https://doi.org/10.1207/S15327892MCP0403_3

Martens, B. K., & Lambert, T. L. (2014). Conducting functional behavior assessments for students. In Walker, H. M., & Gresham, F. M. (Eds.), *Handbook of evidence-based practices for emotional and behavioral disorders: Applications in schools*. (243-261). The Guilford Press.

Martinez, A., McMahon, S. D., & Treger, S. (2016). Individual- and school-level predictors of student office disciplinary referrals. *Journal of Emotional and Behavioral Disorders, 24*(1), 30-41. https://doi.org/10.1177/1063426615588289

Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Organizational Behavior, 2*(2), 99-113. https://doi.org/10.1002/job.4030020205

Mathews, S., McIntosh, K., Frank, J. L., & May, S. L. (2014). Critical features predicting sustained implementation of school-wide positive behavioral interventions and supports. *Journal of Positive Behavior Interventions*, *16*(3), 168-178. https://doi.org/10.1177/1098300713484065

Matias, C. E., Viesca, K. M., Garrison-Wade, D. F., Tandon, M., & Galindo, R. (2014). "What is critical whiteness doing in OUR nice field like critical race theory?" Applying CRT and CWS to understand the white imaginations of white teacher candidates. *Equity & Excellence in Education*, *47*(3), 289-304. https://doi.org/10.1080/10665684.2014.933692

McCrary, D., Lechtenberger, D., & Wand, E. (2012). The effect of schoolwide positive behavioral supports on children in impoverished rural community schools. *Preventing School Failure, 56* (1) 1-7. https://doi.org/10.1080/1045988X.2010.548417

McFarland, J., Hussar, B., Wang, X., Zhang, J., Wang, K., Rathburn, A., Barmer, A., Forrest Cataldi, E., & Bullock Mann, F. (2018). *The condition of education: 2018*. National Center for Education Statistics. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018144

McIntosh, K., Horner, R. H., & Sugai, G. (2009). Sustainability of systems-level evidence-based practices in schools: Current knowledge and future directions. In W. Sailor, G. Dunlap, G. Sugai & R. H. Horner (Eds.), *Handbook of positive behavior support* (pp. 327–352). Springer.

McIntosh, K., MacKay, L., Hume, A., Doolittle, J., Vincent, C. G., Horner, R. H., & Ervin, R. A. (2011). Development and initial validation of a measure to assess factors related to sustainability of school-wide positive behavioral interventions and supports. *Journal of*

Positive Behavior Interventions, 13(4), 208-218. https://doi.org/10.1177/1098300710385348

McIntosh, K., Barnes, A., Eliason, B., & Morris, K. (2014). *Using discipline Data within SWPBIS to identify and address disproportionality: A guide for school teams*. OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. https://www.pbis.org/Common/Cms/files/pbisresources/PBIS_Disproportionality_Data_Guidebook.pdf

McIntosh, K., Predy, L. K., Upreti, G., Hume, A. E., Turri, M. G., & Mathews, S. (2014). Perceptions of contextual features related to implementation and sustainability of school-wide positive behavior support. *Journal of Positive Behavior Interventions*, *16*(1), 31–43. https://doi.org/10.1177/1098300712470723

McIntosh, K., Kim, J., Mercer, S. H., Strickland-Cohen, M. K., & Horner, R. H. (2015). Variables associated with enhanced sustainability of school-wide positive behavioral interventions and supports. *Assessment for Effective Intervention, 40*(3), 184-191. https://doi.org/10.1177/1534508414556503

McIntosh, K., Mercer, S. H., Nese, R. N., & Ghemraoui, A. (2016). Identifying and predicting distinct patterns of implementation in a school-wide positive behavior support framework. *Society for Prevention Research, 17,* 992-1001. https://doi.org/10.1007/s11121-016-0700-1

McIntosh, K., Mercer, S. H., Nese, R. N., Strickland-Cohen, M. K., & Hoselton, R. (2016). Predictors of sustained implementation of school-wide positive behavioral interventions and supports. *Journal of Positive Behavior Interventions, 18*(4), 209-218. https://doi.org/10.1177/1098300715599737

McIntosh, K., Ellwood, K., McCall, L., & Girvan, E. J. (2018). Using discipline data to enhance equity in school discipline. *Intervention in School and Clinic*, *53*(3), 146–152. https://doi.org/10.1177/1053451217702130

Meyers, D. C., Durlak, J. A., & Wandersman, A. (2012). The quality implementation framework: A synthesis of critical steps in the implementation process. *American Journal of Community Psychology*, *50*(3–4), 462–480. https://doi.org/10.1007/s10464-012-9522-x

Mitchell, R. R., Tingstrom, D. H., Dufrene, B. A., Ford, W. B., & Sterling, H. E. (2015). The effects of the Good Behavior Game with general-education high school students. *School Psychology Review*, *44*(2), 191–207. https://doi.org/10.17105/spr-14-0063.1

Molloy, L. E., Moore, J. E., Trail, J., Van Epps, J. .J., & Hopfer, S. (2013). Understanding real-world implementation quality and "active ingredients" of PBIS. *Society for Prevention Research, 14*, 593-605. https://doi.org/10.1007/s11121-012-0343-9

Moore-Partin, T. C., Robertson, R. E., Maggin, D. M., Oliver, R. M., & Wehby, J. H. (2010). Using teacher praise and opportunities to respond to promote appropriate student behavior. *Preventing School Failure*, *54*(3), 172–178. https://doi.org/10.1080/10459880903493179

Mongan, P., & Walker, R. (2012). ''The road to hell is paved with good intentions" : A historical, theoretical, and legal analysis of zero-tolerance weapons policies in American schools. *Preventing School Failure*, *56*, 232–240. https://doi.org/10.1080/1045988.2011.654366

Nakagawa, S., & Schielzeth, H. (2013). A simple and general method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133-142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

National Association of School Psychologists. (2001). *Zero tolerance and alternative strategies: A fact sheet for educators and policy makers.* http://ccsd.net/internal/cms/doc-vault/resources/archive/zero.tolerance.fact.sheet.pdf

National Education Association. (2014). *Early childhood positive behavioral interventions and supports: A multi-tiered framework that works for every student.* [Policy brief]. https://www.nea.org/advocating-for-change/new-from-nea/positive-behavioral-interventions-and-supports

National Professional Development Center on Autism Spectrum Disorders. (n.d.). *Evidence-based practices*. https://autismpdc.fpg.unc.edu/evidence-based-practices

Nelson, J. R., Colvin, G., & Smith, D. J. (1996). The effects of setting clear standards on students' social behavior in common areas of the school. *The Journal of At-Risk Issues, 3(1),* 10-17.

Nelson, J. R., & Roberts, M. L. (2000). Ongoing reciprocal teacher-student interactions involving disruptive behaviors in general education classrooms. *Journal of Emotional and Behavioral Disorders*, *8*(1), 27–37. https://doi.org/10.1177/106342660000800104

Nelson, J. R., Martella, R. M., & Marchand-Martella, N. (2002). Maximizing student learning: The effects of a comprehensive school-based program for preventing problem behaviors. *Journal of Emotional and Behavioral Disorders, 10*, 136– 148. https://doi.org/10.1177/10634266020100030201

New Hanover County Schools. (2018). *Policy 8410: Policies, rules and procedures relative to student discipline in the New Hanover County Schools*. https://boardpolicyonline.com/bl/?b=hanover_county_old

Nguyen, B. M. D., Noguera, P., Adkins, N., & Teranishi, R. T. (2019). Ethnic discipline gap: Unseen dimensions of racial disproportionality in school discipline. *American Educational Research Journal,* 1-29. https://doi.org/10.3102/0002831219833919

Nordstrum, L. E., LeMahieu, P. G., & Berrena, E. (2017). Implementation science: Understanding and finding solutions to variation in program implementation. *Quality Assurance in Education: An International Perspective*, *25*(1), 58–73. https://doi.org/10.1108/QAE-12-2016-0080

North Carolina Department of Public Instruction. (n.d.). *Student accounting data.* https://www.dpi.nc.gov/districts-schools/district-operations/financial-and business-services/demographics-and-finances/student-accounting-data

North Carolina Department of Public Instruction. (2015a). *PBIS Module 1.* [PowerPoint slides]. Retrieved October, 2018, from http://behaviorsupport.ncdpi.wikispaces.net/Professional%20Development

North Carolina Department of Public Instruction. (2015b). *Implementation status criteria*. Retrieved October 2018, from http://behaviorsupport.ncdpi.wikispaces.net/What+is+PBIS%3FAs

North Carolina Department of Public Instruction. (2015c). *Current facts and research*. [PowerPoint slides]. https://ec.ncpublicschools.gov/conferences-profdev/annual-conference/2015/conference-materials/session-08.pdf

North Carolina Department of Public Instruction. (2016a). *NC PBIS state recognition criteria.* Retrieved October 2018, from http://behaviorsupport.ncdpi.wikispaces.net/What+is+PBIS%3FAs

North Carolina Department of Public Instruction. (2016b). *PBIS Module 2.* [PowerPoint slides]. Retrieved October, 2018, from http://behaviorsupport.ncdpi.wikispaces.net/Professional%20Development

North Carolina Department of Public Instruction. (2017a). *PBIS overview*. [PowerPoint Slides]. Retrieved October 2018, from http://behaviorsupport.ncdpi.wikispaces.net/What+is+PBIS%3FAs

North Carolina Department of Public Instruction. (2017b). *Checklist for exemplar recognition*. Retrieved October 2018, from http://behaviorsupport.ncdpi.wikispaces.net/Process+and+Criteria

North Carolina Department of Public Instruction (2017c). *Report to the North Carolina general assembly: Consolidated data report 2015-16*. https://www.dpi.nc.gov/media/6826/download

North Carolina Department of Public Instruction (2018a). *16 reportable criminal offenses.* https://www.dpi.nc.gov/data-reports/dropout-and-discipline-data/16-reportable-criminal-offenses

North Carolina Department of Public Instruction. (2018b). *2017-2018 PBIS recognition changes and timeline*. [Memorandum]. https://ec.ncpublicschools.gov/policies/nc-policies-governing-services-for-children-with-disabilities/ncdpi-communication/2017-2018/ec-division-memos/memo-pbis-changes-2018.pdf/view

North Carolina Department of Public Instruction. (2020a). *North Carolina policies governing services for children with disabilities*. https://ec.ncpublicschools.gov/policies/nc-policies-governing-services-for-children-with-disabilities

North Carolina Department of Public Instruction. (2020b). *Regional consultants*. https://ec.ncpublicschools.gov/directory/regional-consultants

North Carolina Department of Public Instruction. (2021). *Integrated academic and behavior systems.* https://www.dpi.nc.gov/districts-schools/districts-schools-support/district-and-regional-support/integrated-academic-behavior-systems

North Carolina General Assembly. (2011). N.C. Gen. Stat. § 115C-390.1 *Discipline: State policy and definitions.* https://www.ncleg.net/EnactedLegislation/Statutes/HTML/ByArticle/Chapter_115C/Article_27.html

Notlemeyer, A., Palmer, K., James, A., & Petrasek, M. (2019). Disciplinary and achievement outcomes associated with school-wide positive behavioral interventions and supports implementation level. *School Psychology Review*, *48*(1), 81-87 https://doi.org/10.17105/SPR-2017-0131.V48-1

O'Brien, S., & Repp, A. C. (1990). Reinforcement-based reductive procedures: A review of 20 years of their use with persons with severe or profound retardation. *Journal of the Association for Persons with Severe Handicaps (JASH)*, *15*(3), 148–159. https://doi.org/10.1177/154079699001500307

Odom, S. L. (2009). The ties that bind: Evidence-based practices, implementation science, and outcomes for children. *Topics in Early Childhood Special Education*, *29*, 53–61. https://doi.org/10.1177/0271121408329171

Office of Special Education (OSEP) Technical Assistance Center. (2021). *What is tier 1 support?* Positive Behavioral Interventions and Supports. https://www.pbis.org/pbis/tier-1

Pas, E. T., Waasdorp, T. E., & Bradshaw, C. P. (2014). Examining contextual influences on classroom-based implementation of positive behavior support strategies: Findings from a randomized controlled effectiveness trial. *Society for Prevention Research, 16*, 1096-1106. https://doi.org/10.1007/s11121-014-0492-0

Payne, E. B., Gebregziabher, M., Hardin, J. W, Ramakrishnan, V., & Edege, L. E. (2018). An empirical approach to determining a threshold for assessing overdispersion in Poisson and negative binomial models in count data. *Communication in Statistics: Simulation and Computation, 47*(6), 1722-1738. https://doi.org/10.1080/03610918.2017.1323223

Petscher, E. S., Rey, C., & Bailey, J. S. (2009). A review of empirical support for differential reinforcement of alternative behavior. *Research in Developmental Disabilities: A Multidisciplinary Journal*, *30*(3), 409–425. https://doi.org/10.1016/j.ridd.2008.08.008

Public Schools of Robeson County. (2012). *Student code of conduct*. http://robesonnc.apptegy.us/o/robeson/browse/142025

Qin, L. (2019). Factors relating to teachers' intention to change school: A multilevel perspective. *Policy Futures in Education, 17*(3), 318–338. https://doi.org/10.1177/1478210318822184

Regan, K. S., & Michaud, K. M. (2011). Best practices to support behavior. *Beyond Behavior, 20*(2), 40-47.

Reinke, W. M., Herman, K. C., & Stormont, M. (2013). Classroom-level positive behavior supports in schools implementing SW-PBIS: Identifying areas for enhancement. *Journal*

of Positive Behavior Interventions, 15(1), 39-50. https://doi.org/10.1177/1098300712459079

Riddle, T., & Sinclair, T. (2019). Racial disparities in school-based disciplinary actions are associated with county level rates of racial bias. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 116*(17), 8255-8260. https://doi.org/10.1073/pnas.1808307116

Robacker, C. M., Rivera, C. J., & Warren, S. H. (2016). A token economy made easy through ClassDojo. *Intervention in School and Clinic*, *52*(1), 39–43. https://doi.org/10.1177/1053451216630279

Rogers, E. M. (2003). *Diffusion of innovations*. The Free Press.

Ross, S. W., & Horner, R. H. (2007). Teacher outcomes of schoolwide positive behavior support. *Teaching Exceptional Children Plus*, 3*(6)*, 1-13.

Ross, S., Romer, N., & Horner, R. H. (2012). Teacher well-being and the implementation of school-wide positive behavior interventions and supports. *Journal of Positive Behavior Interventions. 14*(2) 118-128. https://doi.org/10.1177/109830071141382

Royer, D. J., Lane, K. L., Dunlap, K., & Ennis, R. P. (2018). A systematic review of teacher-delivered behavior-specific praise on k-12 student performance. *Remedial and Special Education*, *40*(2), 112-128. https://doi.org/10.1177/0741932517751054

Sailor, W., Wolf, N., Choi, H., & Roger, B. (2009). Sustaining positive behavior support in a context of comprehensive school reform. In W. Sailor, G. Dunlap, G. Sugai & R. H. Horner (Eds.), *Handbook of positive behavior support* (pp. 633-669). Springer.

Simon, N. S., & Johnson, S. M. (2015). Teacher turnover in high-poverty schools: What we know and can do. *Teachers College Record, 117*(3), 1-36.

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education & Treatment of Children, 31*, 351–380. https://doi.org/10.1353/etc.0.0007

Simonsen, B., Eber, L., Black, A. C., Sugai, G., Lewandowski, H., Sims, B., & Myers, D. (2012). Illinois statewide positive behavioral interventions and supports: Evolution and impact on student outcome across the years. *Journal of Positive Behavior Interventions*, *14*(1), 5-16. https://doi.org/10.1177/1098300711412601

Skiba, R. J., & Knesting, K. (2001). Zero tolerance, zero evidence: An analysis of school discipline practice. *New Directions for Youth Development*, *92*, 17-43. https://doi.org/10.1002/yd.23320019204

Skiba, R. J., Ritter, S., Simmons, A., Peterson, R., & Miller, C. (2006). The safe and responsive schools project: A school reform model for implementing best practices in violence prevention. In S. R. Jimerson & M. Furlong (Eds.), *Handbook of school violence and school safety: From research to practice.* National Association of School Psychologists.

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. https://doi.org/10.1080/02796015.2011.12087730

Skiba, R. J., Mediratta, K., & Rausch, M. K. (Eds.). (2016). *Inequality in school discipline.* Palgrave MacMillan.

Smolkowski, K., Girvan, E. J., McIntosh, K., Nese, R. N. T., & Horner, R. H. (2016). Vulnerable decision points in school discipline: Comparison of discipline for African American compared to White students in elementary schools. *Behavioral Disorders, 41*, 178–195 https://doi.org/10.17988/bedi-41-04-178-195.1

Southern Coalition for Social Justice. (2017). *The state of discipline in NC Schools.* https://southerncoalition.org/resources/the-state-of-discipline-in-nc-schools/

Southern Poverty Law Center. (2008). *Reducing student and teacher dropout rates in Mississippi*. http://www.splcenter.org/get-informed/publications/reducingstudent-and-teacher-dropout-rates-in-Mississippi

Sprague, J., Colvin, G., & Irvin, L. (1996). *The Oregon School Safety Survey*. University of Oregon.

Sprague, J., Walker, H., Golly, A., White, K., Myers, D., & Shannon, T. (2002). Translating research into effective practice: The effects of a universal staff and student intervention on key indicators of school safety and discipline. *Education and Treatment of Children*, *24*(4), 495-511.

Sugai, G., Horner, R. H., Dunlap, G. Hieneman, M., Lewis, T. J., Nelson, C. M., Scott, T., Liaupsin, C., Sailor, W., Turnbull, A. P., Turnbull, R., Wickham, D., Wilcox, B., & Reuf, M. (2000). Applying positive behavioral support and functional behavioral assessment in schools. *Journal of Positive Behavioral Interventions*, *2*, 131-143. https://doi.org/10.1177/109830070000200302

Sugai, G., Horner, R. H., & Todd, A. W. (2000). *PBIS Self-Assessment Survey*. University of Oregon, OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. https://www.pbis.org/resource/sas

Sugai, G., Sprague, J. R., Horner, R. H., & Walker, H. M. (2000). Preventing school violence: The use of office discipline referrals to assess and monitor school-wide discipline interventions. *Journal of Emotional and Behavioral Disorders, 8*, 94-101. https://doi.org/10.1177/106342660000800205

Sugai, G., Lewis-Palmer, T., Todd, A., & Horner, R. H. (2001). *School-wide Evaluation Tool*. University of Oregon, OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. https://www.pbis.org/resource/school-wide-evaluation-tool-set

Sugai, G., Todd, A. W., & Horner, R. H. (2001). *Team Implementation Checklists (Version 2.2*). University of Oregon, OSEP Technical Assistance Center on Positive Behavioral Supports. https://www.pbis.org/resource-type/assessments

Sugai, G., & Horner, R. H. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Behavior Psychology in the Schools*, *24*, 23–50. https://doi.org/10.1300/J019v24n01_03

Sugai, G., Horner, R. H., & Todd, A. W. (2003). *Effective Behavior Support Self-Assessment Survey (2.0*). University of Oregon. https://ocde.us/EducationalServices/LearningSupports/PBIS/Documents/Tools%20and%20Samples/EBS_Survey.pdf

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

Sugai, G., & Horner, R. H. (2009). Responsiveness to intervention and school-wide positive behavior supports: Integration of multi-tiered system approaches. *Exceptionality*, *17*, 223–237. https://doi.org/10.1080/ 09362830903235375

Sugai, G., & Simonsen, B. (2012). *Positive Behavioral Interventions and Supports: History, defining features, and misconceptions*. OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. https://www.pbis.org/resource/positive-behavioral-interventions-and-supports-history-defining-features-and-misconceptions

Sugai, G., O'Keeffe, B. V., & Fallon, L. M. (2012). A contextual consideration of culture and school-wide positive behavior support. *Journal of Positive Behavior Interventions*, *14*, 197–208. https://doi.org/10.1177/ 1098300711426334

Sugai, G., & Horner, R. H. (2020). Sustaining and scaling positive behavioral interventions and supports: Implementation drivers, outcomes, and considerations. *Exceptional Children, 86*(2), 120-136. https://doi.org/10.1177/0014402919855331

Sullivan, A. L., Van Norman, E. R., & Klingbeil, D. A. (2014). Exclusionary discipline of students with disabilities: Student and school characteristics predicting suspension. *Remedial and Special Education, 35*(4), 199-210. https://doi.org/10.1177/0741932513519825

Sutherland, K. S., Wehby, J. H., & Copeland, S. R. (2000). Effect of varying rates of behavior-specific praise on the on-task behavior of students with EBD. *Journal of Emotional and Behavioral Disorders, 8,* 2–8. https://doi.org/10.1177/106342660000800101

Sutherland, K. S., Alder, N., & Gunter, P. L. (2003). The effect of varying rates of opportunities to respond to academic requests on the classroom behavior of students with EBD.

Journal of Emotional and Behavioral Disorders, 11, 239–248. https://doi.org/10.1177/10634266030110040501

Swain-Bradway, J., Loman, S. L., & Vincent, C. G. (2014). Systematically addressing discipline disproportionality through the application of a school-wide framework. *Multiple Voices for Ethnically Diverse Exceptional Learners*, *14*(1), 2014, 3–17. https://doi.org/10.5555/muvo.14.1.jl626n21408t4846

Swain-Bradway, J., Pinkney, C., & Flannery, K. B. (2015). Implementing schoolwide positive behavior interventions and supports in high schools: Contextual factors and stages of implementation. *TEACHING Exceptional Children*, *47*(5), 245–255. https://doi.org/10.1177/0040059915580030

Tabachnick, B. G. & Fidell, S. (2007). *Using multivariate statistics.* (5th ed.). Pearson.

Taylor, M., West, R., & Smith, T. (2006). *The Indicators of School Quality (ISQ): Survey Manual*. Center for the School of the Future, Utah State University.

Tegano, D. W., Moran, D. J., Ill, & Sawyers, J. K. (1991). *Creativity in early childhood classrooms.* National Education Association.

Thompson, M. T., Marchant, M., Anderson, D., Prater, M. A., & Gibb, G. (2012). Effects of tiered training on general educators' use of specific praise. *Education and Treatment of Children*, *35*(4), 521–546. http://hdl.lib.byu.edu/1877/3995

Tingstrom, D. H., Sterling-Turner, H. E., & Wilczynski, S. M. (2006). The Good Behavior Game: 1969-2002. *Behavior Modification*, *30*(2), 225–253. https://doi.org/10.1177/0145445503261165

Todd, A. W., Kaufman, A., Meyer, G., & Horner, R. H. (2008). The effects of a targeted intervention to reduce problem behaviors: Elementary school implementation of Check In

- Check Out. *Journal of Positive Behavior Interventions, 10*(1), 46-55. https://doi.org/10.1177/1098300707311369

Todd, A., Horner, R. H., & Tobin, T. (2010). *Referral form definitions. School-wide Information System v4.4*. University of Oregon, Educational and Community Supports. https://www.pbisapps.org/Applications/Pages/SWIS-Suite.aspx

Tschannen-Moran, M., Hoy, A., & Hoy, W. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*, 202–248. https://doi.org/10.3102/00346543068002202

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17(7), 783–805. https://doi.org/10.1016/S0742-051X(01)00036-1

Van Houtte, M., & Van Maele, D. (2011). The black box revelation: In search of conceptual clarity regarding climate and culture in school effectiveness research. *Oxford Review of Education*, *37*(4), 505–524. https://doi.org/10.1080/03054985.2011.595552

Vincent, C. G., Cartledge, G., May, S., & Tobin, T. J. (2009). *Do elementary schools that document reductions in overall office discipline referrals document reductions across all student races and ethnicities?* OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. https://www.pbis.org/resource-type/evaluation-briefs

Vincent, C.G., & Tobin, T.J. (2011). An examination of the relationship between implementation of School-Wide Positive Behavior Support (SWPBIS) and exclusion of students from various ethnic backgrounds with and without disabilities. *Journal of Emotional and Behavioral Disorders*, *19*, 217-232. https://doi.org/10.1177/1063426610377329

Vincent, C. G., Swain-Bradway, J., Tobin, T. J., & May, S. (2011). Disciplinary referrals for culturally and linguistically diverse students with and without disabilities: Patterns resulting from school-wide positive behavior support. *Exceptionality, 19*, 175–190. https://doi.org/10.1080/09362835.2011.579936

Vincent, C. G., Randall, C., Cartledge, G., Tobin, T. J., & Swain-Bradway, J. (2011). Toward a conceptual integration of cultural responsiveness and schoolwide positive behavior support. *Journal of Positive Behavior Interventions*, *13*(4), 219-229. https://doi.org/10.1177/1098300711399765

Vincent, C. G., Tobin, T. J., Hawken, L., & Frank, J. (2012). Disciplinary referrals and access to secondary interventions: Patterns across students across African-American, Hispanic-American, and White backgrounds. *Education and Treatment of Children*, *35*, 431-445. https://doi.org/10.1353/etc.2012.0018

Vincent, C. G., Inglish, J., Girvan, E. J., Sprague, J. R., & McCabe, T. M. (2016). School-wide positive and restorative discipline (SWPRD): Integrating school-wide positive behavior interventions and supports and restorative discipline. In R. J. Skiba, K. Mediratta, & M. K. Rausch, (Eds.), *Inequality in school discipline* (pp. 115-134). Palgrave MacMillan.

Wake County Board of Education. (2018). *Code of Student Conduct*. https://boardpolicyonline.com/bl/?b=wake_new#&&hs=208219

Warren, J. S., Edmonson, H. M., Griggs, P., Lassen, S. R., McCart, A., Turnbull, A., & Sailor, W. (2003). Urban applications of school-wide positive behavior support: Critical issues and lessons learned. *Journal of Positive Behavior Interventions*, *5*(2), 80-91. https://doi.org/10.1177/1098300703005002030

What Works Clearinghouse. (n.d.). What Works Clearinghouse. Institute of Education Sciences. http://ies .ed.gov/ncee/wwc/

Wilcox, R. (2017). *Introduction to robust estimation and hypothesis testing*. (4th ed.). Elsevier Inc.

Wolfe, K., Pyle, D., Charlton, C. T., Sabey, C. V., Lund, E. M., & Ross, S. W. (2016). A systematic review of the empirical support for Check-in Check-Out. *Journal of Positive Behavior Interventions*, *18*(2), 74–88. https://doi.org/10.1177/1098300715595957

Wong, C., Odom, S. L., Hume, K., Cox, A. W., Fettig, A., Kucharczyk, S., Brock, M. E., Plavnick, J. B., Fleury, V. P., & Schultz, T. R. (2015). Evidence-based practices for children, youth, and young adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 47*(7), 1951-1966. https://doi.org/10.1007/s10803-014-2351-z.

Woolfolk, A. E., & Hoy, W. K. (1990). Prospective teachers sense of efficacy and beliefs about control. *Journal of Educational Psychology*, 82, 81-91. https://doi.org./10.1037/0022-0663.82.1.81

Yeung, A. S., Craven, R. G., Mooney, M., Tracey, D., Barker, K., Power, A., Dobia, B., Chen, Z., Schofield, J., Whitefield, P., & Lewis, T. J. (2016). Positive behavior interventions: The issue of sustainability of positive effects. *Educational Psychology Review*, *28*, 145-170. https://doi.org/10.1007/s10648-015-9305-7

Table 1

Number of Elementary Schools by Grade Combinations Served

Grades	Schools
Pre-K – K	3
Pre-K – 1	5
Pre-K – 2	19
Pre-K – 3	18
Pre-K – 4	40
Pre-K – 5	731
K – 1	3
K – 2	4
K – 3	2
K – 4	15
K – 5	348
1 – 4	1
1 – 5	3
2 – 3	3
2 – 4	1
2 – 5	2
3 – 5	21
4 – 5	11

Table 2

Average Daily Membership (ADM) of Elementary Schools

ADM	Percentage of schools
36 - 250	7%
251 – 500	39%
501 – 750	41%
751 – 1000	11%
1001 – 1172	2%

Table 3

Concentration of Economically Disadvantaged Students (EDS)

EDS Categories	Percentage of Schools
High	14%
Mid-high	45%
Mid-low	29%
Low	12%

Note: High = ≥ 75%; Mid-high = 50.1% - 75%; Mid-low = 25.1% - 50%; Low = ≤ 25%

Table 4

Percentage of LEAs by Number of Elementary Schools Served

Schools	LEAs
3 – 5	15%
6 – 10	33%
11 – 15	17%
16 – 20	15%
21 – 30	9%
31 – 50	2%
51 – 99	7%
≥ 100	2%

Table 5

GEE Analysis of Factors Associated with Teacher Turnover

Factor	*B*	*SE*	Wald	*p*
Intercept	20.13	4.17	23.3	<.001
PBIS	-0.04	0.65	0.00	.948
Low EDS	-4.29	1.00	18.26	<.001
Mid-low EDS	-3.91	0.88	19.95	<.001
Mid-high EDS	-2.55	0.79	10.16	<.001
Teacher satisfaction	-0.06	0.04	2.01	.156
Long-term suspensions	2.73	1.21	5.08	.024
Short-term suspensions	0.13	0.05	5.91	.015
Crime and violence	-0.38	0.09	19.44	<.001
School size[a]	0.05	0.29	0.03	.862

Intercept = Teacher turnover rate in non-implementing schools with a high concentration of economically disadvantaged students (EDS)

[a] School size: Standardized to account for wide variation in school populations

Following Bonferroni correction procedure, $\alpha = .003$

Table 6

GEE Analysis of Factors Associated with Teacher Satisfaction

Factor	*B*	*SE*	Wald	*p*
Intercept	81.68	1.4	3396.84	<.001
PBIS	0.81	0.58	1.96	.161
Low EDS	9.49	1.58	35.85	<.001
Mid-low EDS	5.06	1.12	20.39	<.001
Mid-high EDS	2.61	1.09	5.72	.016
Teacher turnover	-0.07	0.05	1.88	.170
Long-term suspensions	-4.25	1.22	12.17	<.001
Short-term suspensions	-0.15	0.04	13.49	<.001
Crime and violence	0.02	0.09	0.07	.788
School size[a]	-1.8	0.34	27.62	<.001

Note: Intercept = Teacher satisfaction composite score in non-implementing schools with a high concentration of economically disadvantaged students (EDS)

[a] School size: Standardized to account for wide variation in school populations

Following Bonferroni correction procedure, $\alpha = .003$

Table 7

GEE Analysis of Factors Associated with Annual Teacher Turnover Rates in PBIS-Implementing Schools

Factor	*B*	*SE*	Wald	*p*
Intercept	17.77	5.94	8.95	.003
Model implementation level	0.07	1.34	0.00	.961
Exemplar implementation level	1.50	1.83	0.67	.414
Low EDS	-5.18	2.03	6.52	.011
Mid-low EDS	-3.26	2.24	2.12	.145
Mid-high EDS	-3.45	1.52	5.13	.024
Teacher satisfaction	-0.02	0.07	0.13	.723
Long-term suspensions	2.86	1.48	3.69	.055
Short-term suspensions	0.09	0.08	1.56	.212
Crime and violence	-0.38	1.32	0.08	.773
School size[a]	0.08	0.58	0.02	.888

Note: Intercept = Teacher turnover rate in Green Ribbon implementation level schools with a high concentration of economically disadvantaged students (EDS)

[a] School size: Standardized to account for wide variation in school populations

Following Bonferroni correction procedure, $\alpha = .003$

Table 8

Multiple Linear Regression of Factors Associated with Teacher Satisfaction in PBIS-Implementing Schools

Factor	*B*	*SE*	*t*	*p*
Intercept	76.45	1.81	42.16	<.001
Model implementation level	3.55	0.92	3.87	<.001
Exemplar implementation level	4.78	1.2	3.98	<.001
Low EDS	10.26	2.43	4.23	<.001
Mid-low EDS	7.07	1.55	4.55	<.001
Mid-high EDS	4.72	1.34	3.52	<.001
Teacher turnover	-0.02	0.05	-0.34	.732
Long-term suspensions	-4.12	2.36	-1.74	.083
Short-term suspensions	-0.04	0.04	-0.91	.363
Crime and violence	0.41	0.88	0.47	.640
School size[a]	-1.83	0.53	-3.47	<.001

Note: Intercept = Teacher satisfaction composite score in Green Ribbon implementation level schools with a high concentration of economically disadvantaged students (EDS)

[a] School size: Standardized to account for wide variation in school populations

Following Bonferroni correction procedure, $\alpha = .003$

Table 9

GEE Analyses of Behavioral Outcomes in PBIS Versus Non-Implementing Schools

Factor	*B*	*SE*	Wald	*p*
Non PBIS schools[a]	8.73	1.09	63.78	<.001
Short-term suspensions in PBIS schools	1.53	0.62	6.06	.014
Non PBIS schools[a]	0.01	0.004	9.74	.002
Long-term suspensions in PBIS schools	0.01	0.010	0.57	.452
Non PBIS schools[a]	0.30	0.09	10.67	<.001
Crime and violence in PBIS schools	-0.03	0.10	0.07	.789

[a] Non PBIS Implementing: Represents the intercept (i.e., estimated average rate of short-term suspensions; long-term suspensions; crime and violence rates, respectively, in non-implementing schools,) within each analysis

Following Bonferroni correction procedure, $\alpha = .003$

Table 10

GEE Analysis for Poisson Regression of Short-Term Suspensions across Race

Factor	*B*	*SE*	Wald	*p*
Intercept	-1.94	0.07	855.03	<.001
Hispanic / NonPBIS	-1.68	0.06	915.12	<.001
White / NonPBIS	-1.55	0.06	769.08	<.001
Black / PBIS	0.15	0.10	2.19	.139
Hispanic / PBIS	0.09	0.09	1.10	.295
White / PBIS	0.17	0.09	3.96	.047

Note: Intercept = Short-term suspensions among Black students in non-implementing schools

Following Bonferroni correction procedure, $\alpha = .003$

Table 11

GEE Analysis for Poisson Regression of Crime and Violence Acts across Race

Factor	*B*	*SE*	Wald	*p*
Intercept	-5.61	0.13	2008.1	<.001
Hispanic / NonPBIS	-1.05	0.17	36.80	<.001
White / NonPBIS	-0.944	0.15	37.35	<.001
Black / PBIS	0.06	0.18	0.12	.73
Hispanic / PBIS	0.002	0.30	0.00	.99
White / PBIS	0.272	0.27	1.00	.32

Note: Intercept = Crime and violence acts among Black students in non-implementing schools

Following Bonferroni correction procedure, $\alpha = .003$

Table 12

GEE Analysis for Poisson Regression of Referrals across Disability Status

Factor	*B*	*SE*	Wald	*p*
Intercept	0.08	0.08	0.88	.348
Non-disabled / NonPBIS	-3.02	0.06	2397.23	<.001
With a disability /PBIS	0.16	0.12	1.89	.169
Non-disabled / PBIS	0.26	0.09	8.66	.003

Note: Intercept = Office disciplinary referrals for students with disabilities in non-implementing schools

Following Bonferroni correction procedure, $\alpha = .003$

Table 13

GEE Analysis of Short-Term Suspension Rates across Implementation Levels

Factor	*B*	*SE*	Wald	*p*
Intercept	9.81	1.38	50.79	<.001
Model implementation level	1.66	1.34	1.53	.22
Exemplar implementation level	-0.45	1.23	0.13	.72

Note: Intercept = Short-term suspension rate in Green Ribbon implementation level schools

Following Bonferroni correction procedure, $\alpha = .003$

Table 14

GEE Analysis of Long-Term Suspension Rates across Implementation Levels

Factor	*B*	*SE*	Wald	*p*
Intercept	0.03	0.03	1.19	.27
Model implementation level	-0.02	0.03	0.64	.42
Exemplar implementation level	0.003	0.03	0.01	.93

Note: Intercept = Long-term suspension rate in Green Ribbon implementation level schools

Following Bonferroni correction procedure, $\alpha = .003$

Table 15

GEE Analysis of Crime and Violence Rates across Implementation Levels

Factor	*B*	*SE*	Wald	*p*
Intercept	0.28	0.06	25.30	<.001
Model implementation level	0.003	0.05	0.00	.95
Exemplar implementation level	0.001	0.06	0.00	.98

Note: Intercept = Crime and violence rate in Green Ribbon implementation level schools

Following Bonferroni correction procedure, α = .003

Table 16

GEE Analysis for Poisson Regression of Short-Term Suspensions by Race across Implementation Levels

Factor	*B*	*SE*	Wald	*p*
Intercept	-1.89	0.16	267.86	<.001
Hispanic / Green Ribbon level	-1.51	0.11	191.4	<.001
White / Green Ribbon level	-1.29	0.10	170.14	<.001
Black / Model level	0.21	0.18	1.39	.24
Hispanic / Model level	-0.09	0.15	0.33	.57
White / Model level	-0.15	0.14	1.09	.30
Black / Exemplar level	-0.03	0.19	0.02	.88
Hispanic / Exemplar level	0.008	0.23	0.00	.97
White / Exemplar level	-0.08	0.18	0.21	.65

Note: Intercept = Short-term suspensions for Black students in Green Ribbon implementation level schools

Following Bonferroni correction procedure, $\alpha = .003$

Table 17

GEE Analysis for Poisson Regression of Referrals by Disability Status across Implementation Levels

Factor	*B*	*SE*	Wald	*p*
Intercept	0.21	0.12	2.87	.09
Non-disabled / Green Ribbon level	-2.73	0.08	1075.7	<.001
With a disability / Model level	0.15	0.20	0.55	.46
Non-disabled / Model level	-0.16	0.14	1.35	.24
With a disability / Exemplar level	-0.37	0.23	2.54	.11
Non-disabled / Exemplar level	0.25	0.16	2.29	.13

Note: Intercept = Office disciplinary referrals for students with disabilities in Green Ribbon implementation level schools

Following Bonferroni correction procedure, $\alpha = .003$

APPENDIX

Appendix A: Description of Variables

Average Daily Membership (ADM): The ADM is the total days in membership for all students over the month or school year divided by the number of days school was in session. Average Daily Membership provides a more accurate count of the number of students in a school than enrollment (North Carolina Department of Public Instruction, n.d.). In the mixed effects linear regression models, average daily membership (i.e., school population) was a standardized variable.

PBIS Implementing or Non-implementing Status: In the DPI's original data set containing schools that sought recognition for PBIS implementation, schools were coded by DPI as "yes" or "no" within the "implementing" column, and coded as "approved" or "denied" in the application "status" column. In the principal investigator's database, each school identified as "yes" and "approved" was coded as "1" (implementing), and all other schools were coded as "0" (non-implementing).

PBIS Implementation Level: In the DPI's original data set containing schools that sought recognition for PBIS implementation, schools were identified as Green Ribbon, Model, or Exemplar within the "recognition requested" column, coded as "yes" or "no" within the "implementing" column, and coded as "approved" or "denied" in the application "status" column. In the principal investigator's database, for each school identified as "yes" and "approved," the implementation level was coded as "1" (Green Ribbon), "2" (Model), or "3" (Exemplar). Non-implementing schools were coded as "NA" in the implementation level column.

Concentration of Economically Disadvantaged Students by Category: In the original data set from the DPI, the concentration of economically disadvantaged students per school was derived from the percentage of students eligible to receive free or reduced-price lunch. These data were presented as a range of percentages for each school (e.g., "between 25% to 30%"). The principal investigator converted these data into the four concentration categories offered by the U.S. Department of Education (2015): low (≤ 25%), mid-low (25.1% - 50%), mid-high (50.1% - 75%) or high (≥ 75%)**.** Concentration of economically disadvantaged students was included as a fixed effect in the mixed effects linear regression models.

Overall Composite Mean Rate of Agreement on Teacher Working Conditions Survey:

The North Carolina Teacher Working Conditions Survey is conducted bi-annually and contains 74 items organized into the following eight construct categories: Time, facilities and resources, professional development, school leadership, teacher leadership, instructional practices and support, managing student conduct, and community support and involvement. An overall teaching conditions composite is derived from responses to two additional items. Participants respond to each item on a 4-point scale from "strongly disagree" to "strongly agree." Participants could also select "don't know," and these responses were treated as missing values by the DPI. In the original data set provided by the DPI for the 2015-16 school year, the average rate of agreement (percent of teachers who "agree" plus percent who "strongly agree") for each construct (per school) was presented. The overall composite mean rate of agreement was obtained by averaging these construct values for each school. These data were presented as percentages. The average composite mean rate of agreement for the schools included in this study was 82.6%. Composite mean rate of agreement values across schools ranged from 46.5% to 98.1%.

Annual Teacher Turnover Rates (per school): The percent of teachers within each school in March 2015 who were no longer employed as teachers in March 2016. The original data from the DPI also contains average annual teacher turnover rates per LEA and statewide for elementary schools. The average teacher turnover rate for the schools included in this study was 13.4%. Turnover rates across schools ranged from 0% to 54.1%.

Aggregated Acts of Crime and Violence: Presented in decimal form as the rate per 100 students. Sixteen crime and violence types were included in the totals: (1) homicide, (2) assault resulting in serious bodily injury, (3) rape, (4) sexual offense, (5) sexual assault, (6) kidnapping, (7) robbery with a dangerous weapon, (8) taking indecent liberties with a minor, (9) bomb threat, (10) burning of a school building, (11) possession of alcoholic beverage, (12) possession of controlled substance in violation of law, (13) possession of a firearm or powerful explosive, (14) possession of a weapon, (15) assault involving the use of a weapon, and (16) assault on school personnel (NCDPI, 2017c). The average crime and violence rate for the schools included in this study was 0.29 (per 100 students). Crime and violence rates across schools ranged from 0 to 41.67 (per 100 students).

Short-term Suspensions per 100 Students: Presented in decimal form as the rate per 100 students. The original data from the DPI also contains average short-term suspension rates per LEA and statewide for elementary schools. The average short-term suspension rate for the schools included in this study was 7.95 (per 100 students). Short-term suspension rates across schools ranged from 0 to 123.51 (per 100 students).

Long-term Suspensions per 100 Students: Presented in decimal form as the rate per 100 students. The original data from the DPI also contains average long-term suspension rates per LEA and statewide for elementary schools. The average long-term suspension rate for the

schools included in this study was 0.01 (per 100 students). Long-term suspension rates across schools ranged from 0 to 2.79 (per 100 students).

Number of Office Disciplinary Referrals Issued to Students with Disabilities: Presented as whole number counts. In the original data set from the DPI, the "number of incidents" were disaggregated across disability categories for each school. Prior to analyses, the principal investigator totaled those counts to yield the number of referrals among students with disabilities within each school.

Number of Office Disciplinary Referrals Issued to Non-disabled Students: Presented as whole number counts for each school.

Number of Short-term Suspensions Disaggregated by Ethno-racial Category: Presented as whole number counts for Asian, Black, Hispanic, Indigenous (classified by DPI as American Indian), Multiracial, Pacific Islander, Other or Missing, and White students within each school. The term "Indigenous" refers to students who are Indigenous to North America, and is utilized in light of calls to adjust terminology (e.g., American Psychological Association, 2020).

Number of Long-term Suspensions Disaggregated by Ethno-racial Category: Presented as whole number counts for Asian, Black, Hispanic, Indigenous, Multiracial, Pacific Islander, Other or Missing, and White students within each school.

Number of Crime and Violence Acts Disaggregated by Ethno-racial Category: Presented as whole number counts for Asian, Black, Hispanic, Indigenous, Multiracial, Pacific Islander, Other or Missing, and White students within each school.