

STETSON UNIVERSITY

*Voices of Reform: Educational Research to
Inform and Reform*

Volume 3 • Issue 2 • Article 5



December 2020

PBIS Implementation Fidelity and Student Outcomes in an Urban School District

C. Angelique Scherer

Jefferson County Public Schools

W. Kyle Ingle

University of Louisville

Follow this and additional works at: <http://www.voicesofreform.com>

Recommended Citation

Scherer, C. A. & Ingle, W. K. (2020). PBIS implementation fidelity and student outcomes in an urban school district. Retrieved from <https://www.voicesofreform.com/article/18597-pbis-implementation-fidelity-and-student-outcomes-in-an-urban-school-district> doi: 10.32623/3.10007

<http://dx.doi.org/10.32623/3.10007>

Revisions

Submission date: April 17th, 2020

1st Revision: May 20th, 2020

Acceptance: May 20th, 2020

Publication date: December 29th, 2020

PBIS Implementation Fidelity and Student Outcomes in an Urban School District

C. Angelique Scherer¹
W. Kyle Ingle²

¹Assistant Director of Exceptional Child Education
Jefferson County Public Schools, Louisville, Kentucky, United States
angelique.scherer@jefferson.kyschools.us

²Department of Educational Leadership, Evaluation, and Organizational Development
University of Louisville, United States
william.ingle@louisville.edu

Abstract

The purpose of our study was to examine Positive Behavior Interventions and Supports (PBIS) implementation fidelity and student outcomes for secondary schools in a large, urban school district placed under a corrective action plan due to disproportionate suspension practices for students of color. Drawing upon data over a four-year period on PBIS implementation fidelity, discipline referrals, suspensions, and standardized test scores, we employed repeated measures analysis of variance (ANOVA) and revealed improved outcomes in the first year, but improvements did not continue over time. Additionally, academic outcomes did not show any significant increases. The study concludes with implications for policy, practice, and future research.

Keywords

urban schools, district reform, behavioral interventions, discipline

Introduction

Maintaining safe school environments is a high priority for teachers and educational leaders (Skiba & Sprague, 2008). When students display inappropriate behaviors, school administrators may use a variety of consequences, including detention, in-school suspension, out-of-school suspension, and expulsion. Typically reserved for the most serious breaches of school conduct, out-of-school suspensions exclude students from school attendance and school functions for 1 to 10 days at a

time (Gregory, Skiba, & Noguera, 2010). However, disproportionate suspension of students of color is a nationwide trend in US public schools. Researchers reveal suspension rates that are higher among Black students than their White peers (Raffaele, Mendez, & Knoff, 2003; Skiba et al., 2011; Skiba et al., 2002). Data from district, state, and national levels indicate students of color, encompassing all non-white groups, are overrepresented in office referrals and school expulsions, with suspension rates of two to three times that of other students (Skiba et al., 2002). Students who are suspended miss instructional time and are often further behind peers when they return to school, creating a cycle of lower academic achievement and disengagement (Gregory et al., 2010; US Departments of Education & Justice, 2014).

Positive Behavioral Interventions and Supports (PBIS) began as a strategy for misbehavior for students with disabilities. In the 1997 reauthorization of the Individuals with Disability Education Act (IDEA), Congress added PBIS as an approach to prevent exclusions from school and improve educational results for students with disabilities. PBIS remains in the 2004 amendments of IDEA as a strategy for students who struggle with behavior. Since then, PBIS emerged as a school-wide initiative to improve outcomes for all students. School-wide Positive Behavior Support (SWPBS) is a framework of intervention practices and organizational systems for establishing the social culture, learning and teaching environment, and individual behavior supports needed to achieve academic and social success for all students (Office of Special Education Programs, 2015). Dissemination in schools, districts, and state education agencies resulted largely due to the Technical Assistance (TA) Center on PBIS, funded by the US Department of Education's Office of Special Education Programs (OSEP) since 1998. PBIS gained momentum nationwide as a potential strategy for preventing negative behaviors and improving discipline outcomes. There is evidence that PBIS reduced suspensions and increased student outcomes at the elementary level. However, there are limited data for PBIS implementation for secondary schools in urban settings (Bohanon et al., 2006; Flannery et al., 2014; Sugai & Horner, 2008).

In response, we drew upon data collected over a four-year period from 24 secondary schools in a single urban school district in the southeastern United States. We sought to answer the following research questions:

1. Are schools under corrective action implementing PBIS Tier I with fidelity over time?
2. Is there a statistically significant difference in office discipline referrals (ODR) after PBIS implementation?
3. Is there a statistically significant difference in out-of-school suspensions (OSS) after PBIS implementation?
4. Is there a statistically significant difference in reading achievement after PBIS implementation?
5. Is there a statistically significant difference in mathematics achievement after PBIS implementation?

Review of the Literature

PBIS is not a curriculum or program, but an approach to prevent problem behaviors and minimize the use of punitive responses to inappropriate behaviors. PBIS is a school-wide approach to positive behaviors, but it is dependent on the extent that individual teachers implement classroom PBIS practices with fidelity (Mathews, McIntosh, Frank, & May, 2014). School leaders use the multi-tiered approach to change school culture while providing behavioral supports that will assist in producing safe, orderly, and effective learning organizations (Sugai & Horner, 2009). PBIS is a problem solving and action planning process through which school leadership teams (a) review school data, (b) develop measurable and realistic short and long-term objectives, (c) select practices that have demonstrated efficacy in achieving those objectives, and (d) establish systems to enable adaptation and implementation of those practices (Sugai et al., 2000).

The PBIS framework focuses on a continuum of supports in which students receive assistance based on the intensity of their needs. The framework is comprised of three tiers, which identifies students' needs and level of support received. Data analysis is a critical feature of the PBIS framework because school teams use discipline data to determine problem areas and the level of support needed for students. The first tier is a universal tier of support for all students in a school. The school staff establish expectations and teach the common expectations to all students. Additionally, school staff reinforce positive behaviors and avoid the over-reliance on negative responses to problem behaviors. Identifying and teaching clear expectations for behavior reduces ambiguity for both students and staff as well as working toward the shared goal of a positive school culture (McIntosh et al., 2014). The second tier is for students who do not respond to the first tier of support and require targeted interventions for their problem behaviors. Targeted interventions may be an adult mentor to check in on students or small group social skills lessons. Finally, the third tier is for students with significant problem behaviors requiring intensive support due to the frequency or intensity of their behaviors. Students at this level require individual support to meet their needs. A school staff person completes the functional behavioral assessment (FBA) at this tier to identify the target behavior and the function of the behavior. The multi-tiered approach allows for a customization of support based on student needs.

Research suggests that PBIS is a promising approach to meet the challenge of school suspensions because of its focus on preventive measures for problem behavior. PBIS also has been associated with decreases in office discipline referrals, suspension, and expulsions (Barrett, Bradshaw, & Lewis-Palmer, 2008; Horner, Sugai, & Anderson, 2010). Schools implementing PBIS with higher fidelity report benefits, such as decreases in problem behavior, increases in academic time, and improved perceptions of school safety (Bradshaw, Mitchell, & Leaf, 2010; Swain-Bradway et al., 2013). Scholars find that implementation of PBIS with greater fidelity is associated with positive outcomes for students and schools. In a quasi-experimental design, Freeman et al. (2016) compared school outcomes prior to PBIS implementation to school outcomes after implementation. They reported that the effect of fidelity on ODR rates indicated statistically significant decreases for schools approaching fidelity. Schools that were approaching fidelity or at implementation fidelity had significantly lower ODR rates than schools that were not

implementing. Although they did not find statistically significant differences in academics, they did find significant differences in attendance.

There is a variety of tools available to measure PBIS implementation. Examples include, but are not limited to the Benchmarks of Quality (Childs, Kincaid, & George, 2010), Tiered Fidelity Inventory (Algozzine et al., 2014) and the School-wide Evaluation Tool (Sugai, Lewis-Palmer, Todd, & Horner, 2001). These tools are typically self-assessments completed by school staff to determine if the components of PBIS are in place, such as having a team dedicated to implementing PBIS, data analysis, behavioral expectations taught, and staff buy-in. Schools may assess their PBIS implementation multiple times in a school year depending on the assessment tool and their stage of PBIS implementation. Using the self-assessment tools provides schools with the information to celebrate successes and make improvements in PBIS implementation when needed.

Although PBIS has supporting evidence in elementary schools, there is a lack of data on the outcomes for schools on the secondary level (Sugai & Horner, 2008). Secondary schools may struggle to implement PBIS with fidelity due to factors such as school size, administrative support, or staff buy-in. Middle and high schools tend to have large enrollments with students coming from multiple elementary schools. Students have multiple teachers with varying expectations. Secondary teachers assume the students are able to self-manage and be responsible for their own learning. Teachers may not understand the need to teach the adolescents and teenagers expected behaviors. The lack of school-wide systems, consistent teacher expectations for student behaviors, adequate implementation logistics, and consistent policies are barriers to PBIS implementation in urban schools (Bohanon et al., 2006).

Methodology

To answer our research questions, we used the repeated measures analysis of variance (ANOVA) statistical procedure to determine if there were statistically significant differences in PBIS implementation fidelity and student outcomes. Stevens (2007) indicates the repeated measures ANOVA is appropriate when one is concerned with performance trends over time. In terms of assumptions, the independence of observation assumption may be violated, as the schools participated in training together and had opportunities to share ideas on how to implement PBIS. To account for the potential of this error, a more conservative alpha level was set at .01. We used the Mauchly's test to determine if we met the assumption of sphericity. If the Mauchly's test indicated that we did not meet the assumption of sphericity, we then used the Greenhouse-Geisser correction to account for this type of error. The Greenhouse-Geisser correction is more conservative and decreases the likelihood of error (Stevens, 2007).

Context of the Study

We drew our data from a large, urban school district in the southeastern United States. The district is the largest in the state with approximately 100,000 students, 47 secondary schools (23 middle- and 24 high schools). As part of the development of the corrective action plan (CAP), the district identified 24 middle and high schools as *high suspension schools* with disproportionate suspension rates of African-American students, mandating PBIS training and

implementation in order to reduce high suspension rates. We collected data from these 24 schools. The 2013-14 school year served as the baseline year for the student outcome variables. We analyzed PBIS implementation fidelity for the three academic years immediately following the initial training (2014-15, 2015-16, and 2016-17).

Data Sources

We drew upon multiple sources of data for our study, including a survey implemented to educators in the sampled schools and administrative datasets collected by the district and state's department of education. To assess PBIS fidelity, we used the Benchmarks of Quality (BOQ) and the Tiered Fidelity Inventory (TFI). The district used the BOQ in the initial year of PBIS training to measure implementation. In subsequent years, the district used the TFI as the fidelity instrument to measure PBIS implementation. Both the BOQ and TFI are research-validated measures used to assess PBIS implementation (Cohen et al., 2007; McIntosh et al., 2017). Research suggests that both measures are valid assessments (Cohen et al., 2007; McIntosh et al., 2017). The BOQ has demonstrated strong psychometric properties with internal consistency (.96), test-retest reliability (.94) and interrater reliability (.87) (Cohen et al., 2007).

The TFI is a measure of the extent to which school staff is applying the core features of PBIS (Algozzine et al., 2014). Research indicates survey scores are more accurate when a coach external to the school team facilitates completion of the inventory (Algozzine et al., 2014). The inventory yields summary scores for each scale, representing the percentage of PBIS components in place for each tier and a total for all three tiers. The school team and their coach use the score reports to develop action plans moving forward. A score of 70% is acceptable for implementation (Algozzine et al., 2014). The school teams administered the TFI in both fall and spring of the school year. For data consistency, we used the spring score, which is the final score for the end of that particular school year. Using the spring score also allowed schools to improve upon fall scores from earlier in the school year.

McIntosh et al. (2017) conducted studies to assess the TFI construct validity, usability, reliability, and concurrent validity with existing, validated measures of SWPBIS fidelity of implementation. They determined the TFI is reliable and valid to measure PBIS implementation fidelity as evidenced by a high test-retest reliability (.99) and interrater reliability across all three tiers (.99). They also determined the TFI correlates with other PBIS fidelity implementation measures such as the Benchmarks of Quality (BOQ). Additionally, Mercer, McIntosh, and Hoselton (2017) found that there were no statistically significant differences between the Tier I scores of the TFI and BOQ in their study of PBIS fidelity measures. They found few differences in mean scores of the fidelity assessments with the exception of the School-wide Evaluation Tool (SET). Additionally, they indicated that the TFI demonstrated evidence of convergent validity with other Tier I PBIS fidelity measures ($r = .54-.64$).

We used office discipline referral (ODR) data and out-of-school suspension (OSS) data as indicators of problem behaviors (Irvin et al., 2004). Both the school district and state department of education collect these data as indicators of school quality and school safety. While school and district leaders often use ODR to measure problem behaviors, their use in schools is subjective.

The referral process is subject to influence by students' previous problem behaviors or race/ethnicity (Skiba et al., 2008). Additionally, other teacher factors may influence the use of ODRs, including classroom management skills, tolerance for disorder, or even mood for a particular day. Most ODRs originate in the classroom setting. Pas, Bradshaw, and Mitchell (2011) examined data from 21 elementary schools implementing PBIS, finding that ODRs correlated significantly and moderately ($r = .57$) with teacher ratings of student behavior problems.

Out-of-school suspensions (OSS) may result from student's inappropriate behavior, requiring student exclusion completely from the school environment for a set period of time (Costenbader, & Markson, 1998). The student may not participate in school or any extracurricular activities. School administrators have long used suspensions as a method to improve student behavior; however, there is not conclusive evidence to document the effectiveness of suspension to change student behavior. There is limited evidence to document the reliability and validity of suspensions. In spite of the limited data reliability of the use of suspensions, school districts, states and even the federal government use this data as an indicator of school quality and safety.

In the district that served as the context of our study, a school leader may issue OSS in response to student offenses as outlined in the district's Code of Acceptable Behavior and Discipline, which classifies violations and suggests disciplinary actions according to severity. Behaviors that may warrant suspension include fighting, possession of weapons, and possession of tobacco or drugs. There are also some behaviors, which are not as severe but may also result in a suspension, such as classroom disruption or failure to follow instructions. We calculated OSS events by dividing the total number of suspensions per school by the total enrollment per school year to determine if there was a statistically significant difference in OSS for each school after three years of PBIS implementation.

In terms of student achievement data, the state department of education uses a criterion-referenced test consisting of multiple choice, extended response, and short answer questions for students in grades 3-8. Students receive a scaled score derived from a statistical transformation of the raw scores. The scaled scores represent a metric that is consistent across test forms and allows for comparisons across test administrations within subject and grade. The scores are categorized using cut points into four performance levels: Novice, Apprentice, Proficient, and Distinguished. Students earning Novice or Apprentice are below state benchmarks, while Proficient meets state benchmarks and Distinguished exceeds benchmarks.

State department officials also adopted a new assessment model at the high school level to measure reading and math proficiency. State department officials purchased assessments for English II, Algebra II, Biology and US History as part of ACT's Quality Core program. The state department of education uses these end-of-course assessments for Algebra II and English II in the state accountability model to demonstrate proficiency in reading and math at the high school level. These end-of-course tests (EOC) measure grade appropriate core academic content, basic skills, and higher-order thinking skills. Similar to the tests administered to grades 3-8, the EOC consist of multiple choice, multiple select, short answer, extended response, and essay items. We examined the percentage of students in each school earning a rating of Proficient or Distinguished,

which mirrors the performance ratings for the middle school level. We summarize the variables and their grounding in the research literature in Table 1.

Table 1: *Variables in the Model with Their Grounding in the Literature*

	Variable	Type	Grounding in the Literature
Years of PBIS Implementation	Number of years of PBIS implementation	Independent Categorical	Childs et al. (2016); Simonsen et al. (2012);
PBIS Implementation Fidelity	Benchmarks of Quality	Dependent Continuous	Childs, Kincaid & George (2010); Cohen, Kincaid, & Childs (2007)
	Tiered Fidelity Inventory	Dependent Continuous	Algozzine et al. (2014); McIntosh et al. (2017)
Discipline Outcomes	Office Discipline Referral	Dependent Continuous	Irvin, Tobin, Sprague, Sugai & Vincent (2004); Skiba et al. (2002); Skiba et al. (2011)
	Out-of-School Suspension	Dependent Continuous	Raffaele Mendez (2003); Raffele Mendez & Knoff (2003); Skiba & Rausch (2006); Skiba et al. (2011); Wald & Losen (2003)
Academic Outcomes	State Reading scores	Dependent Continuous	Arcia (2006); Morris & Perry (2016); Perry & Morris (2014)
	State Math scores	Dependent Continuous	Arcia (2006); Morris & Perry (2016); Perry & Morris (2014)

Results

PBIS Implementation Fidelity

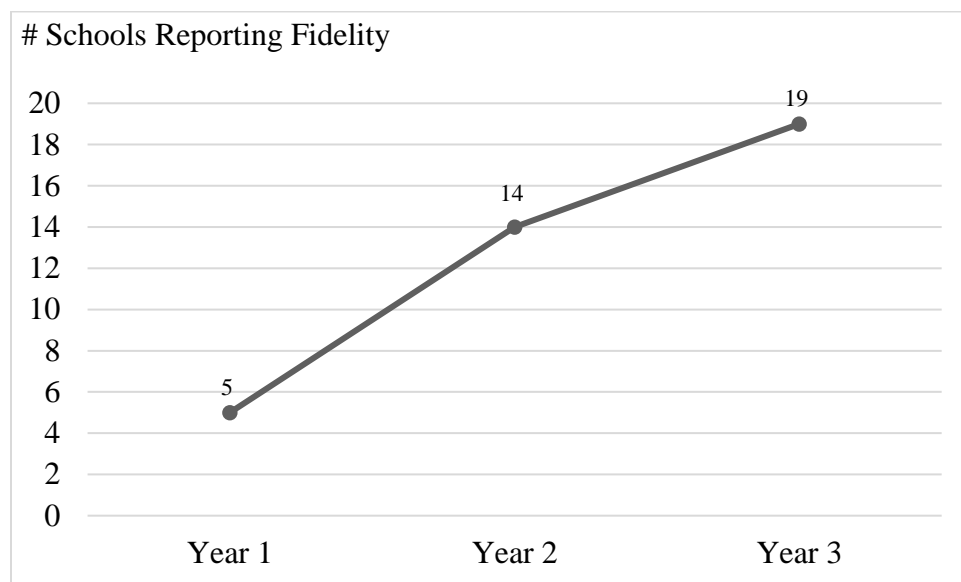
School district personnel measured PBIS implementation fidelity with the BOQ in Year 1 and the TFI in Years 2 and 3. The criteria for implementation fidelity is 70% for both the BOQ and the TFI. All schools included in the study reported PBIS implementation fidelity scores for each year of the study. We summarize the descriptive statistics for PBIS implementation fidelity for each year in Table 2. An examination of the mean scores by implementation year indicates that the total mean scores increased each year. The mean score for Year 1 indicates the district as a whole was not implementing PBIS with fidelity, but improved and reached fidelity with scores 70% or greater in Years 2 and 3. The standard deviation from Year 1 to Year 2 decreased slightly indicating less variation in mean scores from year 1 to year 2. However, there is greater variation from year 2 to year 3. Due to the variability in scores, we calculated the range of the mean scores. A closer examination of the scores indicated that in the first year, five schools reported PBIS implementation scores at fidelity. By the third year, 19 schools reported PBIS implementation scores that met fidelity.

Not all schools reported implementation data for each year in the study. In the first year, one school did not report implementation data. Another school did not report implementation data in the third year. This school closed at the end of the 2015-16 school year; thus, scores were absent in the 2016-17 data. We summarize the schools' progress toward implementation fidelity in Figure 1.

Table 2: *Descriptive Statistics for PBIS Implementation Fidelity*

Year	N	M	SD	Range
Year 1	24	50.54	20.68	7-51
Year 2	24	72.33	17.76	40-100
Year 3	24	77.42	25.93	30-100

Figure 1: *Number of Schools Reporting Fidelity*



We conducted the repeated measures ANOVA comparing the implementation fidelity scores at the end of each of the 3 years. The Mauchly's test revealed a violation of the sphericity assumption ($\chi^2(2) = 15.22, p = .01$). In response, we used the Greenhouse-Geisser correction and a more conservative alpha level at $p < .01$. We found a statistically significant difference in PBIS implementation fidelity for the three years ($F(1.33, 30.68) = 15.42, p < .01$). The effect size ($\eta^2 = .40$) indicated a large difference in overall scores. The differences for PBIS implementation were both statistically significant and practically significant. We used the Bonferroni test to determine in which years there were statistically significant differences. The pairwise comparisons indicated statistically significant increases from Year 1 to both Years 2 and 3. The mean difference from Year 1 to Year 2 increased 21.79. The mean difference from Year 1 to Year 3 increased 26.88. There was not a statistically significant difference between Year 2 and 3. The results suggest a statistically significant difference in PBIS implementation fidelity over time.

Office Discipline Referrals

We analyzed ODR rates to determine if there was a statistically significant difference over the four years included in the study. The first year of ODR data was during the year of PBIS training and prior to full PBIS implementation. The subsequent three years are after PBIS training and implementation. An examination of the mean scores by implementation year indicates that the total mean scores of ODR decreased during the first year after PBIS training. However, after the initial year of PBIS implementation, the mean rates of ODR increased each year (See Table 3).

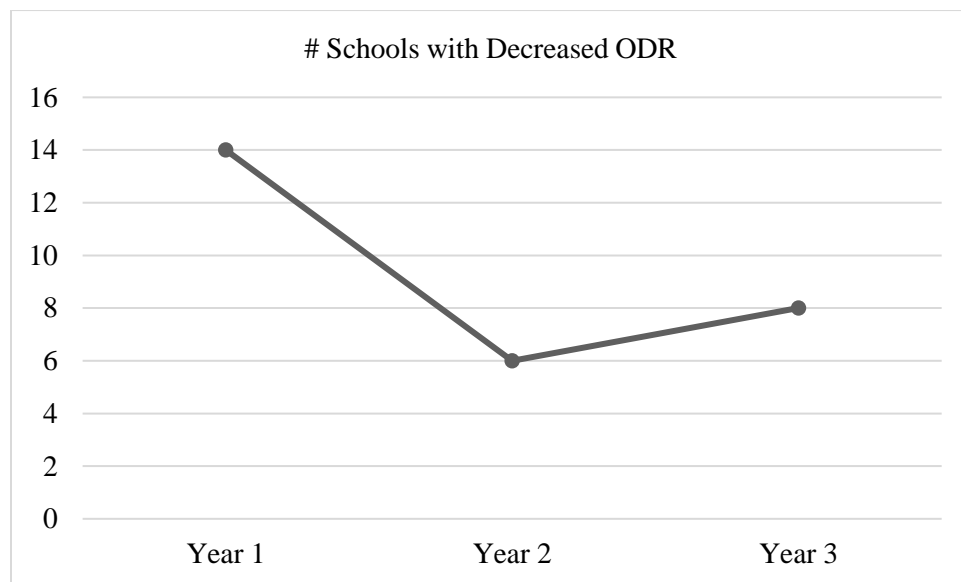
Table 3: *Descriptive Statistics for Office Discipline Referrals*

Years	N	M	SD
Year 0 (Baseline)	24	2.64	2.26
Year 1	24	2.46	1.67
Year 2	24	2.90	1.77
Year 3	24	3.92	2.92

Due to the increase in the mean rates for ODR, we examined the data to determine how many schools experienced a decrease in ODR from the previous school year (See Figure 2). In the first year after PBIS training, 14 schools experienced a decrease in the number of referrals from the previous school year. After the second year of training, six schools experienced a decrease in referrals. After the third year, eight schools decreased referrals from the previous year. One school did not have ODR data for 2016-17, as the school closed in 2015-16. One middle school did not have data for the 2013-14, 2014-15 and 2015-16 school years. District officials included the data for this particular middle school with the high school as the school has grades 6-12. District officials did not disaggregate the discipline data until the 2016-17 school year.

We conducted the repeated measures ANOVA to determine if there was a statistically significant difference in ODR rates over the four years. Again, the Mauchly's test revealed that a violation of the sphericity assumption ($\chi^2(5) = 42.05, p = .01$), thus requiring the Greenhouse-Geisser correction and more conservative alpha level at $p < .01$. There was not a statistically significant difference in ODR rates for the four years ($F(1.67, 38.43) = 4.01, p < .01$).

Figure 2: *Number of Schools with Decreased ODR*



Out-of-School Suspension

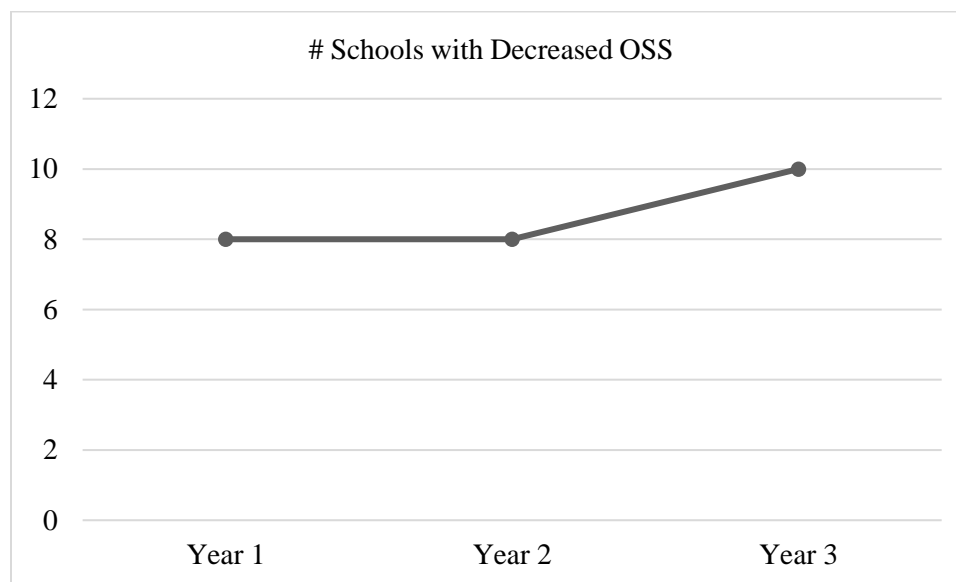
We calculated OSS rates by dividing the total number of suspension events per school by the total enrollment for each school to get a rate for each school. We used descriptive statistics to analyze the means and standard deviations of OSS for all four years (See Table 4). An examination of the mean scores by implementation year indicates that the total mean rates for OSS increased each year. Each year the mean rate of OSS increased from the previous year. The mean OSS rates were highest in Year 3 after PBIS training. The standard deviation increased after the baseline year, suggesting more variation in the mean OSS rates.

Due to variation in mean OSS rates, we examined the number of schools that decreased suspensions from the previous school year (See Figure 3). The first two years after PBIS training, eight schools experienced decreased suspensions from the previous school year. In the third year of PBIS implementation, 10 schools experienced decreased suspensions from the previous school year.

Table 4: *Descriptive Statistics for Out-of-School Suspension*

Years	N	M	SD
Year 0 (Baseline)	24	.29	.19
Year 1	24	.43	.25
Year 2	24	.50	.27
Year 3	24	.52	.26

Figure 3: *Number of Schools with Decreased OSS*



We conducted the repeated measures ANOVA to examine OSS rates both before and after PBIS training and implementation using the Greenhouse-Geisser correction and more conservative alpha level at $p < .01$. There was a statistically significant difference in OSS events for the four years ($F(2.11, 48.59) = 7.60, p < .01$), with OSS rates increasing each year. Additionally, there was practical significance based on the effect size measure ($\eta^2 = .25$). This large effect size indicates large overall differences in scores over the years. The post hoc analyses with Bonferroni correction indicated the significant difference occurred between the baseline year and years 2 and 3. The mean difference from the baseline year increased .22 in Year 2. Additionally, the mean difference from the baseline year increased .24 in Year 3. There was a statistically significant increase in OSS rates from the baseline year to the second and third years of PBIS implementation.

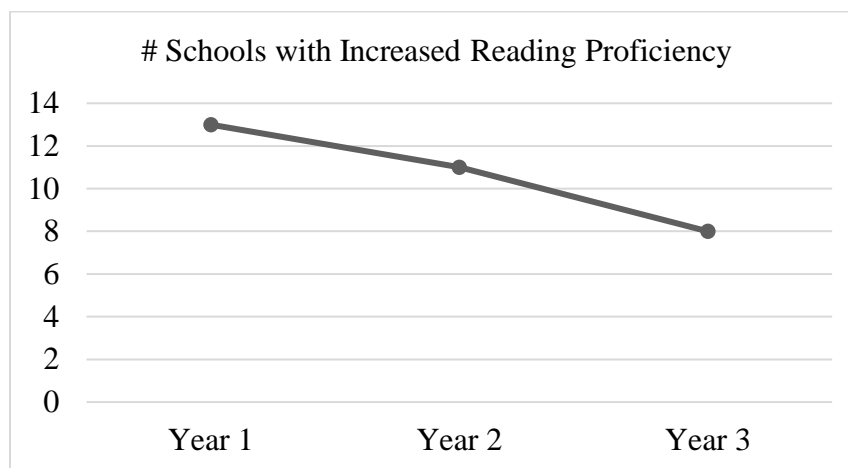
Reading Achievement

The mean reading scores remained relatively consistent over the four years (See Table 5). The mean scores decreased slightly from the baseline year to Year 1 of implementation. There was an increase from Year 1 to Year 2 and then another decrease from Year 2 to Year 3. The standard deviations from year to year do not vary greatly, suggesting that the mean scores were not widely spread over the four years. Due to the flat mean scores, we examined individual school scores to determine if schools increased reading proficiency over the 4 years. We summarize the number of schools with increases in reading proficiency in Figure 4, finding that 13 schools increased the percentage of students meeting reading proficiency by at least 1% from the Baseline year to Year 1 of PBIS implementation. In the following year, eleven schools increased their reading proficiency by at least 1% in Year 2 of PBIS implementation. In the final year, eight schools increased from the previous year.

Table 5: *Descriptive Statistics for Reading Achievement*

Years	N	M	SD
Year 0 (Baseline)	24	35.52	14.48
Year 1	24	34.90	15.04
Year 2	24	36.06	15.34
Year 3	24	35.03	16.68

Figure 4: *Number of Schools with Increased Reading Proficiency*



We conducted the repeated measures ANOVA to examine reading achievement for each year of implementation, seeking to determine if there was a statistically significant difference in reading achievement PBIS implementation. Using the Greenhouse-Geisser correction and more conservative alpha level at $p < .01$, our analysis did not reveal a statistically significant difference in the means for reading achievement during the years of the study.

Mathematics Achievement

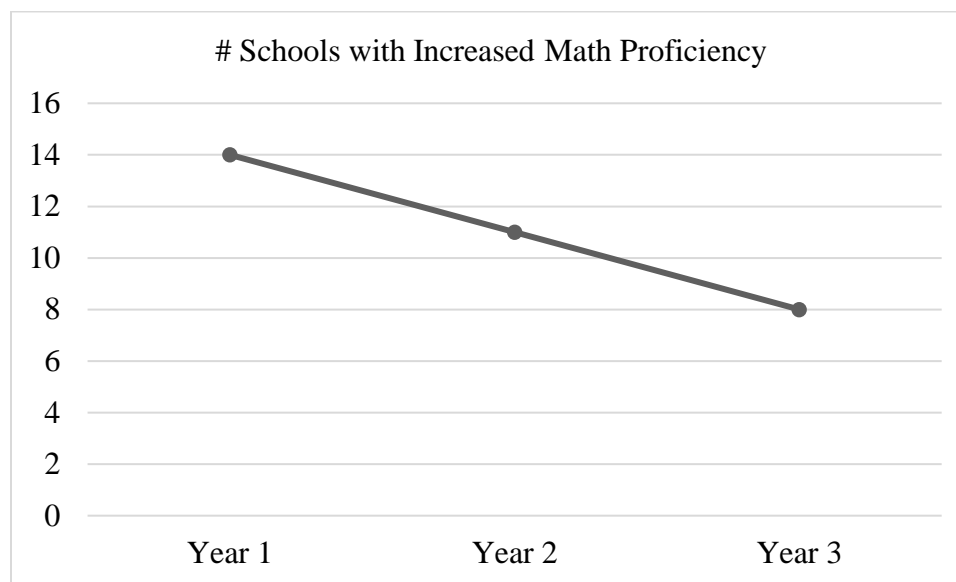
The final outcome we analyzed was mathematics achievement as measured by scores from the state assessment. The mean scores indicate a decrease from the baseline year to Year 1. There was a slight increase from Year 1 to Year 2, but then a decline in Year 3. The standard deviation for Year 3 is also greater than the previous years, suggesting more variation in mean scores than in previous years (See Table 6).

Due to variation in mean scores, we examined the individual school scores (See Figure 5). In total, 14 schools increased the percentage of students meeting mathematics proficiency from the baseline year to Year 1 after PBIS implementation. In Year 2, 11 schools increased proficiency from the previous year and eight increased proficiency scores in Year 3. Each year of the study the number of schools with increased mathematics proficiency decreased.

Repeated measures ANOVA (using the Greenhouse-Geisser correction and more conservative alpha level at $p < .01$) revealed no statistically significant difference between the means for the different years of PBIS implementation ($p > .01$).

Table 6: *Descriptive Statistics for Mathematics Achievement*

Years	N	M	SD
Year 0 (Baseline)	24	30.15	14.12
Year 1	24	26.99	13.80
Year 2	24	29.80	14.93
Year 3	23	25.45	15.23

Figure 5: *Number of Schools with Increased Math Proficiency*

Implications for Policy and Practice

The extent to which schools implement PBIS with fidelity is important for school staff, administrators, district leaders, and policymakers. School leaders should view PBIS as a system to incorporate and enhance all behavioral efforts within a school and not as *just another initiative*. One strategy for improving the likelihood that a school will sustain implementation is to incorporate PBIS into the school improvement plan to address school climate (Warren et al., 2003). The school improvement plan could identify professional development activities, identify a team to support implementation, and serve as an example of how data may inform decision-making. Assessing fidelity and using data to inform action planning may increase sustainability and decrease the likelihood of abandoning effective practices (McIntosh et al., 2015). School PBIS teams must use their fidelity data for continuous improvement in PBIS implementation.

While our results did not indicate significant differences in academic achievement, some schools observed increases in achievement scores. As schools implement PBIS, the school teams may also analyze classroom procedures to determine if classroom management strategies have increased opportunities for students to learn. School administrators may also observe teacher practices to ensure teachers are using evidence-based practices to meet student needs.

Implications for Future Research

In terms of limitations, our study was limited to one school district implementing PBIS. As such, our results are not generalizable beyond the one district and the years reflected in our data set. Additionally, schools may have improved discipline outcomes for reasons unrelated to PBIS implementation fidelity (e.g., administrator turnover). All of the schools in the study were under scrutiny due to high suspension rates that affected students of color disproportionately. School administrators may have employed other interventions in addition to PBIS in their efforts to improve the discipline outcomes. Administrators may have increased the use of in-school

suspension as an alternative to out-of-school suspension. Future research studies may want to explore the extent to which schools utilized other interventions in addition to PBIS. Future research studies may also include administrative (principal and/or assistant principal) personnel changes as variables in their models. High staff turnover—teachers and administrators alike—plagued the schools that served as the context of this study. Finally, district staff changed PBIS fidelity measures during the study period. District staff used the BOQ for the first year of PBIS implementation and the TFI for the two subsequent years. School teams can use both measures to determine if the core features of PBIS are in place and to develop action plans. In their study on PBIS fidelity measures, Mercer et al. (2017) found that there were no statistically significant mean differences between the TFI and the BOQ. While there were no statistically significant differences, there are possible limitations to the change in fidelity measures.

One of the outcome variables we examined was the number of suspension events as reported in the student information system instead of the number of suspension days. The number of suspension days and the number of students contributing to the suspensions may have been more comprehensive data points for analysis. The suspension events could have prompted just one day out of school versus extended days of five or more. The number of suspension days provides more information on the number of days a student is out of school and missing instruction. Additionally, if we had looked at the number of students who were suspended, this may have provided better insight into the discipline needs of the schools. The number of students suspended may have been small when compared to the total enrollment of the schools. These students may need the more intensive supports provided in PBIS Tiers II and III. Future research may focus on the number of students suspended to determine if it is the same students receiving suspensions. Future research could explore the number of suspension days to determine how much instructional time is lost due to suspensions and the impact on students.

Future research could improve on our data collection procedures. While we analyzed school and district data verified by district personnel, we did not examine data collection procedures in individual schools to determine if these changed after PBIS training. School personnel may have made changes in their data collection procedures due to a better understanding of data collection from PBIS training. Future studies could factor in data collection procedures prior to training to determine if schools or districts made changes in their data collection procedures to define and organize their data after PBIS training. Furthermore, other data sources, such as attendance rates and student satisfaction surveys, may be useful. All of the schools reflected in our data were *majority-minority* serving schools and subject to a corrective action plan (CAP) due to their disproportionate suspension rates of African-American students. However, the ODR data provided to us were disaggregated by type (ISS, OSS), but not by race. Future research studies may examine data disaggregated by student subpopulations to determine whether there are significant decreases among students of color in schools that implement PBIS.

Future studies may focus on schools who choose to implement PBIS or compare schools who choose PBIS as an intervention as opposed to those on which it was forced (as was the case in our study). Schools who choose willingly to implement PBIS with fidelity, then outcomes may be more likely to improve. Qualitative studies with key stakeholders (e.g., district leaders, building

leaders, teachers, counselors, parents, and students) may improve our understanding and perceptions of PBIS implementation as well as the pitfalls and successes of PBIS implementation.

As PBIS implementation continues to expand across schools and districts, additional factors may be considered. This study focused on PBIS Tier I but did not consider the other tiers in the framework. Future studies may examine the number of students who require Tier II and III interventions, especially in urban settings. Students in urban settings have complex needs due to environments in which they may live. The students may require intensive supports to make positive changes in behaviors. Additionally, the staff may require more training and technical assistance to meet the complex needs of students living in urban settings. Future studies may examine classroom PBIS implementation more closely to determine if teachers are teaching desired behaviors and recognizing positive behaviors. These potential studies could analyze teacher responses to student misbehaviors to determine which behaviors ultimately lead to referral or suspension. The results could provide information regarding minor behavioral issues and teachers' responses to them.

References

- Algozzine, B., Barrett, S., Eber, L., George, H., Horner, R. H., Lewis, T., Putnam, B., Swain-Bradway, J., McIntosh, K., & Sugai, G. (2014). *School-wide PBIS Tiered Fidelity Inventory*. OSEP Technical Assistance Center on Positive Behavior Interventions and Supports. Retrieved from [https://www.pbisapps.org/Resources/SWIS%20Publications/SWPBIS%20Tiered%20Fidelity%20Inventory%20\(TFI\).pdf](https://www.pbisapps.org/Resources/SWIS%20Publications/SWPBIS%20Tiered%20Fidelity%20Inventory%20(TFI).pdf)
- Arcia, E. (2006). Achievement and enrollment status of suspended students' outcomes in a large, multicultural school district. *Education and Urban Society*, 38(3), 359-369.
- 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.
- Bohanon, H., Fenning, P., Carney, K., Minnis, M., Anderson-Harris, S., Moroz, K., Hicks, K., Kasper, B., Culos, C., Sailor, W., & Pigott, T. (2006). School-wide application of urban high school positive behavior support. *Journal of Positive Behavior Interventions and Supports*, 8(3), 131-145.
- Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2010). Examining the effects of school-wide positive behavior 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.
- Childs, K., Kincaid, D., & George, H. (2010). A model for statewide evaluation of a universal positive behavior support initiative. *Journal of Positive Behavior Interventions*, 12(4), 198-210.
- Childs, K., Kincaid, D., George, H., & Gage, N. (2016). The relationship between school-wide implementation of positive behavior intervention and supports and student discipline outcomes. *Journal of Positive Behavior Interventions*, 18(2), 89-99.
- Cohen, R., Kincaid, D., & Childs, K. (2007). Measuring school-wide positive behavior support implementation: Development and validation of the benchmarks of quality. *Journal of Positive Behavior Interventions*, 9(4), 203-213.
- Costenbader, V., & Markson, S. (1998). School suspension: A study with secondary school students. *Journal of School Psychology*, 36(1), 59-82.
- Flannery, K., Fenning, P., McGrath Kato, 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(2), 111-124.
- Freeman, J., Simonsen, B., McCoach, B., Sugai, G., Lombardi, A., & Horner, R. H. (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.
- Gregory, A., Skiba, R., & Noguera, P. (2010). The achievement gap and the discipline gap: Two sides of the same coin. *Educational Researcher*, 39(1), 59-68.
- Horner, R. H., Sugai, G., & Anderson, C. (2010). Examining the evidence base for school-wide positive behavior support. *Focus on Exceptional Children*, 42(8), 1-14.

- Irvin, L., Tobin, T., Sprague, J., Sugai, G., & Vincent, C. (2004). Validity of office discipline referral measures as indices of school-wide behavioral status and effects of school-wide behavioral interventions. *Journal of Positive Behavioral Interventions*, 6(3), 131-147.
- Mathews, S., McIntosh, K., Frank, J., & May, S. (2014). Critical features of predicting sustained implementation of school-wide positive behavioral interventions and supports. *Journal of Positive Behavior Interventions*, 16(3), 168-178.
- McIntosh, K., Girvan, E. J., Horner, R. H., Smolkowski, K., & Sugai, G. (2014). *Recommendations for addressing discipline disproportionality in education*. OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports. Retrieved from https://static1.squarespace.com/static/56b90cb101dbae64ff707585/v/57e2c31d03596e4c714a1ed6/1474478878439/Keynotehandout1_McIntosh.pdf
- 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.
- McIntosh, K., Massar, M., Algozzine, R., George, H., Horner, R. H., Lewis, T., & Swain-Bradway, J. (2017). Technical adequacy of the SWPBIS Tiered Fidelity Inventory. *Journal of Positive Behavior Interventions*, 19(1), 3-13.
- Mercer, S., McIntosh, K., & Hoselton, R. (2017). Comparability of fidelity measures for assessing tier I school-wide positive behavioral interventions and supports. *Journal of Positive Behavior Interventions*, 19(4), 195-204.
- Morris, E., & Perry, B. (2016). The punishment gap: School suspension and racial disparities in achievement. *Social Problems*, 63(1), 68-86
- OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports (October 2015). *Positive Behavioral Interventions and Supports (PBIS) Implementation Blueprint: Part 1 – Foundations and Supporting Information*. Eugene, OR: University of Oregon. Retrieved from <https://www.pbis.org/resource/pbis-implementation-blueprint>
- Pas, E. T., Bradshaw, C. P., & Mitchell, M. M. (2011). Examining the validity of office discipline referrals as an indicator of student behavior problems. *Psychology in the Schools*, 48(6), 541-555.
- Perry, B. L., & Morris, E. W. (2014). Suspending progress: Collateral consequences of exclusionary punishment in public schools. *American Sociological Review*, 79(6), 1067-1087.
- Raffaele-Mendez, L. (2003). Predictors of suspension and negative school outcomes: A longitudinal investigation. In J. Wald & D. J. Losen (Eds.), *New directions for youth development* (no. 99; Deconstructing the school-to-prison pipeline) (pp. 17-34). San Francisco, CA: Jossey-Bass.
- Raffaele-Mendez, L., & Knoff, H. (2003). Who gets suspended from school and why: A demographic analysis of schools and disciplinary infractions in a large school district. *Education and Treatment of Children*, 26(1), 30-51.
- 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 outcomes across years. *Journal of Positive Behavior Interventions*, 14, 5-16.
- Skiba, R., Horner, R. H., Chung, C. G., Rausch, M., May, S., & Tobin, T. (2011) Race is not neutral: A national investigation of African American and Latino disproportionality in school discipline. *School Psychology Review*, 40(1), 85-107.
- Skiba, R., Michael, R., Nardo, A., & Peterson, R. (2002). The color of discipline: sources of racial and gender disproportionality in school punishment. *Urban Review*, 34(4), 317-342.
- Skiba, R., & Rausch, M. (2006). Zero tolerance, suspension, and expulsion: Questions of equity and effectiveness. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 1063-1089). Mahwah, NJ: Lawrence Erlbaum Associates.
- Skiba, R., & Sprague, J. (2008). Safety without suspension. *Educational Leadership*, 66(1), 38-43.
- Stevens, J. (2007). *Intermediate Statistics: A Modern Approach*. New York, NY: Routledge.
- Sugai, G., & Horner, R. H. (2008). What we know and need to know about preventing problem behavior in schools, *Exceptionality*, 16(2), 67-77.
- Sugai, G., & Horner, R. H. (2009). Defining and describing school-wide positive behavior support. In W. Sailor, G. Dunlap, G. Sugai & R. H. Horner (Eds.), *Handbook of positive behavior support* (pp. 307-324). New York, NY: Springer Science & Business Media.
- Sugai, G., Lewis-Palmer, T., Todd, A., & Horner, R. H. (2001). *School-wide evaluation tool*. Eugene, OR: University of Oregon.

- 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, H. R., III, Wickham, D. Reuf, M., & Wilcox, B. (2000). Applying positive behavioral support and functional behavioral assessment in schools. *Journal of Positive Behavioral Interventions*, 2(3), 131-143.
- Swain-Bradway, J., Swoszowski, N., Boden, L., & Sprague, J. (2013). Voices from the field: Stakeholder perspectives on PBIS implementation in alternative educational settings. *Education and Treatment of Children*, 36(3), 31-46.
- United States Departments of Education & Justice. (2014). Dear colleague letter on nondiscriminatory administration of school discipline. Retrieved from <https://www2.ed.gov/about/offices/list/ocr/letters/colleague-201401-title-vi.html>
- Wald, J., & Losen, D. J. (2003). Defining and redirecting a school-to-prison pipeline. *New Directions for Youth Development*, 2003(99), 9-15.
- Warren, J., Edmonson H., Griggs, P., Lassen, S., McCart, A., Turbull, 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.