Effects of a Multifaceted Classroom Intervention on Racial Disproportionality

Cody Gion

Gresham-Barlow School District

Kent McIntosh

Sarah Fairbanks Falcon

University of Oregon

Gion, C., McIntosh, K., & Falcon, S. F. (2020). Effects of a multifaceted classroom intervention on racial disproportionality. *School Psychology Review.*

Author note: The development of this research was supported by the Institute of Education Sciences, U.S. Department of Education (#R324A170034). The opinions expressed are those of the authors and do not represent views of the Office or U.S. Department of Education.

Address correspondences to: Cody Gion, gion@gresham.k12.or.us, 1550 NW Eastman Parkway Gresham, OR 97030

Abstract

This article reports results from an experimental study of a classroom intervention intended to decrease racial disproportionality in school discipline by focusing on making classroom behavior systems more culturally responsive and changing teacher behaviors (e.g., use of praise and reprimands). The intervention had three main components: (a) an activity to increase the cultural responsiveness of classroom behavior expectations (the personal matrix), (b) a tool to assess individual students' preferred and non-preferred ways to receive immediate acknowledgement for desired behaviors (the praise preference assessment), and (c) coaching with visual performance feedback to heighten awareness of racial disparities in use of praise and reprimands and set goals for racial equity (using a modified version of the Classroom Check-up). The design was a concurrent multiple-baseline single-case design across four general education teachers ranging from kindergarten to seventh grade meeting What Works Clearinghouse standards. Results from the study indicate a functional relation between the intervention and increased rates of praise and decreased rates of reprimands for African American students. Findings from this study suggest that changing teacher use of praise and reprimands may help to increase racial equity in schools.

Effects of a Multifaceted Classroom Intervention on Racial Disproportionality

Research continues to demonstrate the harmful effects of high rates of exclusionary discipline practices (e.g., Office Discipline Referrals, suspensions, expulsions) on students (Noltemeyer, Ward, & Mcloughlin, 2015). For example, students who experience high rates of exclusion are more likely to have future behavioral problems, dropout, and be involved with juvenile justice systems than students who experience none or low rates (Nicholson-Crotty, Birchmeier, & Valentine, 2009; Raffaele Mendez, 2003). When students are removed from the educational environment, they miss critical academic and social content, and unwanted behavior may be reinforced by escaping aversive interactions or unwanted tasks. As a result, it is clear that exclusionary discipline is ineffective in improving student behavior and instead exacerbates behavioral and academic skill deficits (American Academy of Pediatrics Council on School Health, 2013; Morris & Perry, 2016).

The Impact of Race on Exclusionary Discipline

The most disturbing characteristic of the overuse of exclusionary discipline is that it is disproportionally applied to students who are African American. In the United States, African American students are two-to-three times more likely to receive Office Discipline Referrals (ODRs), be suspended, and be expelled than any other racial group (Anyon et al., 2014; Losen, Hodson, Keith, Morrison, & Belway, 2015). Hence, African American students are placed disproportionately at risk for the negative outcomes of exclusionary discipline.

To understand this problem, it is important to consider the variables that may be contributing to racial inequities. Decades of research have examined patterns of

disproportionate discipline to both dispel some popular explanations and identify promising avenues for intervention. For example, multiple studies continue to show that although poverty is a risk factor, race remains a significant predictor of exclusionary discipline even when controlling for socioeconomic status (Anyon et al., 2014; Skiba et al., 2014). Another explanation, that African American students have higher base rates of unwanted behavior, has not been in found in recent research (Scott, Gage, Hirn, & Han, 2019). Other variables, such as attending a predominantly African American school (Hughes, Warren, Stewart, Tomaskovic-Devey, & Mears, 2017), are powerful predictors but less actionable for school psychologists and school teams.

A more detailed examination of patterns of disproportionality point to the classroom as the primary source of disproportionate discipline. The classroom is the location with the highest rates of both exclusionary discipline and disproportionate discipline (Gion, McIntosh, & Horner, 2014; Smolkowski, Girvan, McIntosh, Nese, & Horner, 2016). Additionally, research into the effects of school-wide positive behavioral interventions and supports (SWPBIS) have shown that its implementation is associated with reduced discipline disproportionality (McIntosh, Gion, & Bastable, 2018; Vincent & Tobin, 2011), but SWPBIS alone is not adequate to eliminate racial disproportionality in school discipline. Implementation of classroom SWPBIS systems, however, is the strongest predictor of equity in school discipline compared to other core elements of SWPBIS, such as school-wide expectations, recognition systems, and consistent discipline procedures (Tobin & Vincent, 2011). Thus, a focus on supporting teachers in their interactions with students in the classroom may have the most direct impact on discipline disproportionality.

Understanding Racial Disproportionality in the Classroom

There are a number of potential explanations for why discipline disproportionality is most prevalent in the classroom. An emerging area of research examines the role of educator implicit (i.e., unconscious) bias in discipline disproportionality (Staats, Capatosto, Wright, & Contractor, 2015). Implicit bias represents the cognitive shortcuts that help us react quickly to stimuli, and result from learning histories, media depictions, or incomplete experiences that lead to racial stereotypes. Essentially, teachers may not be aware of inequities in their decisions and actions toward students, leading to disproportionate use of exclusionary discipline (McIntosh, Girvan, Horner, & Smolkowski, 2014). Evidence for the role of implicit bias comes from the patterns of strongest disproportionality for subjective behaviors (e.g., defiance, disrespect, disruption) which require a value judgment, as opposed to objective behaviors (e.g., smoking, truancy), which allow for clear decisions (Girvan, Gion, McIntosh, & Smolkowski, 2017).

If implicit bias leads teachers to watch African American students more closely for misbehavior (Gilliam, Maupin, Reyes, Accavitti, & Shic, 2016), teachers may be more likely to discipline African American students for similar behavior (Scott, Gage, Hirn, & Han, 2018). This inequitable treatment could create an aversive classroom environment for students, resulting in a coercive cycle of increased resistance from African American students and increased use of exclusionary discipline by their teachers (Okonofua, Walton, & Eberhardt, 2016). Hence, implicit bias may influence classroom teacher behavior in a number of ways, including the way they structure the classroom environment and interact with students.

Ameliorating Racial Disproportionality in the Classroom

Because implicit biases operate outside of conscious decision-making, it may be necessary to help educators become aware of their biased interactions. Utilizing existing proven classroom-based innovations, while simultaneously adapting elements to improve awareness, may be an effective strategy for improving equity for African American students (Chu & Leino, 2017). One effective classroom-based intervention for improving classroom environmental outcomes is the Classroom Check-up (CCU; Reinke, Herman, & Sprick, 2011; Reinke et al., 2008). Critical elements of the CCU model are motivational interviewing, teacher implementation of evidence-based classroom management strategies (e.g., explicitly teaching expectations, specific acknowledgment for prosaical behaviors), and ongoing visual performance feedback to teachers.

Motivational Interviewing. The CCU model is predicated on the coach's use of motivational interviewing techniques to build rapport with the teacher and to improve acceptance, effectiveness, and contextual fit of intervention implementation. Motivational interviewing is a non-confrontational approach aimed at collaborative effort to initiate change (Miller & Rollnick, 2002). Due to the sensitivity and complexity of racial disproportionality, a motivational interviewing approach might improve the probability of effective strategies being accepted and implemented with fidelity.

Adaptations for teaching classroom behavior expectations. One way that implicit bias may affect judgment is how teachers define what behaviors are acceptable and unacceptable in the classroom. Common in SWPBIS, defining and teaching a small number of positively-stated behavior expectations is an effective intervention for preventing unwanted classroom behavior (Epstein, Atkins, Cullinan, Kutash, & Weaver, 2008; Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008). However, these definitions

may categorize some behavior from African American students as “inappropriate” in a school setting, whereas it may be an acceptable way of demonstrating values such as safety, respect, or responsibility outside of school (Leverson, Smith, McIntosh, Rose, & Pinkelman, 2016). For example, it may be against the rules to use physical force to defend yourself at school, but in the neighborhood this is an acceptable action. Hence, educators can take steps to learn about their students’ cultures and norms when defining and teaching classroom expectations, creating greater awareness of cultural differences (Emdin, 2016).

Adaptations for acknowledging prosocial behavior. Given research showing that that fidelity of implementation of SWPBIS acknowledgment systems is significantly related to disciplinary equity (i.e., schools with greater use of formal acknowledgement systems had lower disproportionality; Barclay, 2017; Tobin & Vincent, 2011), it seems that increasing praise may be a key component of creating equitable classroom environments. However, there is some research to suggest that type of acknowledgement delivered might matter for students of color (Yeager et al., 2014). Hence, it may be helpful to provide educators with tools to identify students’ preferred forms of acknowledgement (e.g., public, private, tangible, group acknowledgement) to ensure it is reinforcing, particularly for students with whom teachers do not have strong relationships or common interests (Fefer, DeMagistris, & Shuttleton, 2016).

Adaptations to coaching to visual performance feedback. Recent research shows that classroom coaching is a promising avenue to improving equity in exclusionary discipline, with teachers receiving coaching issuing fewer ODRs for African American students (Bradshaw et al., 2018; Gregory et al., 2016). In addition, using disaggregated

data may be a critical mechanism for understanding and improving racial inequities in school discipline (Bastable & McIntosh, 2019; McIntosh, Ellwood, McCall, & Girvan, 2018). Providing teachers visual feedback on their implementation of intervention components has been supported to improve the implementation of effective classroom management (MacSuga & Simonsen, 2011; Solomon, Klein, & Politylo, 2012). However, no research to date has studied the approach of disaggregating student interactions (i.e., praise and reprimand) data by race as a coaching tool.

Purpose

The field is in need of evidence-based interventions to address persistent racial inequities in school discipline. Although some approaches show promise, modifying commonly-used evidence-based approaches (i.e., Classroom Check-Up) to address racial disparities more explicitly could be an avenue for widespread application of equity intervention. The purpose of this study was to test an adaption of the CCU model consisting of (a) providing tools to educators to reflect on the impact bias may play on the development of classroom behavior expectations, (b) providing tools to identify what types of acknowledgement students valued, and (c) providing coaching with visual performance feedback on inequitable treatment of students. It was hypothesized that these tools would surface inequitable treatment, motivating teachers to provide more equitable interactions in their classrooms. Specifically, the study examined the following primary (experimental) research questions:

1. Is there a functional relation between implementation of the intervention and an increase in teacher use of praise for African American students?

2. Is there a functional relation between implementation of the intervention and a decrease in teacher use of reprimands for African American students?

In addition, the study addressed the following secondary (descriptive) research questions:

3. To what extent do teachers find the intervention socially valid?

Method

Setting

Two public neighborhood schools (pseudonyms used throughout) in an urban district located in the Pacific Northwest were approached to be sources for participating teachers and classrooms. Demographic data were obtained through each school's most current state report card. Acadia school was a K-8 school with an enrollment of 451 students (White = 55%, African American = 17%, Hispanic/Latinx = 13%, Asian = 1%), and 29% of these students received free or reduced-price meals. Maple Park was a K-5 school with an enrollment of 334 students (African American = 42%, White = 23%, Hispanic/Latinx = 23%, Multi-Racial = 8%, Asian = 1%, Hawaiian/Pacific Islander = 1%, Native American/Alaska Native = 1%), and 100% of the students in this school received free or reduced-price meals.

ODR records prior to participation were examined to confirm the extent of racial disproportionality in exclusionary discipline. Disproportionality was determined by the ODR Risk Ratio, calculated by dividing the percent of African American students receiving 1 or more ODRs by the percent of all other students receiving 1 or more ODRs (the U.S. Federal Government's recommeded metric; Girvan, McIntosh, & Smolkowski, 2019). Acadia had a risk ratio of 2.67, and Maple Park had a risk ratio of 2.55, meaning

that in both schools, African American students were more than two and a half times more likely to receive ODRs than other students.

Participants

Within these schools, the principals nominated teachers meeting two criteria. First, principals identified teachers needing assistance in classroom behavior support or equity in treatment of students using their own judgement. Second, the classroom racial diversity needed to be adequate to observe racial equity in teacher-student interactions (i.e., between 25% and 75% of students in the class were African American). Nominated teachers were told they would be participating in a study to reduce discipline disproportionality, consisting of classroom observations and implementing an intervention. These teachers were not told what behaviors would be observed in their classrooms or what intervention they would be implementing. After explaining the purpose and commitment required to participate in the study, four of the five nominated teachers agreed to participate, with one declining based on lack of availability for meetings outside of school hours due to child care needs.

Participants for this study included four general education classroom teachers, two from each school. Sofia and Martina (pseudonyms used throughout) were recruited from Maple Park, and Alma and Orien were recruited from Acadia. Classroom racial compositions of the class (i.e., the race of each student) were derived from teacher reports. The researchers chose to use this method over others (e.g., obtaining a class roster), because teacher's perceptions of which students were African American may be the most accurate measure of potential bias. Sofia was a fifth-year teacher, who identified as Pacific Islander/Asian and female. She taught a classroom of 21 second-grade students (African

American = 12, Other = 9). Martina was a 17th-year classroom teacher who identified as Hispanic/Latina and female. She had 15 kindergarten students (African American = 6, Other = 9). Alma was a third-year teacher who identified as Hispanic/Latina and female. She taught fifth grade and had a classroom of 28 total students (African American = 5, Other = 23). Orien was a first-year teacher who identified as White and male. He taught a seventh-grade class of 21 students (African American = 6, Other = 15).

Measures

Observation of teacher behavior. The researcher and a trained observer collected frequency counts of praise and reprimands using the operational definitions of teacher behaviors from the Brief Classroom Interaction Observation – Revised (BCIO-R) measure (Reinke, Stormont, Herman, Wachsmuth, & Newcomer, 2015), with one modification to directly measure racial equity in teacher behaviors. Separate frequency counts of teacher behaviors were tallied based on student race. Observers determined student race as either African American or All Other (i.e., not African American) through an initial conference with the teacher where each participant identified the African American students in their class, based on each teacher's perception.

Teachers were observed during 20-min daily sessions during the time of day when they indicated unwanted behaviors were most likely to occur. To account for differences in classroom racial composition, Praise and Reprimand rates (per 20-min) were calculated by dividing the frequency of praise for a specified sample (i.e., African American or All Other) by the number of students in the sample present during that day's observational session. For example, if five African American students received a total of 10 reprimands during a 20-min observational session, the reprimand rate would be 2.0.

Data collectors (the first and third authors) used the existing operational definitions of each target behavior (i.e., Praise and Reprimands) from Reinke et al. 2015) and developed examples and non-examples of each behavior. The operational definitions are provided below. Praises and reprimands were tallied only when directed toward a single student, and not when directed to a group or whole class.

Praise. Praise included both general praise and behavior specific praise. The operational definition for general praise (GP) was as follows: verbal statements or gestures that indicate approval and *do not name a specific behavior* (e.g. "Kiss your brain," "Give me a bam," "Good job," giving out reinforcement tokens silently, high five to student, clapping, thumbs up). The operational definition for behavior-specific praise (BSP) was as follows: verbal statements that indicate approval and *name a specific behavior* (e.g., "Thank you for sitting quietly," "Maria is showing me she is ready with her eyes on me"). Non-examples of praise included statements regarding correct academic responding without statements of approval for the behavior (e.g., "Yes, 2 + 2 is 4").

Reprimands. Reprimands included both general and harsh reprimands. General reprimands were verbal comments or gestures by teacher to indicate disapproval of behavior, delivered concisely (briefly) in a normal speaking tone. General reprimands included error corrections, where a teacher responds to a social behavior error with the correct response provided by the teacher (e.g., "Tim you need to put your book away and begin working."). Harsh reprimands were reprimands using a voice louder than typical for setting; a harsh, critical, or sarcastic tone; or a reprimand lasting longer than 30-s. Non-

examples of reprimands included stating expectations before activities (i.e., precorrections) and corrections for academic errors (e.g., “That word is Renaissance”).

Classroom Fidelity Checklist. In addition to frequency counts of praise and reprimands, a researcher-developed observational checklist (Appendix A) was completed by observers during each session across all phases (i.e., baseline and intervention) to indicate the presence of defining elements, such as teaching expectations and analysis of the quality of praise and reprimands to inform intervention implementation across phases. Data collectors used a rubric to document fidelity before and after intervention and to select elements of intervention for teachers.

Interobserver Agreement (IOA). The first author met with the additional data collector prior to the study and established reliability through direct observations of non-target classrooms in a neighboring school district. Each observer was required to obtain 85% reliability agreement for all variables before collecting data for the study. IOA was calculated using the total count IOA method for each direct observation measure. Total count IOA was calculated by dividing the smaller total count observed (from one observer, relative to the other) by the larger total count (from the other observer). IOA data were collected for 38% of all sessions across teachers and phases (Range 29% to 60%). Average IOA was 90% for Praise (Range 76% to100%) and 94% for Reprimands (Range 75% to 100%). For the classroom fidelity checklist, total IOA was 85% (Range 75% to 100%). This frequency of data collection and IOA exceeds methodological WWC SCD guidelines that specify IOA needing to be collected for at least 20% of sessions across participants and phases with reliability above at or above 80% (Kratochwill et al.,

2013). Observers met and came to consensus on operational definitions for any sessions with IOA that fell below 80% before the next observation.

Coaching logs. The first author served as the coach for the teachers in this study. He kept logs by the minute and coded each coaching activity completed during both the baseline and intervention phases as a method to assess procedural fidelity. The coach spent an average of 12-hr and 15-min per participant engaging in data collection (*M* = 6hr 15min), meeting with teachers (*M* = 3hr 8min), providing performance feedback via email (*M* = 1hr 25min), scheduling activities (*M* = 36min), and preparing materials (*M* = 29min).

Primary Intervention Rating Scale (PIRS). The PIRS was administered to teachers after intervention completion, and included items pertaining to the acceptability, effectiveness, and contextual fit of the intervention. The PIRS includes 17 items, each rated on a scale of 1 to 6 (Strongly Disagree to Strongly Agree). This scale has previously been used as a descriptive measure of social validity (Lane, Robertson, & Wehby, 2002). The PIRS has demonstrated strong internal consistency (.97 or higher), and high ratings on the PIRS are predictive of higher levels of fidelity of implementation (Lane et al., 2009).

Procedure

Experimental design. A concurrent multiple-baseline, single-case design across four teachers was used in this study. It consisted of two phases (baseline and intervention). The start order in which teachers received intervention was randomly assigned using a random number generator. To meet WWC single-case design standards without reservations, the design allowed for at least three demonstrations of effect, at three

different points in time, and each participant's baseline and intervention phases had at least five data points (Kratochwill et al., 2013). As determined before the study, the intervention was introduced in staggered fashion after at least five data points and stability in baseline responding.

Intake meeting. The first author held an intake meeting with each teacher to obtain consent, build rapport, and ask the teacher about their (a) experience, (b) values, (c) management style, (d) ideal classroom, and (e) past coaching experiences following procedures from the Classroom Check-Up (CCU; Reinke, Lewis-Palmer, & Merrell, 2008). Additionally, the coach and the teacher identified the best time to conduct classroom observations (i.e., time where unwanted behavior was most likely to occur).

Baseline phase. Teachers were urged to provide instruction as usual during baseline. Observers collected frequency data of praise and reprimands, as well as fidelity data using the classroom fidelity checklist. Observations were 20-min each and occurred during the same times of day as the intervention phase. Instruction was similar for each participant across phases.

Intervention phase. The intervention package was selected based on results from baseline data (both the classroom fidelity checklist and the teacher praise and reprimands data). All four teachers had the same general patterns of data. On the classroom fidelity checklist, all scored low on teaching expectations, specifically referring to expectations before and throughout instruction. All had low ratios of praise to reprimands (below 1.0) for both groups during the baseline phase, indicating a need for strategies to increase praise overall and for African American students in particular. The generally low rates of reprimands and almost non-existent harsh reprimands, combined with high baseline scores

for instructional use of reprimands on the classroom fidelity checklist, made potential intervention elements for calm responses to unwanted behavior irrelevant. However, African American students received more reprimands, indicating the need to examine classroom expectations and review their performance. As such, all four teachers received the same intervention package of (a) personal matrix activity (to support culturally responsive expectations), (b) praise preference assessment (to support use of praise), and (c) coaching with visual performance feedback (to increase equity in praise and reprimands).

Personal matrix activity. To help teachers reflect on their classroom expectations and any differences with students, teachers were introduced to a personal matrix activity to implement with their class (Leverson et al., 2016). Each student completed a worksheet that resembled a SWPBIS expectations by settings teaching matrix (Horner, Sugai, Todd, & Lewis-Palmer, 2005), but with only one column completed for school-wide expectations and blank cells for "at home" and "in my neighborhood" (Appendix B). Students were asked to complete the worksheet by identifying what the school/class-wide expectations looked like in their homes and in their neighborhoods. For example, respect might look like raising a hand to answer a question in the classroom but may look like helping mom with the dishes at home. Teachers used the information they gathered from this activity to reflect on their classroom expectations and how they may be similar or different from each individual student's life outside of school. Teachers were encouraged to clarify any significant differences between home and school, make connections across settings using similar language, and adjust their expectations, as needed, to better align with students' background knowledge and cultural values. All four teachers shared the

completed worksheets with the first author to document fidelity of implementation (i.e., via permanent product).

Praise preference assessment. To increase teacher knowledge regarding praise and their students, teachers conducted a praise preference assessment, a survey to better understand the type of acknowledgment that individual students reported as most (and least) desirable, somewhat akin to a reinforcer preference survey, but focusing on the immediate teacher response, not a backup reward. Each teacher generated a list of potential responses to desired student behavior that they would be willing to use (e.g., public praise, private praise or secret gesture, provision of a school-wide reward ticket, edible) and created a questionnaire in which students marked most and least desired teacher responses. Teachers were encouraged to use the information to tailor their reinforcement to meet individual student needs and to increase the use of strategies that were highly preferred, both class-wide and for individual students with whom the teacher lacked strong relationships. For example, teachers increased their use of public praise or school-wide tickets to reinforce desired behavior based on the results of this assessment. All four teachers shared the assessment results with the first author to document fidelity of implementation (i.e., via permanent product).

Coaching with visual performance feedback. An adapted CCU coaching model was used to support intervention implementation throughout the intervention phase. The CCU has been used effectively to change rates of teacher praise and reprimands in previous research (Reinke et al., 2008). The coaching model consisted of (a) an initial meeting (described in Intake meeting); (b) data collection on praise and reprimand ratios and fidelity of implementation of classroom practices in the classroom fidelity checklist;

(c) an action planning meeting where baseline data were reviewed, teacher strengths and weaknesses were discussed, and a goal and an action plan was put in place; (d) visual performance feedback delivered via email after each observational session; and (e) a follow-up meeting to review progress toward goals and make adjustments, as needed.

There were two significant modifications made to the CCU model. First, the researcher reviewed the data from the study-specific classroom fidelity checklist with the teacher, as opposed to the classroom management forms used in previous research (Reinke et al., 2008). Second, praise and reprimand data were presented as disaggregated data by student race (i.e., African American vs. All Other) to depict differences in teacher-student interactions by race. Otherwise, the format of the CCU remained the same, with the key coaching behaviors of motivational interviewing and visual performance feedback.

Action-planning meeting. The first author held an individual action-planning meeting with each teacher at the end of each baseline phase, before the implementation of the intervention components. To prepare for this meeting, the first author compiled the baseline observational data into visual formats. The frequency data of praise and reprimands were graphed separately by race of student as ratios (praise to reprimands for African American students and praise to reprimands for All Other). The reason for choosing this method of presentation was two-fold. First, most teachers are familiar with the "magic ratio" of praise statements to corrections, thus it was hypothesized that this format would be easier to understand than a rate-based representation. Second, the first author wanted to draw attention to the relation between praise and reprimands and avoid misrules of (a) providing frivolous praise to increase rates or (b) avoiding providing

reprimands when necessary. Separation on the graph between African American and All Other students would indicate disproportionality in praise to reprimand ratios.

The action-planning meeting consisted of four discrete steps presented in a guided format. The meeting began with a review of the purpose of the study and rationale as to why measuring praise and reprimands are important for improving student outcomes. Teachers were given documents describing (a) what behavior specific praise is, (b) why it is an important strategy to implement, and (c) tips to help make implementation easier. Following this brief review, the prepared data were reviewed. The coach explained the data to the teacher using a straightforward, neutral tone. Teachers were encouraged to ask clarifying questions and to provide their perspective regarding whether the data seemed to be an accurate representation of what was observed in their classroom.

After the data were reviewed and discussed, the coach and the teacher established a goal based on the data. Each teacher set a goal of achieving a praise to reprimand ratio of over 1.0 for African American students, indicating students were receiving more praise than reprimands during the observational period. The teachers were also introduced to the personal matrix activity and praise preference assessment. Teachers were encouraged to use the basic format of the personal matrix and praise preference activities but adapt them to improve contextual fit for their classroom. Orien (7th grade) and Alma (5th grade) chose to have students complete both activities independently during class. Sofia (2nd grade) chose to complete the activities in a guided small group format, and Martina (Kindergarten) chose to do the activities individually with each student in an interview format. The meeting concluded with establishing action plans with established timelines to enhance the clarity in expected behavior for both the teacher and the coach.

Visual performance feedback. In this study, the first author emailed graphs of disaggregated praise to reprimand ratios and recommendations for improvement after each observation session. The first author used a standard email format consisting of one to three behaviors that were noticed during the observational session (e.g., "You praised other students for following directions before correcting often") and one to three recommendations for improvement (e.g., "Continue to give more praise to students who may be struggling to follow directions") for every observation conducted during the intervention phase. Feedback included brief strategies related to increasing behavior-specific praise or to restate behavioral expectations when providing a reprimand (i.e., to make the response instructional). The email included graphs of the disaggregated praise to reprimand ratios, the current action plan, and the classroom fidelity checklist for reference.

Follow-up meeting. The first author had one follow-up meeting with each teacher within two weeks of initial implementation. The structure of this meeting was to review outcome and fidelity data, the previous action plan, and to create new action items as needed. Action plan revision was based on consistent achievement of previously established goals. Based on performance, additional action items were not needed for Sofia, Martina, or Orien. The coach worked with Alma to revise the original action plan and included the items of providing a script for teaching expectations and a script for responding to unwanted behavior. The previously established goal of a praise to reprimand ratio above 1.0 stayed the same.

Data Analysis

Systematic visual analysis was conducted to determine a functional relation between the intervention and outcome data. Data were analyzed through inspection of level, trend, and variability for each phase and by each participant. Then, vertical analysis was conducted to determine the effect of the intervention across participants.

In addition to visual analysis, we used a design-comparable standardized effect size statistic (Hedge's *g*) designed for single-case research to determine the effect of the intervention on the outcomes of teacher praise and reprimands for African American students. This statistic is equivalent to the usual *d* statistic in between-groups designs and is appropriate for use in single-case multiple baseline designs (Hedges, Pustejovsky, & Shadish, 2013). Criteria for determining the size of the effect of the intervention on specific target outcome variables (i.e., teacher praise and reprimand rates) was determined as follows: small ($g = 0.20$), medium ($g = 0.50$), and large ($g = 0.80$; Cohen, 1977). ODR and social validity data were examined descriptively.

Results

Our primary anticipated outcome was an increase in praise rates and a reduction in reprimands for African American students compared to all other students. Results are provided by dependent variable.

Teacher Praise to Reprimand Ratios (Table 2)

Teacher praise to reprimand ratios were calculated by dividing the sum of praise statements by the sum of reprimands for each group (i.e., African American and All Other) in each session. The results depicted in Table 2 show average ratios for each teacher below 1.0 in the baseline phase and above 1.0 in the intervention phase (the intervention goal). For African American students, three of four teachers met their 1.0

goal in each intervention phase session. However, these data should be interpreted with caution, as ratio data tend to be unstable for repeated comparisons (Girvan et al., 2019). The rate measures described subsequently allow more accurate comparisons across groups.

Teacher Use of Praise (Figure 1)

Sofia. In baseline, Sofia's rates of praise were extremely low and stable across both groups (African American $M = 0.37$ per student, All Other $M = 0.15$), with a slight decreasing trend for African American students. There were few differences in praise by race. Upon intervention, there was a strong immediacy of effect for both groups, with an increase in level (African American $M = 1.04$ per student, All Other $M = 0.86$) and higher, but more variable, praise rates for African American students.

Martina. In baseline, Martina's rates of praise were low, with slight variability across both groups (African American $M = 0.66$ per student, All Other $M = 0.39$) and a stable trend for both African American and All Other students. On average, African American students received more praise than All Other students. Upon intervention, there was a strong immediacy of effect for both groups, with an increase in level (African American $M = 2.73$ per student, All Other $M = 1.98$) and higher praise rates for African American students. Praise for African American students also showed an increasing trend throughout the intervention phase.

Alma. In baseline, Alma's rates of praise were low, with slight variability across both groups (African American $M = 0.69$ per student, All Other $M = 0.50$), and a stable trend for both African American and All Other students. There were few differences in praise by race. Upon intervention, there was a no immediacy of effect for either group, but

there was an increase in trend and a slight increase in variability throughout the intervention phase, resulting in increased in levels of praise for both groups (African American $M = 1.28$ per student, All Other $M = 0.87$), with a stronger effect for African American students.

Orien. In baseline, Orien's rates of praise were extremely low and stable, with a stable trend, across both groups (African American $M = 0.08$ per student, All Other $M = 0.07$). There were few differences in praise by race. Upon intervention, there was an immediate of effect for both groups, resulting in increased in levels of praise for both groups (African American $M = 0.98$ per student, All Other $M = 0.69$), with a slightly stronger effect for African American students.

Vertical analysis. Data showed an immediate positive effect in praise rates across three of the four participants (Sophia, Marina, and Orien). When the intervention was introduced, the effect of the intervention on increasing praise for both racial groups was seen for the teacher receiving intervention, but praise rates for other teachers in baseline remained low. These data demonstrate an immediate effect of the intervention for three of four teachers at three different points in time for both racial groups, supporting a functional relation between the implementation of the intervention and increased praise rates for African American students.

Teacher Use of Reprimands (Figure 2)

Sofia. In baseline, Sofia's rates of reprimands were moderate to high, with substantial variability and a stable trend for both groups (African American $M = 1.06$ per student, All Other $M = 1.09$). There was little differentiation in reprimands by race. Upon intervention, there was a strong immediacy of effect for African American students and no

immediate effect for All Other students; both groups showed a decrease in level (African American $M = 0.18$ per student, All Other $M = 0.50$), with a stronger effect for African American students. Both groups showed less variability and a stable trend across the intervention phase.

Martina. In baseline, Martina's rates of reprimands were high for African American students and moderate to high for other students (African American $M = 1.98$ per student, All Other $M = 0.81$), African American student data had more variability than other student data, and both groups showed an increasing trend across the baseline phase. On average, African American students received more reprimands than other students. Upon intervention, there was a strong immediacy of effect in terms of a decrease for African American students and no immediate effect for other students. Both groups showed a significant decrease in level (African American $M = 0.49$ per student, All Other $M = 0.39$) and less variability during intervention. There was a decreasing trend for other students, with a stable trend for African American students throughout the intervention phase.

Alma. In baseline, Alma's rates of reprimands were high for African American students and moderate for other students (African American $M = 2.56$ per student, All Other $M = 0.72$), with a stable trend for both groups. African American student data showed more variability and African American students received more reprimands than other students during baseline. Upon intervention, there was an immediate effect for both groups, with a substantial change in level for the African American group only (African American $M = 1.28$ per student, All Other $M = 0.66$). There was a stable trend for African American students and a slightly increasing trend for All Other students. African

American student data also showed less variability during intervention, and reprimands were more equitable across groups, compared to baseline.

Orien. In baseline, Orien's rates of reprimands were high for African American students and moderate for other students (African American $M = 1.48$ per student, All Other $M = 0.63$). African American students received more reprimands on average than All Other students during baseline. Both groups had variable data, with a stable trend. Upon intervention, there was an immediate effect for both groups, resulting in decreased levels of reprimands for both groups, (African American $M = 0.31$ per student, All Other $M = 0.07$). Data during intervention were stable with low variability. Additionally, reprimand rates were more equitable across racial groups during the intervention phase, when compared to baseline.

Vertical analysis. In addition to examining the effect of the intervention for individual teachers, it is important to compare the effects of the intervention across participants. Data showed an immediate decrease in reprimand rates for African American students when each teacher received the intervention, whereas rates of reprimands for the other participants still in baseline remained high. This effect was obtained across all participants in the study, indicating a functional relation between intervention implementation and decreased rates of reprimands for African American students.

Second, reprimand differences for African American students compared to other students remained discrepant throughout the baseline phase for Martina, Alma, and Orien, but not for Sophia. When the intervention was introduced, there was an immediate effect, resulting in little differentiation between groups for the teacher receiving intervention, but not for the remaining teachers in baseline, supporting the demonstration of a functional

relation between the implementation of the intervention and improved equity in reprimands between racial groups.

Summary of Visual Analysis

Data indicated a functional relation (i.e., three demonstrations of effect at three different points in time) between the intervention and an increase in the level of praise for African American students and for All Other students. Based on the results from visual analysis, there were no substantial differences between groups during the intervention phase. The results from statistical analysis also support these claims, as evidenced by identical, large effects from the Hedge's *g* analysis (African American = 1.12, All Other = 1.12).

Additionally, data indicated a strong functional relation between the intervention and decrease in reprimands for both groups, with a stronger effect for African American students. Hedge's *g* results for reprimands were -1.16 for African American students and -0.97 for All Other students, both large effects that supported the results from visual analysis.

Primary Intervention Rating Scale (PIRS)

The PIRS was administered post-intervention after study completion to obtain teacher feedback about the (a) acceptability, (b) effectiveness, and (c) contextual fit of the intervention. Overall, participants in this study scored the intervention as acceptable, effective, and a good fit for their setting (mean = 5.32 on a scale of 1 to 6, between Agree and Strongly Agree).

Discussion

The purpose of this study was to provide an experimental test of the efficacy of a multicomponent intervention to increase disciplinary equity by changing teacher-student interactions with African American students from negative to positive. The intervention included a brief set of feasible strategies to make classroom systems more culturally responsive and increase attention to equity in praise and reprimands, with data-driven coaching to ensure fidelity of implementation. The training and coaching increased fidelity of implementation, and implementation of the intervention package resulted in increased praise and decreased reprimands for African American students. Upon training, the intervention increased the generally low rates of praise for all students, including African American students, and decreased or eliminated the racial gap in reprimands. Finally, the intervention was rated as socially valid by teachers.

It was hypothesized that disproportionate rates of exclusionary discipline could be the result of disproportionate rates of praise and reprimands for African American students. Findings from this study suggest that all students received low rates of praise, regardless of race, and African American students were more likely to receive higher rates of reprimands when compared to students from other racial backgrounds, for three out of four teachers. Additionally, although rates of exclusion were low, all instances of exclusionary discipline occurred for African American students. The findings of higher rates of reprimands for African American students in baseline is consistent with previous research (Scott et al., 2019) and potentially supports teacher-student interactions as the basis of contributing to a coercive cycle of inequity for disciplinary discipline (Okonofua et al., 2016).

Results from this study provide initial support for the potential of this intervention to increase praise rates and decrease reprimand rates for African American students. These students received more praise than reprimands, on average, during the intervention phase, as opposed to the opposite pattern in baseline.

Teachers implementing the classroom intervention thought it was acceptable, effective, and fit well within their school and classroom contexts. This finding, coupled with the high fidelity of implementation, suggests that practitioners could find this intervention useful and feasible to implement in their local settings.

Contributions to the Field

Recent intervention research in the area of discipline disproportionality has focused largely on exclusionary discipline (Bradshaw et al., 2018, Cook et al., 2018; Gregory et al., 2016), but this study examined the effects of an intervention to reduce precursor behaviors (i.e., teacher-student interactions) that may be contributing to discipline disproportionality. This focus may allow for further development of preventive intervention strategies that are more effective and less resource intensive than intervention focused on solely on behaviors that result in exclusion from the educational environment. Additionally, the finding that African American students received more negative feedback (i.e., reprimands) than other students before intervention is also consistent with previous research (Scott et al., 2018). This might suggest that reducing negative teacher-student interactions for African American students is an appropriate and feasible avenue to increasing equity in the classroom.

Finally, this study adds to the growing body of literature of promising classroom-based interventions to reduce discipline disproportionality (e.g., Bradshaw et al., 2018,

Cook et al., 2018; Gregory et al., 2016). The findings from this study support a multicomponent approach rooted in effective coaching strategies, with a reliance on the use of disaggregated data to guide decision-making that has also been supported in previous research (McIntosh, Ellwood, et al., 2018). Somewhat unique to this study is that the intervention package consisted of adaptations of existing, evidence-based interventions (SWPBIS and CCU) that are already employed in schools across the U.S.

Limitations and Implications for Future Research

Although the results of this study are promising, there are some considerable limitations that need to be expressed. First, the dependent variables in this study were teacher behavior, and there was no repeated measure of the effects of the intervention on student behavior. The ODR data, although promising, was descriptive and had low rates before intervention, making these results tentative. Future research will need to examine these effects to be able to determine the extent to which this intervention has a significant impact on student classroom behavior and more distal measures of equity beyond the descriptive effects on ODRs shown in this study.

Second, the study included a small sample, and more demonstrations of these effects need to be studied to establish this intervention as an evidence-based practice to improve disciplinary equity for African American students. For single-case research, Kratochwill et al. (2013) described the need for findings to be replicated with at least 20 participants, across at least five separate studies, and examined by at least two different research groups to establish a practice as an evidence-based intervention. One single-case study, as reported here, can provide credence and support to the approach used in this

study to eliminate discipline disproportionality, but it would be inappropriate to draw firm conclusions from only this study.

Third, the first and third authors served as data collectors for this study and thus were not blind to the intervention. Although steps were taken to improve objectivity of the data being collected (i.e., establishing IOA in more sessions than is recommended), findings from this study may be influenced by confirmation bias. Future research should examine the replication of these findings with data collectors who are blinded to the intervention and the aims of the study.

Fourth, the coaching intervention components were implemented by the first author, who has significant experience in coaching classroom management. The results of this study need to be replicated with coaches with varying backgrounds and experiences to determine the generalizability of these findings. In addition, future research should examine the feasibility and acceptability of the coaching components by coaches who are internal to the school environment. There may be drastic differences in terms of feasibility and acceptability from coaches within these contexts.

Finally, the intervention focused on individual classroom implementation, independent of larger school systems. Implementation of intervention in isolation, without consideration for systematic support, is limited in its impact and sustainability (Scheirer, 2005; Wiltsey Stirman et al., 2012). Future research and implementation of the classroom intervention will have to be considered within a larger context of system implementation. Because the intervention elements were adaptations of SWPBIS, such an effort seems feasible.

Intervention research to reduce exclusionary discipline disproportionality is in its infancy, but the initial results are promising. Disproportionality is a multifaceted and complex problem, requiring various interventions depending on environmental contexts. This study is merely one supporting branch in a larger body of research. Findings and conclusions made here need to be taken into context within the larger and ever-changing educational context. Future research will want to take steps to further refine the mechanisms and contributing factors that may be contributing to discipline disproportionality.

Researchers could examine more precisely the coercive cycle of teacher-student interactions that may be leading to high rates of exclusion. It would be helpful to know to what extent these interactions lead to exclusion or increased engagement. It would also be helpful to know to what extent other elements of classroom management and instruction contribute to exclusionary discipline.

Implications for Practice

Practitioners will want to use their knowledge of their context and their local data to determine potential contributing factors to discipline disproportionality. After identifying the root cause of disproportionality, they may choose to implement some or all of the intervention components described here or other promising approaches that have been developed. They may want to test varying methods for effectiveness and feasibility, guiding the development of supports that are equitable and inclusive for all students.

References

American Academy of Pediatrics Council on School Health. (2013). Policy statement: Out-of-school suspension and expulsion. *Pediatrics, 131*, e1000-e1007. doi:10.1542/peds.2012-3932

Anyon, Y., Jenson, J. M., Altschul, I., Farrar, J., McQueen, J., Greer, E., . . . Simmons, J. (2014). The persistent effect of race and the promise of alternatives to suspension in school discipline outcomes. *Children and Youth Services Review, 44*, 379-386.

Barclay, C. M. (2017). *Benchmarks of equality? School-wide positive behavior interventions and supports and school discipline risk and disparities for Black and Hispanic students*. (Unpublished doctoral dissertation). University of South Florida, Tampa, FL.

Bastable, E., & McIntosh, K. (2019). Understanding educators' commitment to equity in school discipline. *Manuscript submitted for publication*.

Bradshaw, C. P., Pas, E. T., Bottiani, J. H., Debnam, K. J., Reinke, W. M., Herman, K. C., & Rosenberg, M. S. (2018). Promoting cultural responsivity and student engagement through Double Check coaching of classroom teachers: An efficacy study. *School Psychology Review, 47*(2), 118-134.

Chu, J., & Leino, A. (2017). Advancement in the maturing science of cultural adaptations of evidence-based interventions. *Journal of Consulting and Clinical Psychology*, *85*(1), 45.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences (Rev. ed.). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Emdin, C. (2016). *For White folks who teach in the hood... and the rest of y'all too: Reality pedagogy and urban education*. New York: Beacon Press.

Epstein, M., Atkins, M., Cullinan, D., Kutash, K., & Weaver, R. (2008). *Reducing Behavior Problems in the Elementary School Classroom: A Practice Guide (NCEE #2008-012)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides.

Fefer, S., DeMagistris, J., & Shuttleton, C. (2016). Assessing adolescent praise and reward preferences for academic behavior. *Translational Issues in Psychological Science, 2*, 153-162. doi:10.1037/tps0000072

Gilliam, W. S., Maupin, A. N., Reyes, C. R., Accavitti, M., & Shic, F. (2016). *Do early educators' implicit biases regarding sex and race relate to behavior expectations and recommendations of preschool expulsions and suspensions?* Research Study Brief. Yale University, Yale Child Study Center, New Haven, CT.

Gion, C., McIntosh, K., & Horner, R. (2014). Patterns of minor office discipline referrals in schools using SWIS *PBIS Evaluation Brief. OSEP National Technical Assistance Center on Positive Behavioral Interventions and Supports. Eugene, OR.*

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. doi:10.1037/spq0000178

Girvan, E. J., McIntosh, K., & Smolkowski, K. (2019). Tail, tusk, and trunk: What different metrics reveal about racial disproportionality in school discipline. *Educational Psychologist*. doi:10.1080/00461520.2018.1537125

Gregory, A., Hafen, C. A., Ruzek, E., Mikami, A. Y., Allen, J. P., & Pianta, R. C. (2016). Closing the racial discipline gap in classrooms by changing teacher practice. *School Psychology Review, 45*(2), 171-191.

Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2013). A standardized mean difference effect size for multiple baseline designs across individuals. *Research Synthesis Methods, 4*, 324-341.

Horner, R. H., Sugai, G., Todd, A. W., & Lewis-Palmer, T. (2005). School-wide positive behavior support. In L. Bambara & L. Kern (Eds.), *Individualized supports for students with problem behaviors: Designing positive behavior plans* (pp. 359-390). New York: Guilford Press.

Hughes, C., Warren, P. Y., Stewart, E. A., Tomaskovic-Devey, D., & Mears, D. P. (2017). Racial threat, intergroup contact, and school punishment. *Journal of Research in Crime and Delinquency*. doi:https://doi.org/10.1177/0022427816689811

Irvin, L. K., Tobin, T. J., Sprague, J. R., Sugai, G., & Vincent, C. G. (2004). Validity of office discipline referral measures as indices of school-Wide behavioral status and effects of school-wide behavioral interventions. *Journal of Positive Behavior Interventions, 6*(3), 131-147.

Kratochwill, T., R. , Hitchcock, J., H., Horner, R., H., Levin, J., R., Odom, S., L. , Rindskopf, D., M., & Shadish, W., R. . (2013). Single-case intervention research design standards. *Remedial and Special Education, 34*(1), 26-38. doi:10.1177/0741932512452794

Leverson, M., Smith, K., McIntosh, K., Rose, J., & Pinkelman, S. (2016). PBIS cultural responsiveness field guide: Resources for trainers and coaches.

Losen, D. J., Hodson, C., Keith, M. A., Morrison, K., & Belway, S. (2015). Are we closing the school discipline gap? *Center for Civil Rights Remedies at The Civil Rights Project at UCLA*. Los Angeles, CA.

MacSuga, A. S., & Simonsen, B. (2011). Increasing teachers' use of evidence-based classroom management strategies through consultation: Overview and case studies. *Beyond Behavior, 20*(2), 4-12.

May, S. L., Ard, W. I., Todd, A. W., Horner, R. H., Glasgow, A., Sugai, G., & Sprague, J. R. (2013). School-Wide Information System. Educational and Community Supports, University of Oregon, Eugene, OR. Retrieved from http:/www.pbisapps.org

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.

McIntosh, K., Gion, C., & Bastable, E. (2018). *Do schools implementing SWPBIS have decreased racial disproportionality in school discipline?* PBIS evaluation brief. OSEP TA Center on Positive Behavioral Interventions and Supports. Eugene, OR.

McIntosh, K., Girvan, E. J., Horner, R. H., & Smolkowski, K. (2014). Education not incarceration: A conceptual model for reducing racial and ethnic disproportionality in school discipline. *Journal of Applied Research on Children, 5*(2), 1-22.

Morris, E. W., & Perry, B. L. (2016). The punishment gap: School suspension and racial disparities in achievement. *Social Problems, 63*, 68-86.

Nicholson-Crotty, S., Birchmeier, Z., & Valentine, D. (2009). Exploring the impact of school discipline on racial disproportion in the juvenile justice system. *Social Science Quarterly, 90*, 1003-1018.

Noltemeyer, A. L., Ward, R. M., & Mcloughlin, C. S. (2015). Relationship between school suspension and student outcomes: A meta-analysis. *School Psychology Review, 44*, 224-240. doi:10.17105/spr-14-0008.1

Okonofua, J. A., Walton, G. M., & Eberhardt, J. L. (2016). A vicious cycle: A social–psychological account of extreme racial disparities in school discipline. . *Perspectives on Psychological Science, 11*, 381-398.

Raffaele Mendez, L. M. (2003). Predictors of suspension and negative school outcomes: A longitudinal investigation. *New Directions for Youth Development, 2003*(99), 17-33.

Reinke, W. M., Lewis-Palmer, T., & Merrell, K. (2008). The Classroom Check-up: A classwide teacher consultation model for increasing praise and decreasing disruptive behavior. *School Psychology Review, 37*(3), 315-332.

Reinke, W. M., Stormont, M., Herman, K. C., Wachsmuth, S., & Newcomer, L. (2015). The Brief Classroom Interaction Observation–Revised: An observation system to inform and increase teacher use of universal classroom management practices. *Journal of Positive Behavior Interventions, 17*(3), 159-169.

Scheirer, M. A. (2005). Is sustainability possible? A review and commentary on empirical studies of program sustainability. *American Journal of Evaluation, 26*, 320-347.

Scott, T. M., Gage, N., Hirn, R., & Han, H. (2018). Teacher and student race as a predictor for negative feedback during instruction. *School Psychology Quarterly*, Advance online publication. . doi:10.1037/spq0000251

Scott, T. M., Gage, N., Hirn, R., & Han, H. (2019). Teacher and student race as a predictor for negative feedback during instruction. *School Psychology, 34*, 22-31. doi:10.1037/spq0000251

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children, 31*(3), 351-380.

Skiba, R. J., Chung, C.-G., Trachok, M., Baker, T. L., Sheya, A., & Hughes, R. L. (2014). Parsing disciplinary disproportionality: Contributions of infraction, student, and school characteristics to out-of-school suspension and expulsion. *American Educational Research Journal, 51*, 640-670. doi:10.3102/0002831214541670

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.

Solomon, B. G., Klein, S. A., & Politylo, B. C. (2012). The effect of performance feedback on teachers' treatment integrity: A meta-analysis of the single-case literature. *School Psychology Review, 41*, 160-175.

Staats, C., Capatosto, K., Wright, R. A., & Contractor, D. (2015). *State of the science: Implicit bias review 2015*. Columbus, OH: Kirwan Institute.

Tobin, T. J., & Vincent, C. G. (2011). Strategies for preventing disproportionate exclusions of African American students. *Preventing School Failure, 55*, 192-201. doi:10.1080/1045988X.2010.532520

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

Wiltsey Stirman, S., Kimberly, J., Cook, N., Calloway, A., Castro, F., & Charns, M. (2012). The sustainability of new programs and innovations: a review of the empirical literature and recommendations for future research. *Implement Sci, 7*, 7-17. doi:10.1186/1748-5908-7-17

Yeager, D. S., Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., . . . Cohen, G. L. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General, 143*, 804-824.

Table 1

Classroom Fidelity Checklist Percent Implementation Across Phases

	Teaching Expectations			Praise			Reprimands		
	Baseline	Intervention	Difference	Baseline	Intervention	Difference	Baseline	Intervention	Difference
Sophia	89%	95%	+ 6%	52%	89%	+ 37%	90%	94%	+ 4%
Martina	73%	99%	+ 26%	42%	96%	+ 54%	86%	95%	+ 9%
Alma	44%	60%	+ 16%	39%	45%	+ 6%	52%	67%	+ 15%
Orien	39%	94%	+ 55%	33%	92%	+ 59%	78%	95%	+ 17%

Table 2

Ratio of Teacher Praise to Reprimands by Race in Baseline and Intervention Phases

	African American Students				All Other Students			
	Baseline		Intervention		Baseline		Intervention	
	Average P:R Ratio	Percent of Sessions ≥ 1.0	Average P:R Ratio	Percent of Sessions ≥ 1.0	Average P:R Ratio	Percent of Sessions ≥ 1.0	Average P:R Ratio	Percent of Sessions ≥ 1.0
Sophia	0.35	0%	5.12	100%	0.19	0%	2.97	91%
Martina	0.38	14%	10.11	100%	0.63	29%	7.94	100%
Alma	0.33	8%	1.08	25%	0.71	17%	1.69	63%
Orien	0.20	8%	3.50	100%	0.24	15%	8.20	100%

Note. Average P:R Ratio = praise / reprimands (< 1.0 = more reprimands than praise). Percent of Sessions ≥ 1.0 = percent of sessions with ratios at or above 1.0 (each teacher's goal).

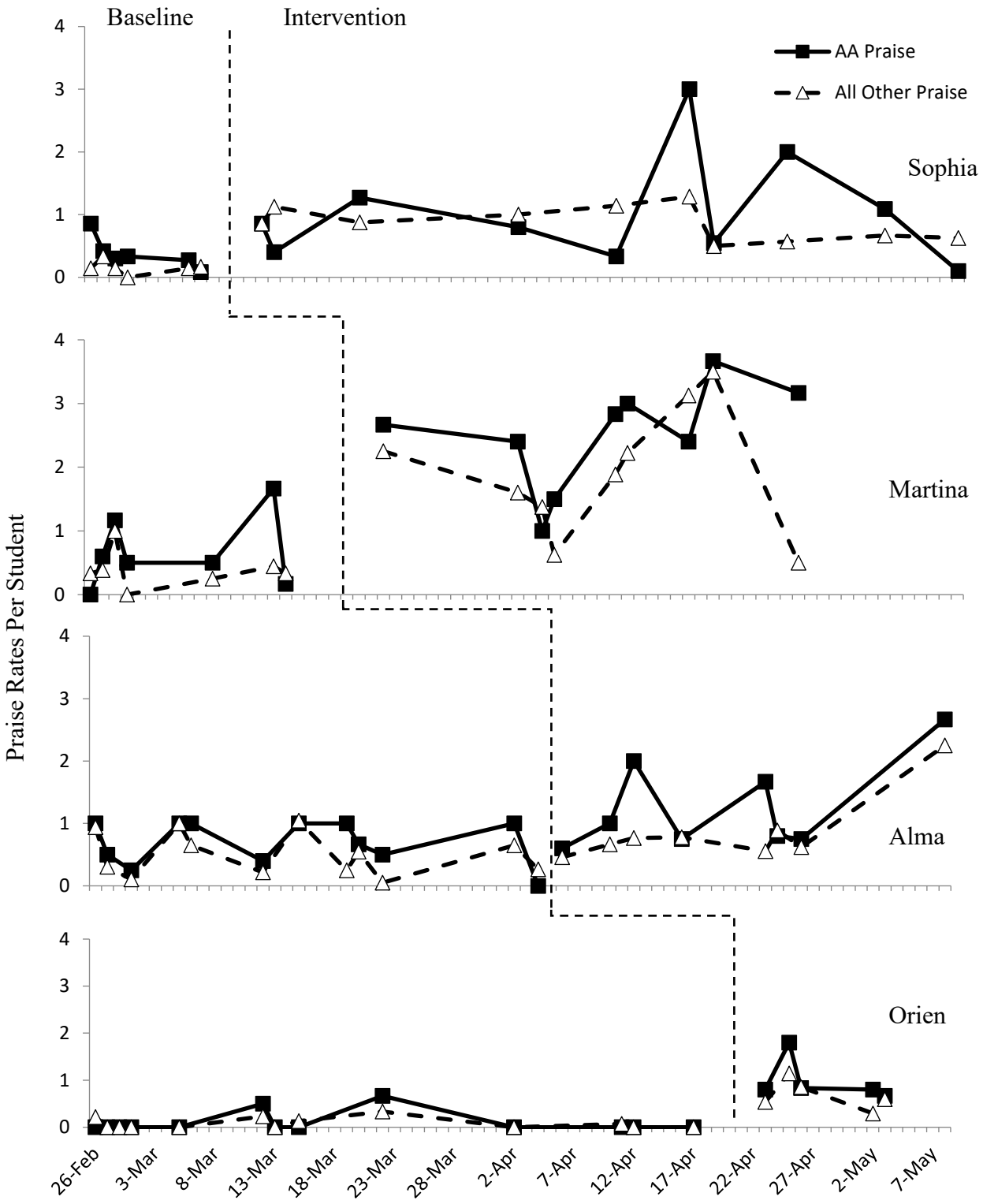

Figure 1. Frequency Rates for Teacher Praise Per Student Across Classrooms

Teacher praise rates per student across four classrooms for African American (AA) students and students who were not African American (All Other).

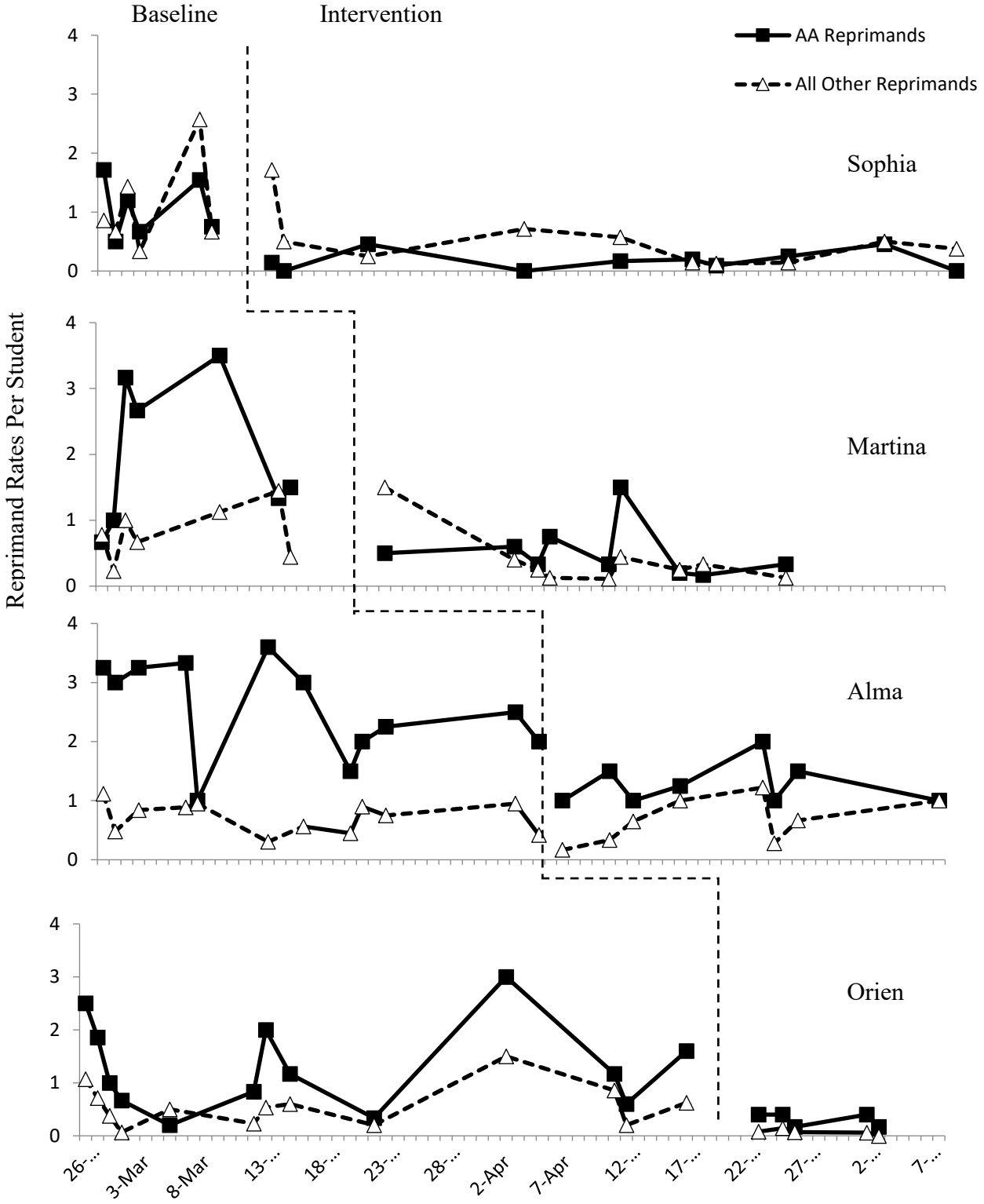

Figure 2. Frequency Rates for Teacher Reprimands Per Student Across Classrooms

Teacher reprimand rates per student across four classrooms for African American (AA) students and students who are not African American (All other).

APPENDIX A

CLASSROOM FIDELITY CHECKLIST

Element	Never	Inconsistent	Consistent	N/A
Teaching Culturally Responsive Expectations				
1. Expectations are posted	Not Posted		Posted	
2. States clear expectations before directions (e.g., when we start our math lesson, I want you to have your voices off and your materials out and ready, if you need help please raise your hand)	Never states specific behavioral expectations during the observation.	States expectations for some but not all activities.	States expectations for all activities.	
3. States or refers to class-wide or school-wide expectations	Never states or refers to posted expectations.	States or refers to **posted** expectations for some but not all activities.	States or refers to **posted** expectations for all activities	Expectations are not posted
Praise				
4. Greets students at the door by name.	Greets		Doesn't greet	Entering is not observed
5. Uses or refers to more than one strategy to acknowledge student behavior (e.g., praise, point system)	Uses only one strategy (e.g., verbal praise)		Uses or refers to more than one strategy (e.g., points, tickets, praise, gestures).	
6. Provides praise/acknowledgement more frequently than correction	Correction is more frequent than acknowledgement for AA **and** Other Students	Praise is more frequent than correction for AA **or** Other Students	Praise is more frequent than correction for both AA **or** Other Students	
7. When problems occur, uses praise around strategy	Does not use praise around strategy	Uses praise around strategy, but has missed opportunities to use praise around strategy	Uses praise around strategy every time when appropriate	No problem behavior
8. Scans and interacts with students throughout the observation	Teacher stays at desk or only interacts with one student or group of students	Scans the room but spends a significant amount of time	Constantly scans the room and interacts with the entire	

		with a few students	class (i.e., does not get bogged down).	
9. Moves around between students during the observation	Teacher remains in one area throughout the observation	Teacher moves around but spends a significant amount of time in one location	Teacher covers the whole room and does not get bogged down	Carpet time or small group
Reprimands				
10. Uses more than one strategy to correct student behavior (modeling, proximity)	Teacher only provides verbal correction		Teacher uses multiple strategies to correct behavior (e.g., proximity, crouching by student, verbal correction, modeling, gesturing).	No Corrections
11. Corrects behavior quickly, explicitly, quietly, & as situationally inappropriate, not wrong	Consistently uses harsh reprimands to correct behavior.	Generally, corrects behavior explicitly, but has harsh reprimands.	Does not use harsh reprimands	No Corrections
12. Provides specific feedback or practice in response to social and academic behavior errors	Teacher frequently does not state desired behavior (e.g., Shhh, no don't) when correcting behavior.	Teacher has students model desired behavior and states desired behavior when correcting about half the time.	Teacher has students model desired behavior and states desired behavior when correcting almost every time.	No Corrections
13. Uses friendly and firm tone	Teacher uses harsh tone consistently	Teacher generally uses a friendly tone, but also uses a harsh tone.	Teacher does not use a harsh tone when correcting	No Corrections
14. De-escalates conflicts	Teacher escalates conflicts that occur.	Teacher does not de-escalate conflicts, but does not make it worse.	Teacher effectively de-escalates conflicts	No conflicts occur

APPENDIX B
Personal Matrix Activity

Expectation	At SCHOOL it looks like…	At HOME it looks like…	In my NEIGHBORHOOD it looks like…
Be Safe	• Keep hands and feet to self • Tell an adult if there is a problem	• Protect your friends and family • Don't talk back	• Stick up for your friends • Don't back down • Look the other way
Be Respectful	• Treat others how you want to be treated • Include others • Listen to adults	• Do exactly what adults tell you to do • Don't stand out • Don't bring shame	• Text back within 30 seconds • Be nice to friends' parents • Share food
Be Responsible	• Do my own work • Personal best • Follow directions • Clean up messes	• Help your family out first • Own your mistakes • Share credit for successes	• Have each other's backs • Own your mistakes • Check in about what to do