Document Title:	**Examining the Efficacy of Circles on School Safety and Student Outcomes in Boston Public Schools: Final Report**
Author(s):	**Roger Jarjoura, Ph.D, Christina LiCalsi, Ph.D, Melissa Yisak, Lauren Stargel, Ph.D, Amelia Auchstetter**
Document Number:	**308394**
Date Received:	**January 2024**
Award Number:	**2017-CK-BX-0013**

Examining the Efficacy of Circles on School Safety and Student Outcomes in Boston Public Schools: Final Report

Roger Jarjoura, Ph.D, Christina LiCalsi, Ph.D, Melissa Yisak, Lauren Stargel, Ph.D, Amelia Auchstetter

American Institutes for Research

National Institute of Justice Grant 2017-CK-BX-0013

MARCH 2023

American Institutes for Research® | AIR.ORG

AIR®
Advancing Evidence.
Improving Lives.

Contents

Exhibits

Introduction

School violence is harmful to student and school staff well-being and mental health as well as detrimental to the learning environment (Capp, Astor & Gilreath, 2020). There is increasing evidence that promoting social-emotional learning, positive teacher-student relationships and a supportive climate can prevent school violence (Bradshaw, Cohen, Espelage & Nation, 2021; Durlak, Weissberg, Dymnicki, Taylor & Schellinger, 2011; Hong, Espelage & Lee, 2018; Nickerson, 2018; Thapa, Cohen, Guffey & Higgins-D'Alessandro, 2013; Voight & Nation, 2016; Wang, Berry & Swearer, 2013). In this context, many schools are turning to restorative practices to promote safe, inclusive, caring schools. Restorative justice, or restorative practices more broadly, are rooted in relationality and Indigenous ways of knowing (Antara, Evans & Lester, 2013; Gregory & Evans, 2020; González, Sattler & Buth, 2019; Vaandering, 2014). Schools that use restorative practices aim to bring students, teachers, staff, family and local residents together to build community and address students' (unmet) needs. Whereas traditional disciplinary responses (e.g., detention) often focus on punishment of the individual who caused harm with neither involvement of the person who was harmed nor discussion of the incident, restorative practices emphasize learning and growth. When harm has occurred, restorative practices emphasize dialogue, accountability for those who caused harm and repair for the individual(s) who were harmed. Restorative practices in schools have been associated with decreased bullying and school violence, improvements in positive school climate and decreased use of suspensions (Boulton & Mirsky, 2006; Gregory et al., 2018; Hantzopoulos, 2013; Ingraham et al., 2016; Jain, Bassey, Brown & Kalkra, 2014; Schumacher, 2014; Zakszeski & Rutherford, 2021). Yet, there are challenges to implementing restorative practices in school-wide initiatives and causal evidence of their effectiveness are limited. The current study presents implementation and impact findings of a cluster randomized control trial examining a school-wide restorative practices model called Circle Forward.

Implementation of restorative practices in schools

Restorative principles reflect beliefs about relationships, justice, mutual understanding and equity that often challenge traditional mindsets in school settings (Boyes-Watson & Pranis, 2014; Antara et al., 2013; Gregory & Evans, 2020; González et al., 2019). For example, restorative principles reframe how student (mis)behavior is understood (Antara et al., 2013; Gregory & Evans, 2020; González et al., 2019). Instead of viewing misbehavior as morally wrong, student behavior can be understood through a developmental lens related to students' unmet needs and emotional self-regulation. Further, restorative principles encourage a focus not on punishment for wrong caused but on repair of both harm in the context of relationships and conflict as a learning opportunity (Antara et al., 2013; González, 2016; Zehr, 2002). More broadly, restorative principles emphasize that individual positive relationships among students, teachers and staff can enhance the

relational health of the school community (Antara et al., 2013; González, 2016; McCluskey, 2018; Vaandering, 2014). Finally, restorative practices also must address systemic power imbalances and structural inequities in school environments and beyond, since these power dynamics are relevant at the interpersonal, social and structural levels (Antara et al., 2013; González, 2016). For example, the use of a talking piece, an object that designates the speaker, can help shift power between teachers and their students, as everyone is expected to respect its use and listen to the speaker (Boyes-Watson & Pranis, 2014).

Circles are a common restorative approach used most frequently to build community or address harm (such as incidents that have damaged a relationship) with students, teachers, other school staff, families and community members. Educators and practitioners applying restorative practices such as circles aim to connect participants and recognize the humanity of all involved (McCammon, 2020). Circles support a relational context in school settings where power dynamics (e.g., adults may wield more power; inequities related to race/ethnicity, gender, sexual orientation and disability) may affect the ability to promote authentic relationships among all members of school communities.

Increasingly, schools are implementing whole-school restorative practices initiatives. To do so often requires changes at many levels, from individual mindsets to school climate and culture to formal practices and policies (González et al., 2019; Gregory & Evans, 2020; Martinez, Villegas, Hassoun Ayoub, Jensen & Miller, 2022). Although prior research has frequently focused on restorative practices as an alternative disciplinary intervention, studies have noted the difficulty of transforming disciplinary practices in the context of isolated interventions that do not address broader structures and policies (e.g., González et al., 2019; Gregory, Ward-Seidel & Carter, 2020; Schiff, 2018). Whole-school restorative practices approaches, on the other hand, include a focus on community building and proactive approaches to harm prevention; they also intentionally focus on the relational and structural context of the school among students, educators, community and family members and the broader district (González et al., 2019), allowing schools greater potential for culture and systems change (Martinez et al., 2022).

Researchers and practitioners have found that implementing whole-school restorative practices brings its own set of school-wide challenges. First, for school-wide restorative practices, buy-in at all levels of the school is needed; this can be challenging, and it takes time (Gregory et al., 2020; González, 2014). Educators may struggle to implement and fully buy into restorative practices because of other competing demands on their staff time (Martinez et al., 2022). In addition, buy-in and ownership from students and other community members is important for whole-school implementation (Garnett, Kervick et al., 2022; Gregory et al., 2020) but often neither prioritized nor measured. Further, few restorative practices implementation studies have examined broader policies at the school and district levels and their relationship with

implementation (Gregory et al., 2020; Martinez et al., 2022). Finally, implementation challenges may vary based on school and district context.

For this reason, whole-school approaches to restorative practices are often shaped by the school and community context (e.g., González et al., 2019; Gregory et al., 2020). This leads to variation in the actual practices and activities in whole-school restorative practice initiatives and how they are staffed and implemented. This variation often exists even among schools implementing a single model, making it especially difficult to understand and evaluate the implementation of restorative practices in schools (Zakszeski & Rutherford, 2021). In turn, for the study of whole-school restorative practices initiatives, the variation in what is implemented has led to a relative dearth of information about what constitutes a successful implementation with fidelity (Gregory & Evans, 2020; Zakszeski & Rutherford, 2021). For example, in a systematic review of research on restorative practices in education, Zakszeski and Rutherford (2021) identified 71 previous studies, but only seven of those described implementation approaches. Of those seven, three reported evaluating any aspect of fidelity of implementation. Taken together, the infrequent, incomplete and inconsistent measurement of implementation affects the ability of researchers and evaluators to assess the effectiveness of restorative practices in schools more broadly.

In recent years, scholars have developed a multi-tiered model to conceptualize both restorative practices and implementation supports at multiple levels, with clear implementation supports for a whole-school approach (Garnett, Moore et al., 2020; Mayworm et al., 2016). For example, Mayworm and colleagues (2016) present a multi-tiered model of restorative practices implementation that includes school-wide, universal restorative practices (Tier 1), targeted professional development (group consultation; Tier 2) and one-on-one teacher consultation (Tier 3). Gregory and colleagues (2020) identified twelve practice-based qualitative indicators of restorative practices implementation within the three tiers, including restorative practices infrastructure and capacity building as foundational for all tiers (e.g., practice and policy level; staff restorative mindset development and growth; inclusion of students, family and community members). These approaches consider multiple levels of schools' ecology and implementation supports at each level. However, these conceptual models do not actually define measurement approaches for whole-school restorative practices. In terms of actual measurement of fidelity of implementation, there is a lack of attention to the whole-school context: some researchers have assessed implementation of individual practices (e.g., the quality of responsive circles; Wang & Lee, 2019), or a single dimension of implementation fidelity (Acosta et al., 2016). Thus, researchers and educators engaged in whole-school restorative practices initiatives need rigorous research that systematically examines restorative practices implementation across tiers or levels.

Evidence supporting restorative practices in schools

Despite the increased implementation of restorative practices in schools, there is scant causal evidence of their effectiveness. A recent review of school-based restorative practices studies published in academic journals between 2000 and 2020 (Zakszeski and Rutherford, 2021) found only two peer-reviewed studies used a cluster randomized controlled trial design. The first, a two-year study in which 40 middle schools in England were randomly assigned to either implement a restorative practices and social-emotional learning intervention (20 schools), or business-as-usual (20 schools), found that students in schools assigned to the intervention group experienced less bullying victimization, less substance abuse (e.g. smoking, alcohol), and had better psychological well-being (Bonell et al., 2018). No statistically significant differences were found between students in the intervention or business-as-usual schools in aggression. The second, a two-year CRCT in which 13 middle schools in Maine were randomly assigned, seven to implement a restorative practices intervention, and six to business-as-usual found no statistically significant differences between students on any of 11 outcomes related to social skills, bullying victimization, school connectedness, or school climate (Acosta et al., 2019).

Results of a third CRCT, published by the Rand Institute, which was conducted with 44 public schools in Pittsburg (22 treatment and 22 control) found that students in intervention schools reported improved perceptions of school climate and fewer days suspended compared to students in control schools. Further, the reductions in days suspended for were greater for Black students as compared to White students and were driven by reductions in suspensions in the elementary but not the middle or high schools. The study did not find any changes in overall student behavior, student arrest rates, or daily attendance. In addition, math achievement of middle school students, in intervention schools with a predominantly Black student enrolment, worsened (Augustine et al, 2018).

Results of a recent CRTC conducted with 18 K-12 schools in a Northeastern US urban school district found no differences in the likelihood of suspension between students in the intervention and business-as-usual schools after one year of the intervention (Huang, Gregory, and Ward-Seidel, 2023). These results were similar for students regardless of race/ethnicity, gender, or student disability status. However, for students who had previously been suspended, students in the intervention schools were less likely to receive an out-of-school suspension.

The Circle Forward Model of Whole-School Change

The Suffolk University Center for Restorative Justice (CRJ) developed the Circle Forward Model of Whole-School Change (hereafter referred to as 'CF model'; Boyes-Watson & Pranis, 2014). Circle Forward is a multi-tiered model of restorative practices influenced by Indigenous values of community and interconnectedness, and using the circle as its main tool in building relationships among school community members and resolving conflicts.

Circles begin and end with ceremonies so participants can forge connections with one another. By focusing on community building, circles seek to address issues through a strengths-based approach, rather than immediately delving into the conflict. The following elements comprise a CF circle: (a) opening ritual (e.g., a poem); (b) use of a talking piece, an item used to designate whose turn it is to speak if they choose; (c) use of a centerpiece, which includes items that represent core values (e.g., community); (d) a facilitator who oversees the space and guides the discussion; (e) decision-making by consensus which promotes equity and (f) a closing ceremony marking the end of the circle. During the circle, participants are expected to follow circle guidelines and norms, including waiting to speak until they are holding the talking piece, giving all participants the opportunity to share and being present in the circle through active listening or sharing if they choose. Within a circle, all participants are viewed as equals and their contributions carry the same weight. The flow of the circle seeks to honor and nourish all parts of the human self: emotional (e.g., feelings), mental (e.g., thoughts and self-reflection), spiritual (e.g., purpose/values) and physical (e.g., meeting the needs of your body; Boyes-Watson, 2005).

CF trains school community members in three types of circles, with each tier increasing in intensity of use. In Tier 1 circles, participants establish norms, build connections between students and their peers and school staff and address academic and non-academic topics. In Tier 2 circles, facilitators resolve less serious issues, such as classroom disruptions and interpersonal conflicts, using restorative questions. In Tier 3 circles, administrators, deans or other staff who oversee school discipline may hold a harm circle or conference to address serious harm and serve as an alternative to exclusionary discipline practices, such as suspensions.

A core component of the CF model is the restorative leadership team (RLT). The team consists of approximately five to ten school staff (e.g., teachers, guidance and support staff, administrators), as well as caregivers and students. Coaches employed by CRJ comprise another component of the CF model. The work of the RLT begins with a retreat, where CRJ coaches provide an overview of restorative practices and their role on the team, and provide opportunities for staff to connect with each other. At the end of the retreat, the RLT drafts a vision of how restorative practices can be implemented in their schools. After the retreat, the RLT works closely with their assigned CRJ coach to determine how to implement the CF model within the context of the school. Although each coach has the flexibility to guide and support schools based on their needs, in this study all CRJ coaches were expected to do the following: Lead the RLT meetings, act as a model and resource for the implementation of circles for the school, lead staff professional development trainings focused on restorative practices and visit the school on a weekly basis. CF follows a gradual release model, with the RLT continuing to implement CF once the CRJ coach has left the school.

Two-day trainings are a core CF model component, with trainings focused on each circle tier. All staff who interact directly with students (administrators, teachers and guidance and support staff) are expected to attend the Tier 1 circle training, which covers the key concepts of restorative practices and circles, teaches core elements of circles and gives participants the tools to plan and lead circles as well as address common challenges (e.g., time). After the Tier 1 training, staff then would have the tools to lead circles, such as those focused on classroom norms or building relationships among staff and students. The Tier 2 training focuses on classroom management and de-escalation techniques. Only certain staff (e.g., student support, guidance, administrative and some teaching staff) are invited to the Tier 2 training, based on the discretion of the school. At the end of Tier 2 training, attendees should have the knowledge and resources to implement restorative responses to classroom disruptions, use restorative language and understand how to handle larger classroom concerns using circles. Lastly, the Tier 3 training focuses on restorative processes, such as circles conferences, for addressing conflict and harm. School staff who oversee discipline, such as administrators and deans, are invited to participate in Tier 3 training.

Study Overview

As part of the National Institute of Justice's (NIJ's) Comprehensive School Safety Initiative grant, the American Institutes for Research (AIR) partnered with CRJ and a large, urban district to conduct a mixed-methods cluster randomized control trial (RCT) evaluation of the CF model. This evaluation has two components: an implementation evaluation and an impact evaluation. The implementation evaluation drew upon data collected from teachers, students, administrators, and other stakeholders using surveys, interviews, focus groups, observation data, and attendance records. The impact evaluation drew upon administrative data collected by the partner school district.

CRJ led the implementation of the CF model, which included scheduling and leading the circle trainings, and overseeing the work of CRJ coaches in intervention schools. The district supported the implementation of CRJ by assisting in the recruitment of schools and supporting study data collection activities. AIR led the RCT external evaluation of the CF model, which included an implementation and impact evaluation.

The goal of the implementation evaluation was to examine how CF was implemented, how participants perceive the program and the extent to which implementation was aligned with existing programs and school policies. The research questions were as follows:

1. To what extent is CF implemented as intended, and how does implementation vary across schools?
2. What facilitators and challenges to program implementation arose? How did schools resolve challenges?
3. In what ways does implementing the CF model lead to changes in practice for all of those trained?

4. How is CF integrated into and coordinated with preexisting programs and aligned to school policies and practices, so as to become routine?

The goal of the impact evaluation was to examine whether implementation of the CF model led to positive impacts on school climate and student behavior and academic achievement. The research questions were as follows:

1. Do students in schools implementing CF report more positive perceptions of school climate than students in schools not implementing CF?
2. Do schools implementing CF see improvements in attendance?
3. Do schools implementing CF experience reductions in use of suspensions? Are there reductions in disparity in use of suspensions by race?
4. Do schools implementing CF see improvements in student academic achievement?

Study Approach

The CRJ worked with district leaders to recruit from among the highest-need middle and high schools. Schools were targeted for recruitment if they received Title I funding, had a high percentage of children from low-income families, and were among the lowest performing 20% of schools. AIR employed a mixed-methods, randomized controlled trial to study the implementation and impact of the Circle Forward model among high-need schools serving middle and high school students.

In this section, we further describe the randomization approach, sample, data collection and analyses.

Randomization approach

The research team randomly assigned 30 schools, with blocking at the grade level (i.e. schools serving 6-8 grade students and schools serving 9-12 grade students were randomly assigned, separately) to either the intervention group (n=15 schools) or the business-as-usual group (n=15) in February of 2018. The intervention and business-as-usual schools were equivalent on all variables at the time of random assignment. Planning for the intervention began in early 2018, with initial RLT formation in late spring and staff starting Tier 1 trainings during the summer. Full implementation of CF began in treatment schools during the 2018–2019 school year. During the 2018–2019 school year, a CRJ coach worked with treatment school RLT members to implement the CF model at their school (approximately 10–12 hours per week per school). During the 2019–2020 school year, the CRJ coach worked with RLT members to gradually transition implementation responsibilities to the RLT. By spring 2020, RLT members were expected to lead CF implementation at their school.

Implementation and impacts were intended to be analyzed during year 1 (2018–2019) and year 2 (2019–2020) of implementation in treatment schools. Before implementation began, one control school dropped out of the study; however, we use an intent-to-treat approach, and therefore this school remains in our analytic sample. Two additional treatment schools closed during the first year of the study. As it is not possible to measure outcomes in schools that no longer exist, the final analytic sample includes students in the 13 treatment schools that remained open and the 15 control schools. Given the high overall (6.7%) and differential (13.3%) cluster-level attrition in our sample, we reexamined individual-level baseline equivalence between our two analytic samples (2018–2019 school year and 2019–2020 school year analytic samples) on all outcomes assessed in our analysis (See Exhibits 1 and 2). Differences between the treatment and control students ranged from 0.04 to 0.18 standard deviations; thus we applied a statistical adjustment to our analysis.[1] This statistical adjustment involved including a baseline measure of the appropriate outcome in each of our models.

Exhibit 1. Baseline Equivalence Between Intervention and Control Students in the 2018–2019 School Year Analytic Sample

	Intervention	Control	
	M(*SD*) or *%*	*M*(*SD*) or *%*	Hedges' g or Cox's Index
Total Number of Days Absent	13.60 (14.91)	12.88 (16.55)	0.05
Total Number of Unexcused Days Absent	10.64 (13.19)	9.95 (14.91)	0.05
Chronic Absenteeism	26.7	23.0	0.12
Truancy	36.8	31.3	0.15
Percent of Unexcused Absences	0.77 (0.26)	0.75 (0.28)	0.07
Total Number of Days Missed Due to Suspensions	0.22 (0.95)	0.16 (0.79)	0.07
Number of In-School Suspension Incidents	0.02 (0.15)	0.01 (0.12)	0.08
Days Missed due to In-School Suspension Incidents	0.03 (0.23)	0.01 (0.15)	0.11
Number of Out-of-School Suspension Incidents	0.11 (0.45)	0.08 (0.40)	0.07
Days Missed Due to Out-of-School Suspension Incidents	0.20 (0.88)	0.15 (0.76)	0.06
MCAS Math	485.26 (23.06)	489.06 (24.42)	-0.16
MCAS English Language Arts	487.46 (25.49)	492.05 (25.05)	-0.18

Note. We present means and standard deviations for continuous variables and percentages for categorical variables. We present Hedges' g for continuous variables and Cox's Index for dichotomous variables.

[1] What Works Clearinghouse (WWC) 4.1 standards require statistical adjustment for baseline differences greater than .05 but less than .25 standard deviations when attrition from an RCT is differential and high, in order to meet WWC standards with reservations.

Exhibit 2. Baseline Equivalence Between Intervention and Control Students in the 2019–2020 School Year Analytic Sample

	Intervention	Control	
	M(*SD*) or *%*	*M*(*SD*) or *%*	Hedges' g or Cox's Index
Total Number of Days Absent	12.61 (13.22)	11.78 (14.40)	0.06
Total Number of Unexcused Days Absent	9.79 (11.63)	8.90 (12.52)	0.07
Chronic Absenteeism	24.2%	20.9%	0.11
Truancy	35.1%	29.7%	0.15
Percent of Unexcused Absences	0.77 (0.26)	0.75 (0.28)	0.07
Total Number of Days Missed Due to Suspensions	0.21 (0.94)	0.17 (0.81)	0.05
Number of In-School Suspension Incidents	0.02 (0.16)	0.01 (0.13)	0.07
Days Missed Due to In-School Suspension Incidents	0.03 (0.23)	0.01 (0.15)	0.11
Number of Out-of-School Suspension Incidents	0.10 (0.44)	0.08 (0.41)	0.05
Days Missed Due to Out-of-School Suspension Incidents	0.18 (0.87)	0.15 (0.77)	0.04

Note. We present means and standard deviations for continuous variables and percentages for categorical variables. We present Hedges' g for continuous variables and Cox's Index for dichotomous variables.

Sample

Of the 28 schools, 13 served K-8 students, two served middle school students, four served a combination of middle and high school students and nine served exclusively high school students (See Exhibit 3). The students in the study schools were racially and ethnically diverse, with about a third of students identifying as Black and more than half identifying as Hispanic; also, about a third of students identified as English language learners. Racial and ethnic student compositions varied by school, with some schools primarily serving Hispanic students or Black students, whereas others had within-school racial and ethnic diversity.

Exhibit 3. School Grade Range by Treatment Status

	Overall	Intervention	Control
School Type			
High School	9	4	5
Middle School	2	1	1
Middle/High School	4	–	4
K–8	13	8	5

Analytic sample for impact evaluation

The analytic sample for the 2018–2019 school year outcomes analyses included all students who were enrolled in schools that were randomized in the 2017–2018 school year as well as any students who moved into an intervention or control school in the 2018–2019 school year, for a total of 10,650 students (intervention = 4,111, control = 6,539). The analytic sample for the 2019–2020 school year outcomes analyses included all students who were enrolled in schools that were randomized in the 2017–2018 school year as well as any students who moved into an intervention or control school in the 2018–2019 or 2019–2020 school years, for a total of 12,256 students (intervention = 4,513, control = 7,743). Exhibits 4 and 5 present demographic information on the analytic samples.

Exhibit 4. Demographic Information for the 2018–2019 Outcomes Analytic Sample

	Overall	Intervention	Control
Race/Ethnicity**			
Asian	11.5	5.4	15.4
Black	33.9	32.0	35.1
Hispanic	44.5	54.8	38.0
White	9.3	7.1	10.7
Multiracial/Other	1.0	1.0	1.0
Gender (Female)	47.1	46.3	47.6
Individualized Education Plan	23.4	26.2	21.7**
English Language Learner	24.7	32.9	19.9**
Economically Disadvantaged	79.7	84.3	77.1**
Grade**			
6	15.0	20.7	11.6
7	13.0	15.4	11.6
8	13.6	14.4	13.2
9	15.3	12.9	16.8
10	15.0	13.3	16.0
11	12.9	10.6	14.2
12	15.2	12.7	16.6

Note. We present percentages for all categorical variables.

*$p<.05$

**$p<.01$

Exhibit 5. Demographic Information for the 2019–2020 Outcomes Analytic Sample

	Overall	Intervention	Control
Race/Ethnicity**			
Asian	11.3	5.2	14.8
Black	33.8	32.0	34.8
Hispanic	45.1	55.3	39.1
White	9.0	6.8	10.4
Multiracial/Other	1.0	1.0	1.0
Gender (Female)	47.4	46.2	48.1
Individualized Education Plan	23.2	25.9	21.7**
English Language Learner	22.3	30.5	17.5**
Economically Disadvantaged	81.7	86.4	79.0**
Grade**			
6	14.4	19.1	11.7
7	14.9	16.6	13.9
8	12.1	13.7	11.1
9	18.4	17.8	18.7
10	13.7	11.5	15.0
11	13.1	10.6	14.5
12	13.5	10.8	15.1

Note. We present percentages for all categorical variables.
*$p<.05$
**$p<.01$

Sample for implementation evaluation site visits

At the start of the study, a CRJ staff member met with intervention school staff, including administrators, to discuss what motivated them to participate in the study and gauge their awareness of restorative justice and prior experiences with restorative justice. After those conversations, the staff member rated the intervention schools using a three-point scale on two school-level characteristics: their exposure to restorative practices before CF (little/no exposure, some exposure and great exposure) and their perceived openness to restorative practices (somewhat resistant, open, very open). When the study team selected a subset of intervention schools to gather in-depth qualitative data through in-person site visits (hereafter referred to as 'site visit schools'), we used a stratified sampling approach based on a school's rating on those two characteristics because we expected those traits to influence implementation. In addition, we accounted for grade levels (elementary, middle and high school) in selecting schools. In this way, the site visit schools included a diversity of experiences,

such as schools that were very open to restorative practices but had no exposure, and schools that were perceived as somewhat resistant but had some exposure to restorative practices.

Data sources and measures

This mixed-methods study used a combination of data sources to answer the research questions. To answer implementation evaluation research questions, we employed attendance records from the CF Tier 1 trainings, in-person observations of one Tier 1 training, implementation surveys from all intervention schools and interviews and focus groups with staff members and students in the eight site visit schools.

To answer the impact research questions, we relied on extant administrative data from our partner district for the 2017–2018 school year to the 2019–2020 school year. Data included student-level data on suspensions, attendance, academic achievement, and demographic data. We then calculated school-level variables using the student-level data. The remainder of this section provides additional detail on all data sources.

Tier 1 professional development attendance

To understand the extent to which school staff attended Tier 1 trainings (the training most school staff were expected to attend), the study team relied on CRJ attendance data compiled and organized by the district from the 2017-2018 school year to the 2019-2020 school year. CRJ expected that all student-facing staff (e.g., administrators, teaching staff and guidance or support staff) would attend a Tier 1 training and any newly hired teachers in the middle of CF implementation were expected to attend a Tier 1 training. However, some schools focused CF training on particular grades: Some K-8 schools aimed to have middle school teachers trained first, whereas some high schools had teachers serving predominantly grades 9 and 10 students trained first.

Tier 1 observation

In spring 2019, two members of the study team observed one two-day Tier 1 training with sixteen school members across the intervention schools. Study team staff were trained to use the study-specific observation rubric, which used CRJ's training agenda as the basis for the rubric. Observers noted whether the professional development followed the agenda by marking whether those activities occurred as planned (yes/no), the level of engagement of attendees using a closed-ended scale,[2] and whether the training met its key goals (yes/no). Observers also took open-ended notes throughout the training to note participants' body language, types of interactions between the trainer and staff and among staff, and signs of

[2] At the end of Day 1 and Day 2, observers answered the question, 'To what extent did the session actively engage participants?' with one of the following response options: Twenty percent of participants or less were actively engaged most of the day, 40% of participants were actively engaged, 60% were actively engaged, 80% t were actively engaged, and 100%were actively engaged.

engagement (e.g., participants making eye contact with trainers) or disengagement (e.g., staff on their phones). After each day, the observers independently completed the rubric and found high levels of agreement.

Year 1 implementation survey

At the end of the first year of implementation (spring 2019), the study team administered a survey to all school staff who directly support student learning in the intervention schools. The survey included constructs related to fidelity of circle setup, following of circle norms, frequency of circle use, frequency of perceived challenges during circles, perceived implementation challenges, perceived buy-in of stakeholders and perceived changes in school climate. The study team created survey items by reviewing CF materials (e.g., professional development presentations and the CF logic model) and also modified survey items from prior studies of restorative practices use in schools (Augustine et al., 2018; Guckenberg, Hurley, Persson, Fronius & Petrosino, 2016, Jain et al., 2014). Overall, 34% of staff (*n* = 364) across the thirteen schools completed the survey, the majority of whom were teachers (77%). Due to COVID-19-related school closures, the study team was unable to collect a second round of data collection planned for 2020.

Interviews and focus groups

At the end of the first year of implementation (spring 2019), the study team conducted site visits to the sample of eight intervention schools. The team conducted in-person interviews with members of the RLT and student and teacher focus groups to collect in-depth qualitative data about how the CF model was implemented and how it may have varied across schools.

The study team interviewed 46 of 48 RLT members (95%) and conducted one teacher and one student focus group at each school. In total, 50 teachers and 49 students participated in focus groups across the eight site visit schools.

Due to the COVID-19 pandemic, the study team was unable to conduct site visits at the end of the second year of implementation (i.e. spring 2020). Instead, the study team conducted virtual interviews with staff at the eight site visit schools during spring 2021 and asked interviewees to reflect on how their school implemented the CF model during the 2019-2020 school year. The study team invited RLT members and school administrators to participate in interviews if they had been at the school during the 2019-2020 school year. Following district research policy to limit research activities due to COVID-19, the study team did not invite students and teachers to participate in focus groups. In total, 28 of 41 (68%) RLT members and nine of ten (90%) administrators participated in interviews about their second year of CF implementation.[3]

[3] RLT members at each school participated in interviews. At two schools, both a principal and assistant principal were invited to and completed administrator interviews. One school did not have an administrator complete an interview.

School climate survey

School climate outcomes data were to be measured utilizing the school climate survey that our partner district administers to students and teachers each spring. Unfortunately, during spring 2019, the end of the first year of implementation, there were very low response rates and extensive missing data in the surveys. The lack of and incompleteness of the data meant we could not analyze for impacts on school climate. The school climate survey was not given during spring 2020 due to the COVID-19 pandemic.

Student level administrative data

For each student level outcome domain, we included one primary outcome measure supplemented by several secondary outcome measures. Due to the COVID-19 pandemic we were unable to examine academic achievement at the end of the second year of implementation (the state our partner district is located in did not give statewide mathematics and English language arts assessments). For a list of all outcomes used in analyses, see Exhibit 6.

Exhibit 6. Primary and Secondary Outcomes

	Outcome Domain	Primary Outcome Measure	Secondary Outcome Measures
Student Level	Attendance	Total Number of Days Absent in School Year	• Total Number of Unexcused Absences in a School Year • Chronic Absenteeism • Truancy • Percent of Unexcused Absences
	Suspensions	Total Number of Days Missed Due to Suspensions	• Total Number of In-School Suspension Incidents • Total Number of Days Missed due to In-School Suspensions • Total Number of Out-of-School Suspension Incidents • Total Number of Days Missed due to Out-of-School Suspensions
	Achievement	MCAS ELA	
		MCAS Math	
School Level	Suspensions	Overall Suspension Rate	• Difference in Suspension Rates for Black Students Compared to Other Students • Difference in Suspension Rates for Students With an IEP Compared to Students Without an IEP

Attendance

Our primary outcome was the total number of days absent in the school year. Secondary outcomes included total number of unexcused absences, an indicator of chronic absenteeism (whether students missed 10% or more of the school year), an indicator of truancy (whether students had 9 or more unexcused absences), and the percentage of all absences that were unexcused.

Suspensions

Our primary outcome was the total number of days missed due to all suspensions (i.e., both in-school and out-of-school suspensions). Secondary outcomes included total number of in-school suspension incidents, number of days missed due to in-school suspensions, total number of out-of-school suspensions, and number of days missed due to out-of-school suspensions.

School Suspension Rate Disparities

Our primary outcomes were the difference in suspension rates between Black students and all other students and the difference in suspension rates between students with an individualized education plan (IEP) and students without.

Student Achievement

We examined student achievement in both English language arts (ELA) and math using the Massachusetts Comprehensive Assessment System (MCAS), administered in the spring of the 2018–2019 school year. Due to the COVID-19 pandemic, the MCAS was not administered in the 2019–2020 school year. Students' raw scores were standardized across grade level to create a standardized score for each test subject. Because the MCAS is not administered to students in all grades in high school, these data were restricted to students in Grades 6–8.

Covariates

We used a standard set of covariates in all analyses. Demographic variables used as covariates included race/ethnicity (Hispanic [reference], Black, non-Hispanic White, Asian/multiracial/other), gender (male = 0, female = 1), whether students had an individualized education program in place (no = 0, yes = 1), whether students were considered economically disadvantaged (no = 0, yes = 1), whether students were English language learners (no = 0, yes = 1), and school type (0 = middle school, 1 = high school). For race/ethnicity, Asian and multiracial/other categories were collapsed due to small cell sizes. For school-level analyses, we included school-level percentages of all demographic variables as covariates.

Data analysis

Data were analyzed to answer all implementation evaluation and impact evaluation research questions. The remainder of this section provides additional detail on all data analysis.

Implementation evaluation data analysis

To understand the fidelity of the Tier 1 trainings, the study team examined the attendance of intervention schools at those trainings and analyzed the observation data from them. For the attendance records, the study team calculated the percentage of school staff who attended a Tier 1 training from spring 2018 through 2019-2020 with the denominator comprising all staff[4] who worked in the intervention schools at some point during implementation[5] (2018-2019 through 2019-2020). For the observation data, the study team descriptively analyzed the observation rubric data by reporting counts and reviewed the open-ended notes to identify common themes.

To analyze the implementation survey data, the study team created Rasch scale scores (Andrich, 1978; Rasch, 1980; Wright & Masters, 1982; Wright & Stone, 1979) using Winsteps, a Rasch analysis software program (Linacre, 2015). The study team reviewed the psychometric output for item fit and internal consistency (see the Appendix for further psychometric information and for the crosswalk of individual survey items to scaled constructs). The study team converted the scale scores back into their original metric (i.e. the Likert scale), allowing for more meaningful interpretation of the scores. The study team conducted descriptive analyses on the converted scale scores as well as on individual survey items.

To analyze the qualitative interview and focus group data, analysts coded transcripts using a set of a priori codes based on protocols that correspond to core components of the CF model and to the research questions. To ensure reliability across coders, members of the study team each coded one transcript of each type of protocol in common and subsequently met to compare coding and mediate any areas of disagreement. The team also created additional codes after reviewing transcripts to identify emergent themes in the data. We employed a cross-case design and pattern-matching technique to identify patterns of practice (Yin, 2009). Our findings show themes present in more than one school community group (e.g., RLT members, school administrators, teacher focus groups and student focus groups) and across four or more schools. Also, the study team discovered a few themes that the study team linked only to RLT members. In these cases, the study team reported findings that arose across four or more schools and

[4] District roster data classified staff into the following categories: teacher, administrator, support and unknown.
[5] Because the CF model aimed to have all staff trained by the end of the intervention, we did not analyze attendance rates by school year.

represented the sentiments of at least two RLT members at each school. Lastly, the study team reported teacher and student focus group findings that arose across four or more schools.

Impact evaluation data analysis

To examine student-level outcomes, we fit a series of linear regressions. We regressed each outcome variable onto the indicator of intervention status (0 = control, 1 = intervention), the appropriate outcome variable measured at baseline (the 2017–2018 school year), and baseline demographic variables. We estimated cluster robust standard errors.

We also conducted separate analyses to examine whether the intervention impact varied for students who were in middle or high school. Each model described above was subsequently fit using two models: (1) a subset of students who were in middle school, and (2) a subset of students who were in high school.

To examine school-level outcomes, we conducted a difference-in-differences analysis. We regressed each outcome variable onto our indicator of intervention status (0 = control, 1 = intervention), school year indicator (pre = 0, post = 1), an interaction between the intervention and school year indicators (intervention indicator X school year indicator), and school demographic covariates. We estimated robust standard errors.

Findings

Overall, we found that CF was generally implemented as intended. We found positive impacts on attendance during the second year of implementation but no impact on exclusionary discipline use, disparities in exclusionary discipline use, or academic achievement. In the remainder of this section, we provide additional detail on our findings, first for the implementation evaluation, and then for the impact evaluation.

Implementation evaluation findings

Circle Forward was, for the most part, implemented as intended, though there was disagreement about the extent to which CF aligned with school policies. Staff and students felt positively about the implementation of CF and its effects on relationships and school climate, but noted a few challenges including time constraints and student participation and behavior. In the following sections, we provide additional details about each of the findings.

To what extent is CF implemented as intended, and how does implementation vary across schools?

To understand if the CF Tier 1 training was implemented as intended and the extent to which school staff attended Tier 1 trainings the study team used observation data, attendance records, qualitative data (interviews and focus groups) and survey data. The study team also used qualitative and survey data to examine implementation in schools, including the extent to which staff engaged in circles, the types of circles that they engaged in and whether school staff followed circle rituals and norms.

The two research team observers found that the CRJ trainers followed the agenda as intended during the observed Tier 1 training. The observers agreed that the observed Tier 1 training met all of its goals: Participants (a) understood the key values, concepts, and practices of a whole school approach to restorative practices; (b) experienced the circle process and (c) developed skills to plan and facilitate talking circles. During Day 1, CRJ trainers covered all eight key topics[6] and completed all seventeen planned activities (including the whole group discussing the values that guide participants as educators, as well as watching a Tier 1 community-building circle video). During Day 2, trainers covered eight of nine topics[7] and thirteen of fourteen activities (such as small group work using the circle planning sheet, and participants going in a circle and sharing how they will incorporate CF). Across activities, participants were highly engaged with 100% of participants engaged for most of Day 1, and 80% of participants engaged in Day 2. Observers said participants exhibited behavior that indicated active listening and interest, such as looking at the CRJ trainer, sharing thoughtful answers (rather than one-word responses), and nodding or snapping in agreement. The CRJ trainers exhibited strong facilitation skills by answering participant questions, repeating the responses of other participants, and during small group activities (e.g., pair-shares), they joined groups to listen to their discussions. In interviews and focus groups, RLT members, teachers and students similarly reported that CF professional development helped them understand the background, key concepts and importance of restorative justice.

Although all staff who directly interact with students (e.g., teachers, paraprofessionals, counselors, administrators) were expected to attend the two-day Tier 1 training, attendance rates varied by school and were generally low based on CRJ records. By spring 2020, the percentage of school staff trained ranged from 11% to 73%, with only four of thirteen schools having a staff attendance rate of at least 50%. To understand why staff may not have attended

[6] The eight topics covered on Day 1: (1) welcome, (2) introduction to participants, (3) values, (4) agreements, (5) restorative justice, (6) self-care and balance, (7) restorative discipline and restorative questions and (8) closing.
[7] The nine topics to be covered on Day 2 were (1) welcome, (2) reflections on yesterday's session, (3) seven core assumptions, (4) share a personal item, (5) setup of circles, (6) circle practice and reflection, (7) circle planning, (8) circle problem solving and (9) check out and closing. In the observed professional development, CRJ trainers skipped topic 4.

professional development, the study team also examined survey and interview data. In the survey, the study team asked respondents to self-report their professional development attendance and included an item regarding why a respondent had not attended any CF training (including Tier 1 circle training). The Tier 1 circle attendance rate was high among survey respondents, with 77% (*n* = 281) reporting that they had attended the Tier 1 training. Sixteen percent (*n* = 60) had never attended *any* professional development. The most common reasons respondents gave for not attending included being unaware of professional development offerings and seeing no offerings that fit their schedules. Also, some teacher focus group participants felt like they heard about offerings at the last minute and did not have enough time to plan to attend the trainings. Looking closer at the schools with the lowest attendance rates, the study team found one school already had trained most staff in restorative justice practices, though not using the CF program, prior to the study period and another school was onboarding several new staff at the time of professional development.

When examining implementation of CF at schools, the study team found that the majority of survey respondents reported that they had either led or participated in circles (91%; *n* = 332). Of those who had led circles, most survey respondents (94%; *n* = 188) reported that they had attended the Tier 1 training. In addition, when examining the percentage of respondents who reported leading circles by school, the study team found that the school with the lowest percentage of survey respondents (40%; *n* = 25) leading circles also had the lowest staff attendance (11%) at Tier 1 trainings according to district attendance records. Of the survey respondents who had never led or participated in circles (nine percent; *n* = 32), the most common reason selected[8] was that they had no opportunities to participate in circles.

In the survey, interview and focus group data, the most frequently implemented circles reported in Year 1 were advisory circles with students (for relationship building), pedagogy circles (as an instructional tool), problem-solving circles with students (to address minor behavioral issues), and classroom norm circles (to set norms and address any violations). RLT members and teachers alike reported implementing advisory circles with students and pedagogy circles, but teachers reported implementing problem-solving circles less often with students. In Year 2, RLT members frequently mentioned implementing advisory circles with students, and RLT members and school administrators frequently mentioned advisory circles with staff.

Survey respondents who had experience with circles responded to two fidelity constructs to examine the extent to which core circle components were implemented. The first construct, fidelity of circle setup, measured whether the expected circle rituals were happening (e.g.,

[8] Other options that survey respondents could have selected were 'time constraints,' 'I don't believe circles can work,' 'I have not received any or enough training' and 'Other.' Respondents could select more than one reason.

using the centerpiece, talking piece). The second construct, following of circle norms, measured whether respondents perceived that circle participants were behaving as expected (e.g., participants wait to speak until they are holding the talking piece). Most respondents reported a high level of fidelity of circle setup and following of circle norms; 87% (*n* = 283) and 74% of responses fell in the *most of the time* or *always* categories, respectively (*n* = 242) (see Exhibit 7). Schools with high fidelity in circle setup tended to have a similarly high fidelity in following circle norms.

Exhibit 7. Distribution of responses for the scaled constructs: Fidelity of circle setup and following of circle norms

	Fidelity of circle setup	Following of circle norms
Always	57%	29%
Most of the time	30%	45%
Sometimes	11%	24%
Rarely/never	2%	2%

Note. Surveyed staff (*n* = 326)

What facilitators and challenges to program implementation arose? How did schools resolve challenges?

Interview and focus group participants noted that lack of time for circle implementation, student participation, and administrators' and teachers' hesitance to implement circles were barriers to circle implementation. Stakeholders at each school also described their unique approaches to addressing these challenges.

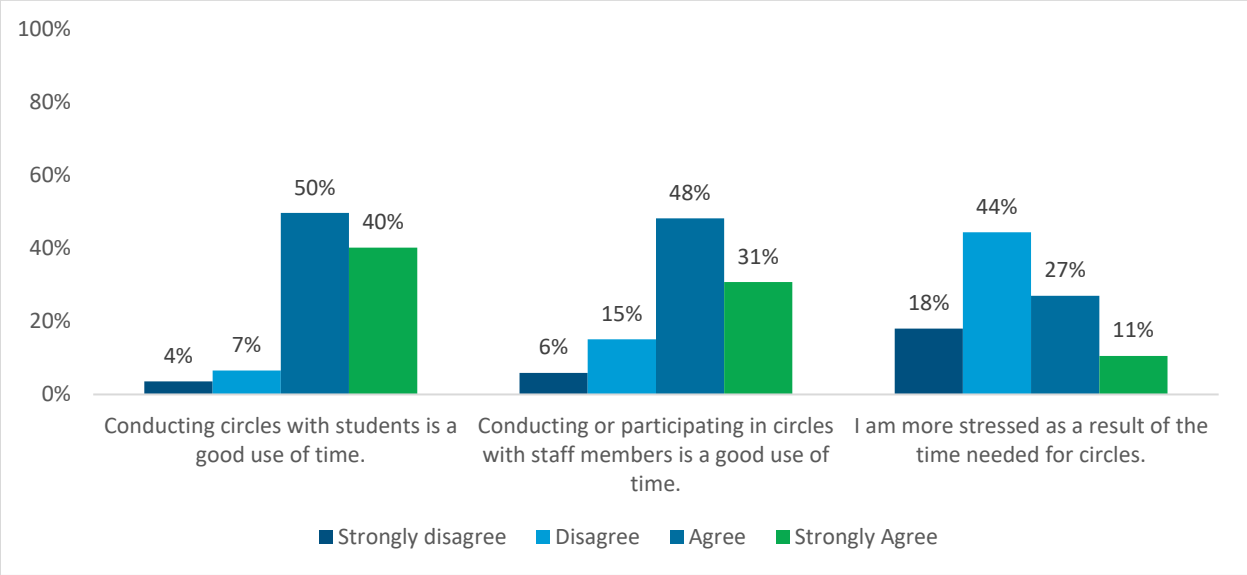
Logistically, finding time for circles was a common barrier to implementation. Forty-nine percent (*n* = 157) of survey respondents who led or participated in circles reported that 'time constraints' were a *major* or *moderate* problem. During Year 1 and Year 2 interviews, RLT members also said they did not have enough time to complete the circle process in their classes. Interview participants explained that circles invite each student to share their thoughts with the rest of the group. However, this became a lengthy process when circles were used with the entire class. Furthermore, a few RLT members noted that circles used time that was needed for other school priorities, such as school climate team meetings, the Comprehensive Behavioral Health Model and their academic course requirements. One RLT member said:

> *I couldn't do circle as often as I would like as well because I had content that I had to teach too. If I have 45 minutes to an hour, it's time-consuming and I couldn't spend a lot*

of time on it if I were to get my curriculum taught … Therefore, I might have aimed to do a circle with every class once a month. How effective that is I don't know.

In another set of survey items focused on the use of circles and time (see Exhibit 8), a sizable minority (38%; *n* = 121) of respondents *agreed* or *strongly agreed* that they 'felt more stressed as a result of time needed for circles.' A large and consistent majority of survey respondents, however, reported that they *agreed* or *strongly agreed* that circles with students were a good use of time (90%; *n* = 302), and in four schools 100% of responding staff *agreed* or *strongly agreed* with this statement. Overall, a sizable majority of respondents also *agreed* or *strongly agreed* that circles with staff were a good use of time (79%; *n* = 267). There was greater variation among schools to this item, however; the school with the lowest percentage of respondents agreeing or strongly agreeing was 53%, and one school had 100% of staff agreeing. Survey respondents also reported on average, spending less than two hours per week planning (81%; *n* = 166) or participating (73%; *n* = 236) in circles (see Exhibit 9).

Exhibit 8. Utility of time spent conducting circles



Note. Surveyed staff (*n* = 322–338).

Exhibit 9. Time spent planning or participating in circles

	Time spent planning for circles	Time spent leading or participating in circles
Less than 2 hours per week	81%	73%
2-4 hours per week	6%	10%
More than 5 hours per week	1%	2%
None	12%	15%

Note. The left column (*n* = 206) only includes respondents who led circles, and the right column (*n* = 322) includes respondents who led or participated in circles

Year 1 RLT members, teachers and students also reported student behavior sometimes posed a challenge to circle implementation. In Year 1 interviews and focus groups, RLT members, teachers, and students described instances where students became 'off-task' or did not take circles seriously. However, respondents said these disruptions were isolated incidents and did not pose a widespread challenge to circle implementation. Furthermore, RLT members said they addressed misbehavior by using 'trial and error' with different strategies of presenting circles to students and worked to increase student buy-in for circles. One RLT member described how they worked through student misbehavior:

> *Not forcing any participation, letting people join and leave the circle when they want. I think also having students who struggle with behavioral issues sit near the circle leader, have them be the keeper, give them jobs and things like that.*

Year 1 RLT members, teachers and students also noted difficulties with the same students participating and needing to engage students who were shyer than their peers. Student focus group participants said 'it's the same kids talking' or 'sometimes it's kind of hard because some people keep on talking.' Similarly, one RLT member said they struggled to find a balance with students who would talk for 'five, six, seven minutes' and others who would not participate. One student focus group member described this barrier:

> *They want to participate, but when they do, they say two or three words. They don't really expand on what they're saying, and you want to get a good understanding about what they're going to say and how they feel about the certain topic ... They usually speak really low. Then when you tell them to say it again, they're, like, 'Oh, no, never mind.'*

RLT members and teachers addressed this challenge by using alternative methods to get more students to participate. For example, one RLT member used a ball of yarn as the talking piece and had students create a 'web' of yarn by throwing the yarn to the next speaker. The RLT member said this method was 'more engaging' and indicated that students were eager to get

the yarn talking piece. One teacher also described using different question prompts that encouraged better student discussion.

In the survey, staff were asked a series of questions about possible challenges (e.g., none of the students share, student misbehavior derails the circle) they may have experienced when conducting circles with students. The survey results showed that the majority of respondents perceived that issues with student circles occurred *sometimes* or *rarely/never* (see Exhibit 10).

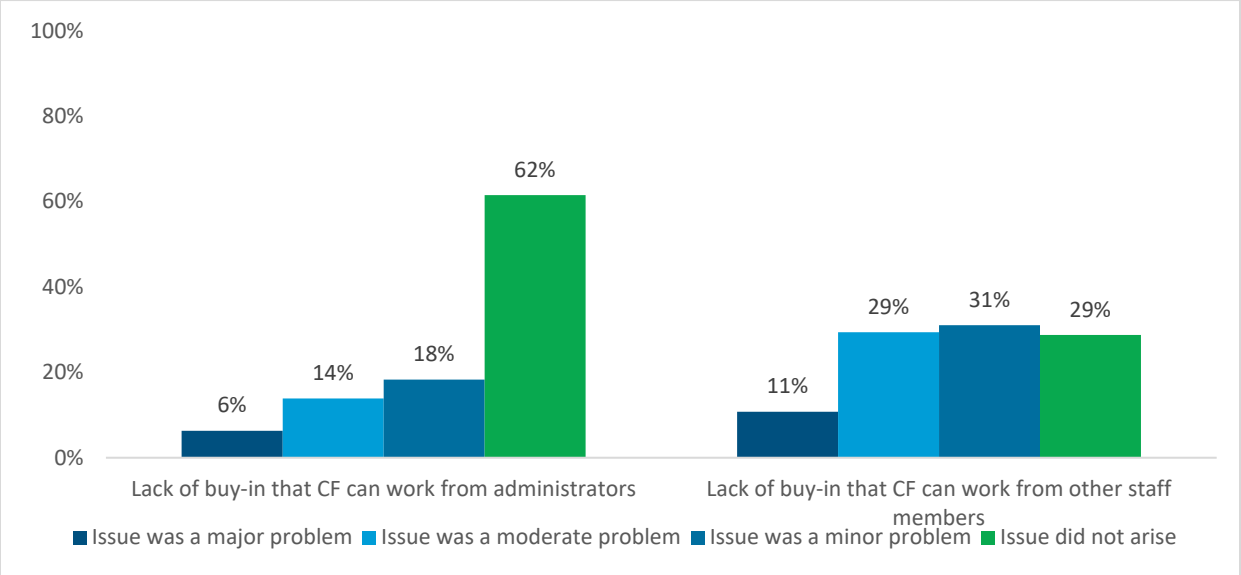
Exhibit 10. Distribution of responses for the scaled construct: Frequency of perceived challenges when leading circles with students

Always	0%
Most of the time	2%
Sometimes	56%
Rarely/never	42%

Note. Surveyed staff (*n* = 201). Only staff who led circles with students responded to the set of items.

Finally, school community members described how a handful of staff expressed hesitance in using circles. In the survey, 20% (*n* = 64) of respondents viewed buy-in from administrators as a *major* or *moderate* problem, and 40% (*n* = 123) of respondents felt that buy-in from other staff was a problem (see Exhibit 11).

Exhibit 11. Extent to which buy-in from staff and administrators was perceived as a problem



Note. Surveyed staff (*n* = 306–317).

During Year 1 interviews, RLT members reported that a small number of administrators and teachers were reluctant to adopt circles. Schools with a perceived lack of buy-in experienced some initial resistance to restorative justice or some or no exposure to it before the study period began. RLT members said they perceived a lack of buy-in from school administrators when they did not communicate their support of circles to the wider school community or create dedicated time in the school schedule for circle implementation or professional development. Also, RLT members said some teachers expressed reluctance toward circles because they 'were stuck in their thinking' about what restorative justice meant and thought of it as just one of many initiatives brought to the school. RLT members said this hesitancy arose mainly when circles were first introduced to the school, and that the school community subsequently was able to move past these challenges.

In Year 2 interviews, though, RLT members and school administrators still described barriers related to consistent buy-in with a handful of teachers. Some RLT members and administrators said these teachers hesitated to fully adopt circles due to the 'personalities of the teachers and their philosophies.' Another RLT member acknowledged that 'you can't get 100% buy-in. It's just not something that you're going get in the education field.' A few RLT members also said some of the teachers who fully bought in were new to the school and may not have received training on restorative practices. One RLT member described their thoughts on school buy-in:

> *I think circles are effective for any classroom, but as a school it has to be a whole-school buy-in. If I'm in first grade, I know I'm [going to] have a circle, and when I'm in eighth grade I know that I'm still [going to] have that same kind of circle. It's really difficult when the whole entire school is not buying in or [says] "I don't have time."*

RLT members mentioned several strategies that they used to increase schoolwide support for the CF model. They described the importance of providing school staff with resources and training in restorative practices so that they understand the key principles of circles. Also, RLT members said they offered concrete examples of different types of circles so that teachers could integrate the CF model more easily into their classrooms. RLT members also said they received support from teachers and school administrators, such as helping implement circles with students, addressing challenges together and exhibiting an openness to making circles a priority. Finally, some RLT members worked with school administrators to dedicate space in the school schedule to implement circles. One RLT member noted the importance of school leader support in creating time for circles:

> *I think that [school administrators] created space in the schedule once a week for us to do circles. That was a huge, huge piece of being able to ground practices. That was really*

monumental … We weren't just talking about restorative justice. We were actually creating space in the weekly schedule for kids to participate.

In what ways does implementing the CF model lead to changes in practice for all those trained?

In Year 1 and Year 2 interviews, RLT members described improved relationships among staff and between staff and students. Interviewees said they developed closer relationships with and more holistic views of their colleagues and students. Year 1 teacher focus group participants also described improved relationships with their students. Also, RLT members in Year 1 and Year 2, as well as Year 2 school administrators, reported having a more reflective or empathetic mindset after participating in circles. Interview participants said circles helped them understand what others were going through and allowed them to approach relationship building and conflict resolution from a different perspective. One RLT member shared their experience:

> *It's helped me really think through the way I show up for children every single day. We're human, and there are moments where I don't show up perfectly … Being in any circle with staff, I've learned some things that I didn't know before, which makes me have a little more empathy for my coworkers. Also, this year in general, through everything I've learned, particularly through the [restorative justice] leadership team, I approach people with a little more grace and a little more humility and understanding. There's always a story that I don't know.*

In the survey, staff were asked a series of questions about the extent to which CF had succeeded in changing their practice and improving aspects of school climate (e.g., gaining skills in alternative responses to student behavior, improving relationships among staff members, improving relationships between staff and students). Echoing the interview responses, survey respondents said they believed CF led to improvements in school climate as it pertained to staff and student behavior and relationships. Fifty-five percent (*n* = 166) of surveyed staff felt that CF improved these measures to a *moderate* or *great extent* (see Exhibit 12).

Exhibit 12. Distribution of responses for the scaled construct: Perceived improvement in school climate

	Perceived improvement in school climate
Great extent	11%
Moderate extent	44%
Small extent	32%
Not at all	12%

Note. Surveyed staff (*n* = 299)

In reflecting on changes in students, 61% (*n* = 151) of survey respondents perceived that CF was successful to a moderate or great extent in increasing student voice. Similarly, Year 1 RLT members and teachers said students developed better relationships with their classmates, and their use of circles increased student voice (or students' comfort sharing their thoughts with others). RLT members and teachers said circles helped the class become like a 'community' or 'family' because students got to know each other on a deeper level. Also, RLT members and teachers said circles helped students become more comfortable expressing their opinions, experiences, and areas where they need support. One RLT member said that 'many times adults are letting them down because they're not listening to them. Just giving them an opportunity to actually share … Their parents may not be listening. Their parents may be too busy. I think it's just a great opportunity for everybody to actually speak.' Year 2 RLT members also reported an increase in student voice during their interviews. One Year 2 RLT member said:

> *At the beginning, [students] could be very timid or shy or insecure or scared. By having these circles and keep practicing it, it's like, "If there's an issue, how do people go about resolving it and hearing other people's perspectives?" It's made them feel more comfortable and also being receptive. You know what I mean? Being able to say, "Wow. Now I get it." It makes them feel more comfortable, and it's relationship building.*

During Year 1 focus groups, students also reported seeing such changes in themselves and their peers. Student focus group participants said circles helped them feel comfortable talking in front of groups of people and 'be more open with their feelings.' Students said they could have 'deep conversations' with their peers and learn about their background experiences, mental health, or whether they were going through challenging times. One student focus group member described the impact of circles:

> *Circle has helped me be more open with myself, and with other students. When I first came to this school, I was a bit more reserved, and I didn't really talk to other students that weren't in my close circle. Now, because of circles, I've come out more, and I've been talking to kids in other grades more, as well as other kids in my class.*

How is CF integrated into and coordinated with preexisting programs and aligned to school policies and practices to become routine use?

To understand how participants integrated circle implementation into their existing school routines, the study team used RLT interviews and interviews with school administrators to understand the roles of the CRJ coach and RLT team members as well as the extent to which circles were perceived as aligning to school policies and practices. The study team also used survey data and teacher focus group data to understand the extent to which perceptions were aligned across school community members.

During the 2018-2019 school year, a CRJ coach was assigned to each school implementing the CF model. Year 1 RLT members said they held biweekly team meetings, and the CRJ coach always or almost always attended them. RLT members said the CRJ coach helped plan their team meetings and professional development, acted as a resource to other staff, and helped lead circles and professional development. As with the interview findings, survey respondents who had met with the CRJ coach said the CRJ coach most frequently led professional development workshops (65%; *n* = 175) and helped plan circles with staff (59%; *n* = 159).

As for the RLT work, team members said they planned professional development, acted as a resource to school staff, and implemented circles. Overall, RLT members said the CRJ coach helped move circle implementation forward and keep their team on track. One Year 1 RLT member said,

> *We would not exist without [the CRJ coach]. [The coach] teaches the team. [They] did so much of the planning, planning the meetings, helping to plan for PD sessions. Often what would happen is we would come up with a general idea of what we wanted to see happen for PD, and then because everyone else in the building has their own other full-time job, [they] would flesh stuff out. [The coach] did a lot of legwork type stuff, but also just educated us about things.*

During the first half of the 2019-2020 school year, CRJ coaches then worked with RLT members to build their capacity and gradually transition responsibilities to the RLT. When asked how their CRJ coach supported the RLT during that school year, RLT members said the coach acted mainly as a resource to school staff. CRJ coaches passed on their CF work in schools in December 2019, expecting RLT members to take full responsibility for implementing and planning the CF model for the rest of the 2019-2020 school year. RLT members said they supported circle implementation during that school year by acting as a resource to school staff and actively implementing circles.

Because the pandemic delayed data collection from spring 2020 to spring 2021, the study team asked Year 2 RLT and school administrator interview participants about the use of circles after CRJ coaches had ended their support of CF at schools. In Year 2 interviews, RLT members said they felt moderately prepared to implement CF after their CRJ coach had left their school. However, RLT members perceived that CF implementation decreased due to the transition to remote learning and a lack of training for new staff. RLT members said 'so many things were changing' during the pandemic, and that circles had to take a 'back seat' to other priorities. Some RLT members also described significant turnover of school staff who had led CF implementation at their school. One RLT member said their professional development meetings

essentially came to a halt during the pandemic, so they were unable to plan CF implementation. One RLT member described some challenges:

> *We have advisory, which is basically [restorative justice] time. I think there's new staff that really have never seen each other, who have never done Tier 1 training. It's hard for them to say, "Hey, I'm just [going to] start doing circles, and I've never done this before." I think it's changed a little bit for sure in this aspect. I think anybody who's done the circles before, I think they're doing a pretty solid job.*

Also, some RLT members described challenges with implementing circles over Zoom, including difficulty leading circles without a physical talking piece and engaging students who were burnt out from hours of screen time. Despite these challenges, other RLT members reported continued CF implementation and discussion of relevant topics during circles (e.g., COVID-19, grief, race, social justice). These RLT members felt students and school staff built stronger relationships during a particularly difficult time. For example, one RLT member described circles on COVID-19, self-care and processing one's feelings. Another RLT member said:

> *We were running circles on Zoom throughout COVID, and it actually was one of the most unifying pieces of our remote learning because you would all come together as a group to talk and say how things are going, and that type of thing.*

Overall, members of Years 1 and 2 RLTs and Year 2 school administrators said consistently that circles aligned with their district or school's policies regarding discipline and school climate. Interview participants said their school's focus on equity and goal of reducing the school-to-prison pipeline were examples of how their school's policies were aligned closely with circles. In addition, interviewees mentioned their school's commitment to reducing the number of suspensions. One RLT member said:

> *[The district] is looking to move away from suspensions and lower the number of suspensions for students who are missing class time. This new approach, Circle Forward, and restorative justice approaches in general have helped staff members find ways to not let students "off the hook" but find ways to help them build agency within the classroom, help them [avoid] removals from the classroom.*

However, this was not a consistent theme in either the Year 1 teacher focus groups or the survey results. Teacher focus group participants said their school did not have distinct policies regarding discipline and school climate. Thus, it was unclear how circles aligned to their school's existing policies. In addition, 49% of survey respondents (n = 155) reported 'It is unclear how CF fits into the school's discipline policy' as a *major* or *moderate* problem. Therefore, teachers may have a different perspective than RLT members and school administrators. Teachers may not

have participated in as much training or circle implementation as RLT members and school administrators and may not perceive a connection between the principles of restorative practices and their school's policies.

Impact evaluation findings

The impact evaluation aimed to answer four primary research questions. First, we intended to examine whether students in schools implementing CF reported more positive school climate and feelings of safety than students in schools not implementing CF. However, due to low response rates and lack of administration due to the COVID-19 pandemic we were unable to use the school climate survey to construct these measures. Additionally, we set out to examine whether schools implementing CF saw improvements in attendance, reductions in the use of exclusionary discipline, reductions in disparities in exclusionary discipline use, and improvement in academic achievement during the 2018-2019 and 2019-2020 school years, but were unable to measure academic achievement during the 2019-2020 school year due to the COVID-19 pandemic.

Overall, the impact evaluation found positive impacts on attendance during the second year of implementation, but no impact on exclusionary discipline use, disparities in exclusionary discipline use, or academic achievement. In the remainder of this section we describe our findings in more depth.

Do schools implementing CF see improvements in attendance?

We found fewer days absent and fewer unexcused days absent for students in schools implementing the CF intervention compared to students in control schools (Exhibit 13). We found no differences in attendance between students who attended a school implementing the CF intervention and those students who attended control schools in the 2018–2019 school year.

Exhibit 13. Circle Forward Overall Impact on Middle and High School Students' School-Day Attendance in the 2019–2020 School Year

	Estimate	*SE*	*p* value	95% CI	
Number of Total Days Absent	-1.61*	0.63	0.01	[-2.86, -0.37]	--
Number of Unexcused Days Absent	-1.62*	0.70	0.02	[-3.00, -0.25]	--
Chronic Absenteeism	-0.15	0.12	0.24	[-0.39, 0.10]	OR = 0.86
Truancy	-0.29	0.16	0.07	[-0.61, 0.02]	OR = 0.75
Percent Unexcused Absences	0.02	0.03	0.63	[-0.05, 0.08]	--

Note. We present unstandardized coefficients, cluster-robust standard errors, *p* values, and 95% confidence intervals. In our last column, we present odds ratios (OR) as appropriate for analyses with binary outcomes. Covariates included baseline attendance, race/ethnicity, gender, IEP status, economically disadvantaged status, English language learner status, and school type.

*$p<.05$

Do schools implementing CF experience reductions in use of suspensions?

We found no differences in suspensions between students who attended a school implementing the CF intervention and those students who attended a control school in either the 2018–2019 or 2019–2020 school years (Exhibits 14 and 15). Subgroup analyses by grade range also found no statistically significant effects.

Exhibit 14. Circle Forward Overall Impact on Middle and High School Students' Suspensions in the 2018–2019 School Year

	Estimate	*SE*	*p* value	95% CI
Total Days Missed Due to Suspensions	0.05	0.06	0.41	[-0.07, 0.17]
In-School Suspension Incidents	-0.002	0.01	0.78	[-0.01, 0.01]
Days Missed Due to In-School Suspensions	-0.01	0.01	0.52	[-0.02, 0.01]
Out-of-School Suspension Incidents	0.03	0.02	0.15	[-0.01, 0.08]
Days Missed Due to Out-of-School Suspensions	0.06	0.06	0.32	[-0.06, 0.17]

Note. We present unstandardized coefficients, cluster-robust standard errors, p values, and 95% confidence intervals. Covariates included prior-year suspensions, race/ethnicity, gender, IEP status, economically disadvantaged status, English language learner status, and school type.

Exhibit 15. Circle Forward Overall Impact on Middle and High School Students' Suspensions in the 2019–2020 School Year

	Estimate	*SE*	*p* value	95% CI
Total Days Missed Due to Suspensions	0.001	0.03	0.96	[-0.06, 0.06]
In-School Suspension Incidents	-0.01	0.003	0.13	[-0.01, 0.00]
Days Missed Due to In-School Suspensions	-0.004	0.01	0.38	[-0.01, 0.01]
Out-of-School Suspension Incidents	0.01	0.01	0.42	[-0.02, 0.04]
Days Missed Due to Out-of-School Suspensions	0.01	0.03	0.82	[-0.05, 0.07]

Note. We present unstandardized coefficients, cluster-robust standard errors, p values, and 95% confidence intervals. Covariates included baseline suspensions, race/ethnicity, gender, IEP status, economically disadvantaged status, English language learner status, and school type.

Are there reductions in disparity in use of suspensions by race?

We found no impact on disparities in school-level suspension rates by race or IEP status in either the 2018–2019 or 2019–2020 school years (Exhibits 16 and 17). Subgroup analyses by grade range also found no statistically significant effects.

Exhibit 16. Circle Forward Impact on School Suspension Rates in the 2018–2019 School Year

	Estimate	*SE*	*p* value	95% CI
Overall Suspension Rates	-0.03	0.02	0.16	[-0.07, 0.01]
Disparity in Suspension Rates for Black Students	-0.03	0.03	0.27	[-0.08, 0.02]
Disparity in Suspension Rates for Students with an IEP	0.01	0.04	0.85	[-0.07, 0.08]

Note. We present unstandardized coefficients, robust standard errors, *p* values, and 95% confidence intervals. Covariates included percentages of female students, students with an IEP, students identified as ELL, students identified as economically disadvantaged, Black students, Hispanic students, and White students.

Exhibit 17. Circle Forward Impact on School Suspension Rates in the 2019–2020 School Year

	Estimate	*SE*	*p* value	95% CI
Overall Suspension Rates	0.00	0.02	0.96	[-0.04, 0.04]
Disparity in Suspension Rates for Black Students	0.01	0.03	0.62	[-0.04, 0.06]
Disparity in Suspension Rates for Students with an IEP	0.06	0.04	0.18	[-0.03, 0.13]

Note. We present unstandardized coefficients, robust standard errors, *p* values, and 95% confidence intervals. Covariates included percentages of female students, students with an IEP, students identified as ELL, students identified as economically disadvantaged, Black students, Hispanic students, and White students.

Do schools implementing CF see improvements in student academic achievement?

We found no differences in assessment scores on ELA or mathematics between students who attended a school implementing the CF intervention and those students who attended a control school in the 2018–2019 school year (Exhibit 18).

Exhibit 18. Circle Forward Impact on Middle School Students' Standardized Assessment Scores in English Language Arts and Math

	Estimate	*SE*	*p* value	95% CI
MCAS ELA	0.03	0.04	0.47	[-0.05, 0.12]
MCAS Math	-0.58	0.34	0.09	[-1.25, 0.09]

Note. We present unstandardized coefficients, cluster-robust standard errors, *p* values, and 95% confidence intervals. Covariates included prior-year standardized assessment scores, race/ethnicity, gender, IEP status, economically disadvantaged status, and English language learner status. ELA = English language arts.

Limitations

This study provides important evidence on the implementation and impacts of a promising whole-school approach to restorative practices. Staff in schools implementing Circle Forward reported having a more reflective and empathetic mindset toward students after participating in circles. The large majority of staff also reported believing that Circle Forward led to improvements in school climate, particularly as it related to relationships between staff and students and among students. Students also reported that circles helped them feel more comfortable sharing their thoughts and feelings and connect better with their classmates. Compared to control schools, schools that implemented CF had higher rates of attendance after the second year of implementation. The study has a number of limitations, however, which we will explore further in the remainder of this section.

Implementation evaluation limitations

First, the study team had to scale back data collection activities because of the COVID-19 pandemic. The team could not conduct the implementation survey of teachers and staff, or the teacher and student focus groups, at the end of Year 2. The study team therefore could not report on broad perceptions of implementation and effectiveness after two years of CF implementation (analyses it would have conducted on Year 2 survey responses). The study team also could not report more in depth on student and teacher perceptions of implementation and effectiveness after two years of CF implementation (analyses it would have conducted on Year 2 student and teacher focus groups). Because whole-school interventions like the CF model require time to permeate school culture, Year 2 survey and focus group data would have provided valuable information on successes and challenges and perceived impacts during the second year of implementation.

Second, response rates to the implementation survey were low (34%) and do not represent the opinions of all teachers, administrators, and other support staff in intervention schools. Because the majority of respondents had attended the Tier 1 training, and our review of Tier 1 attendance records found most schools had attendance rates below 50%, Tier 1–trained school staff were overrepresented in the survey sample.

Third, although the study team was able to collect Year 2 RLT interview data, the response rate was lower than that of Year 1 RLT interviewees. The makeup of the RLT staff varied from Year 1 to Year 2, but the study team also understood that the timing of the Year 2 interviews in spring 2021 coincided with a particularly hectic period during the pandemic. In addition, the study team had to transition from in-person interviews to virtual ones, which were more difficult to schedule.

Last, the intervention took place in one large, urban district and therefore is not generalizable to every district. Because the study team had targeted schools that had received Title I funding and had a high percentage of children from low-income families, the findings may be useful for considering implementation in high-need schools and districts that typically have inadequate resources.

Impact evaluation limitations

First, because of the size of the research grant and the intensity of the intervention, the scope of the study was limited to 30 schools, which provided lower power to detect impacts than would be possible with a larger sample. The diversity of age ranges served means that analyses are combining changes in outcomes among high school students with those of middle school students on measures where students are starting in developmentally different places. This heterogeneity also reduces power. Further, two intervention schools closed during the first year of implementation. This attrition further reduces the study's power. It is possible that, given a larger sample, we would have been able to detect impacts on exclusionary discipline or academics that were not statistically significant within our sample. Additionally, a larger sample size would have allowed for greater power for subgroup analyses, allowing us to better examine whether CF was more or less successful in impacting examined outcomes for particular subgroups of students (e.g., middle school students, high school students, Black students, and students with an IEP).

Second, the closure of two intervention schools resulted in intervention and control groups that were more different on average than they had been at the onset of the study. Although we established baseline equivalence between intervention and control schools on all of our outcomes during the 2017–2018 school year, the closure of two intervention schools resulted in the remainder of intervention schools differing from control schools. No outcomes differed by more than .25 standard deviations, however, and thus we included baseline measures of all outcomes in models to statistically control for differences between the two groups of schools at baseline. This statistical correction allows our study to meet WWC standards with reservations, but we were not able to meet WWC standards without reservations—a designation reserved for the highest quality RCTs with low attrition and low differential attrition. This also means that we cannot rule out that differences between the two groups of schools resulted in some bias that affected our impact analyses.

Third, response rates to BPS's school climate survey were extremely low among our study schools during the 2018–2019 school year, and there was no school climate survey during the 2019–2020 school year due to the COVID-19 pandemic. As a result, we are not able to examine impacts on school climate and perceptions of school safety—an important and proximal outcome measure for a restorative practices intervention.

Lastly, we are only able to examine impacts over a short timeframe; one or two years depending on the outcome. Whole school interventions such as Circle Forward can take years to permeate the school culture enough to impact outcomes like disciplinary responses, or even more distal outcomes like student achievement. Additionally, although Circle Forward was generally implemented as intended, we did observe that a substantial minority of staff in most schools had not been trained in Tier 1 circles. It is possible that given more time, larger numbers of staff would have had the opportunity to be trained, and promising early perceptions of school culture shifts would have translated into observable impacts on the use of exclusionary discipline or student achievement. The COVID-19 pandemic further exacerbated this limitation because we were unable to measure impacts on student achievement at the end of the second year of implementation.

Discussion and Implications

Overall, the study team found implementation of Circle Forward generally was happening as intended in the intervention schools. The Tier 1 training occurred as planned and the circles that staff were implementing in schools fit the CF model both in terms of format and expected behavior. The one implementation area that was lower than expected were staff attendance rates at the Tier 1 training. Previous research has indicated that multi-tiered professional development is important for implementation of restorative practices, including initial Tier 1 training that is school-wide and covers basic restorative principles to build knowledge and buy-in (Garnett, Moore et al., 2020; Gregory et al., 2020; Mayworm et al., 2016).

Each intervention school had a dedicated group of RLT members who implemented circles, planned professional development about circles, provided ongoing support for implementation, and helped encourage staff to use circles in their classrooms. With this support from RLT staff as well as knowledge gained from the Tier 1 training, staff at our study schools could create avenues for students to share their thoughts. Often, school staff implemented advisory circles with their students, which encouraged them to connect with others by voicing their opinions and discussing personal and nonacademic topics. Through this process, students also learned more about their peers and teachers and developed stronger relationships. In turn, these positive staff-student relationships and relationships between students helped build a positive, inclusive classroom climate. This finding is consistent with Garnett, Kervick and colleagues (2022), who described teacher and student perspectives in Tier 1 circle implementation and how it created space for student ownership and voice, as well as the clear connection to broader social-emotional learning. Further, others have indicated the importance of explicitly considering student participation and inclusion as a basic indicator of restorative practices implementation (Gregory et al., 2020).

When asked to reflect on whether circles aligned to their school and district policies, school administrators and RLT members perceived strong alignment between circles and policies. In interviews, these groups of respondents said their school's focus on equity and reducing both the school-to-prison pipeline and the number of suspensions matched the goals and principles of the CF model. However, teacher focus group members and survey respondents did not perceive such alignment. These groups of respondents, who were less involved in CF implementation, said it was unclear how CF fit into their school's policies. Because they may not have participated in as much training as school administrators and RLT members, teachers may not perceive a connection between CF and their school's policies. Previous research has highlighted the importance of alignment between school-wide implementation of restorative practices and broader school and district-level policies (Gregory et al., 2020; Martinez et al., 2022). Gregory and colleagues (2020) identify these as necessary restorative practices infrastructure for implementation.

For system change efforts like school-wide restorative practices, constraints related to educators' time are commonly noted (e.g., Martinez et al., 2022). In our findings, concerns around the lack of adequate time to devote to circles were a nuanced issue. Although some survey respondents said it was stressful finding time for circles, survey respondents also generally agreed it was a good use of time. In future implementations, schools may consider offering resources for circles of different lengths. Another idea that arose from our interviews: Have school administrators intentionally set aside time for circles so it does not fall to individual teachers to plan lessons around circles. Also, interview and focus group respondents said other school priorities, such as academic course requirements, prevented them from spending more time on circles. In their efforts to concentrate on academic achievement, school staff may sacrifice time spent on relationship-building through circles, although a recent literature review of restorative practice programs in schools found that restorative practices can improve staff perceptions of school climate (Darling-Hammond et al., 2020).

Also, a subset of staff expressed reluctance to implement the CF model, though overall staff and administrator buy-in was strong. RLT members indicated that this reluctance occurred mainly when circles were first introduced to the school, and the school community was able to move past these challenges. Buy-in has been identified as a critical indicator of school-wide restorative practices implementation (Gregory et al., 2020). As part of this, schools should implement strategies for increasing buy-in across the school, including learning by doing, staff participation in adult circles, transparent norms related to power and hierarchy, and open discussion of whether values conflict with initiatives. In addition to these strategies to increase buy-in, Gregory and colleagues (2020) identify other characteristics of school-wide buy-in, including the blending and integration of restorative practices with other programs (e.g., social-emotional learning, equity initiatives), distributed leadership across school groups (including

students, administrators, teachers and staff), a school team leading strategic decisions for Restorative practices, and clear processes are in place for staff to collaborate when applying restorative practices. Further, regardless of the approaches used, whole-school buy-in may take several years to reach full implementation and school-wide buy-in (González, 2014). At our site visit schools, RLT members echoed these findings from the literature and reported that full buy-in from all staff was likely not possible. To increase buy-in, RLT members provided suggestions for the future such as having administrators attend trainings on restorative practices and offering resources and examples of circle implementation to teachers and staff.

Last, the COVID-19 pandemic affected school staff members' abilities to continue implementation of the CF model. Although RLT members felt moderately prepared to continue implementation after the CRJ coach left the school, they perceived that implementation of circles decreased. This was largely due to the transition to remote learning, staff turnover during the pandemic and a lack of CF training for new staff. Despite these challenges, some RLT members said they were able to continue CF implementation over Zoom and used circles to build community during a particularly difficult time. To assist staff, CRJ staff also created a virtual circle guide and provided additional materials for how to adapt in-person circles amid the pandemic. For example, CRJ created a 'prioritizing safety' handout that included tips on adapting the talking piece (e.g., tapping feet instead of having a communal talking piece) and example circles.

Appendix. Psychometrics of scaled constructs

Construct	Items	Rasch reliability	Cronbach's alpha
Fidelity of circle setup	When leading or participating in circles, how often does the following occur? a. Participants gather in a circle. b. There is a centerpiece within the circle. c. The facilitator starts the circle with an opening (e.g., poem, quote, activity) d. A talking piece is used. e. The facilitator ends the circle with a closing (e.g., song, story)	0.61	0.95
Following of circle norms	When leading or participating in circles, how often does the following occur? a. Everyone waits to speak until he or she is holding the talking piece. b. There is enough time for all participants to have an opportunity to share. c. Most participants share/are active members of the circle. d. Participants adhere to circle guidelines.	0.77	0.96
Frequency of perceived challenges during circle	When leading circles with students, how often does the following occur? a. It takes too long for students to get in a circle. b. The circle process itself takes too long. c. None of the students share. d. Student misbehavior derails the circle. e. One or a few students do most of the sharing. f. Students make negative or inappropriate expressions. g. Content that is shared in confidence is later shared outside the circle. h. Language or developmental barriers present barriers to sharing.	0.63	0.77
Perceived changes in school climate	To what extent has implementation of Circle Forward in your school been successful with the following items related to school climate? a. Improving relationships with other staff members b. Improving relationships with students c. Gaining skills in alternative responses to student behavior d. Creating a space to talk about issues of social justice, race, and equity e. Improving staff morale f. Improving classroom culture/climate g. Improving overall school culture/climate	0.88	0.98

References

Acosta, J.D., Chinman, M., Ebener, P., Phillips, A., Xenakis, L. & Malone, P.S. (2016). A cluster-randomized trial of restorative practices: An illustration to spur high-quality research and evaluation. *Journal of Educational and Psychological Consultation*, 26(4), 413-430. doi: 10.1080/10474412.2016.1217488

Acosta, J., Chinman, M., Ebener, P., Malone, P. S., Phillips, A., & Wilks, A. (2019). Evaluation of a whole-school change intervention: Findings from a two-year cluster-randomized trial of the restorative practices intervention. Journal of Youth and Adolescence, 48(5), 876–890

Andrich, D. (1978). A rating scale formulation for ordered response categories. *Psychometrika,* 43(4)*,* 561–573.

Antara, V.A. Jr., Evans, K.R. & Lester, J.N. (2013). Restorative justice in education: What we know so far. *Middle School Journal,* 44(5), 57–63. doi: 10.1080/00940771.2013.11461873

Augustine, C.H., Engberg, J., Grimm, G.E., Lee, E., Wang, E.L., Christianson, K. & Joseph, A.A. (2018). *Can restorative practices improve school climate and curb suspensions? An evaluation of the impact of restorative practices in a mid-size urban school district.* RAND.

Balfanz, R., Byrnes, V., & Fox, J. (2014). Sent home and put off-track: The antecedents, disproportionalities, and consequences of being suspended in the ninth grade. *Journal of Applied Research on Children: Informing Policy for Children at Risk,* 5(2), 13.

Bayram Özdemir, S.B., Yanagida, T. & Özdemir, M. (2022). Bystanders of ethnic victimization: Do classroom context and teachers' approach matter for how adolescents intend to act? *Child Development*, 93(5), 1540-1558. doi: 10.1111/cdev.13822

Bonell, C., Allen, E., Warren, E., McGowan, J., Bevilacqua, L., Jamal, F., Legood, R., Wiggins, M., Opondo, C., & Mathiot, A. (2018). Efects of the Learning Together intervention on bullying and aggression in English secondary schools (INCLUSIVE): A cluster randomised controlled trial. The Lancet, 392(10163), 2452–2464.

Boulton, J. & Mirsky, L. (2006). Restorative practices as a tool for organizational change. *Reclaiming Children & Youth*, 15(2), 89-91.

Boyes-Watson, C. (2005). Seeds of change: Using peacemaking circles to build a village for every child. *Child Welfare*, 84(2), 191–208.

Boyes-Watson, C. & Pranis, K. (2014). *Circle forward: Building a restorative school community.* St. Paul, MN: Living Justice Press.

Bradshaw, C.P., Cohen, J., Espelage, D.L. & Nation, M. (2021). Addressing school safety through comprehensive school climate approaches. *School Psychology Review*, 50(2-3), 221-236. doi: 10.1080/2372966X.2021.1926321

Capp, G., Astor, R.A. & Gilreath, T.D. (2020). Advancing a conceptual and empirical model of school climate for school staff in California. *Journal of School Violence*, 19(2), 107-121.

Cohen, J., McCabe, L., Michelli, N. M., & Pickeral, T. (2009). School climate: Research, policy, practice, and teacher education. *Teachers College Record*, *111*(1), 180–213.

Darling-Hammond, S., Fronius, T.A., Sutherland, H., Guckenburg, S., Petrosino, A. & Hurley, N. (2020). Effectiveness of restorative justice in US K-12 schools: A review of quantitative research. *Contemporary School Psychology*, 24, 295-308. doi: 10.1007/s40688-020-00290-0

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

Garnett, B. R., Kervick, C. T., Moore, M., Ballysingh, T. A., & Smith, L. C. (2022). School staff and youth perspectives of Tier 1 restorative practices classroom circles. *School Psychology Review*, *51*(1), 112-126. https://doi.org/10.1080/2372966X.2020.1795557

Garnett, B., Moore, M., Kidde, J., Ballysingh, T.A., Kervick, C.T., Bedinger, L. ... & Sparks, H. (2020). Needs and readiness assessments for implementing school-wide restorative practices. *Improving Schools*, 23(1), 21-32. doi: 10.1177/1365480219836529

González, T. (2014). Socializing schools: Addressing racial disparities in discipline through restorative justice. In D.J. Losen (Ed.), *Closing the School Discipline Gap: Equitable Remedies for Excessive Exclusion*.

González, T. (2016). Restorative justice from the margins to the center: The emergence of a new norm in school discipline. *Howard LJ*, *60*, 267. https://ssrn.com/abstract=2901108

González, T., Sattler, H. & Buth, A.J. (2019). New directions in whole-school restorative justice implementation. *Conflict Resolution Quarterly*, 36(3), 207-220.

González, T., Etow, A., & De La Vega, C. (2019). Health equity, school discipline reform, and restorative justice. *The Journal of Law, Medicine & Ethics*, *47*(2_suppl), 47–50.

Gregory, A., Huang, F.L., Anyon, Y., Greer, E. & Downing, B. (2018). An examination of restorative interventions and racial equity in out-of-school suspensions. *School Psychology Review*, 47(2), 167-182.

Gregory, A. & Evans, K.R. (2020). *The starts and stumbles of restorative justice in education: Where do we go from here?* National Education Policy Center. Retrieved from https://files.eric.ed.gov/fulltext/ED605800.pdf (last accessed 27 December 2022).

Gregory, A., Ward-Seidel, A.R. & Carter, K.V. (2020). Twelve indicators of restorative practices implementation: A framework for educational leaders. *Journal of Educational and Psychological Consultation*, 31(2), 147-179. doi: 10.1080/10474412.2020.1824788

Guckenberg, S., Hurley, N., Persson, H., Fronius, T. & Petrosino, A. (2016). Restorative justice in U.S. schools: Practitioners' perspectives. San Francisco, CA: WestEd. Retrieved from https://www.wested.org (last accessed 27 December 2022).

Hantzopoulos, M. (2013). Building relationships to engage at-risk youth: A case study of a New York City public high school. In M. Matsumoto (Ed)., *Education and Disadvantaged Children and Young People*, 11-30. Bloomsbury.

Hong, J.S., Espelage, D.L. & Lee, J.M. (2018). School climate and bullying prevention programs. In H. Shapiro (Ed.), *The Wiley handbook on violence in education: Forms, factors, and preventions*, 359-374. Wiley.

Ingraham, C.L., Hokoda, A., Moehlenbruck, D., Karafin, M., Manzo, C. & Ramirez, D. (2016). Consultation and collaboration to develop and implement restorative practices in a culturally and linguistically diverse elementary school. *Journal of Educational and Psychological Consultation*, 26(4), 354-384. doi: 10.1080/10474412.2015.1124782

Jain, S., Bassey, H., Brown, M. & Kalkra, P. (2014). Restorative justice in Oakland schools: Implementation and impacts. Oakland, CA: Oakland Unified School District and Data in Action.

LiCalsi, C., Osher, D., & Bailey, P. (2021). *An empirical examination of the effects of suspension and suspension severity on behavioral and academic outcomes*.

Linacre, J.M. (2015). Winsteps® (Version 3.91.0) [Computer software]. Beaverton, OR: Winsteps.com.

Losen, D. J., & Whitaker, A. (2018, August 24). *11 million days lost: Race, discipline, and safety at U.S. public schools* (Part 1). American Civil Liberties Union.

Martinez, A., Villegas, L., Hassoun Ayoub, L., Jensen, E. & Miller, M. (2022). Restorative justice and school-wide transformation: Identifying drivers of implementation and system change. *Journal of School Violence*, 21(2), 190-205. doi: 10.1080/15388220.2022.2039682

Mayworm, A.M., Sharkey, J.D., Hunnicutt, K.L. & Schiedel, K.C. (2016). Teacher consultation to enhance implementation of school-based restorative justice. *Journal of Educational and Psychological Consultation*, 26(4), 385-412. doi: 10.1080/10474412.2016.1196364

McCammon, B. (2020). *Restorative practices at school.* Berkeley, CA: Ulysses Press.

McCluskey, G. (2018). Restorative approaches in schools: Current practices, future directions. In J. Deakin, A. Kupchik & E. Taylor (Eds.), *The Palgrave International Handbook of School Discipline, Surveillance, and Social Control* (pp. 573-593). Cham, Switzerland: Palgrave Macmillan.

Morgan, E., Salomon, N., Plotkin, M., & Cohen, R. (2014). *The school discipline consensus report: Strategies from the field to keep students engaged in school and out of the juvenile justice system.* The Council of State Governments Justice Center.

Nickerson, A. B. (2018). Can SEL Reduce School Violence?. *Educational Leadership*, *76*(2), 46-50.

Noltemeyer, A. L., Ward, R. M., & Mcloughlin, C. (2015). Relationship between school suspension and student outcomes: A meta-analysis. *School Psychology Review*, *44*(2), 224–240.

Osher, D., Bear, G. G., Sprague, J. R., & Doyle, W. (2010). How can we improve school discipline? *Educational Researcher*, *39*(1), 48–58.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests* (Exp. ed.). Chicago, IL: University of Chicago Press.

Schiff, M. (2018). Can restorative justice disrupt the 'school-to-prison pipeline?'. *Contemporary Justice Review*. 21(2). 121-139. Doi: 10.1080/10282580.2018.1455509.

Schumacher, A. (2014). Talking circles for adolescent girls in an urban high school: A restorative practices program for building friendships and developing emotional literacy skills. *SAGE Open*, 4(4). doi: 10.1177/2158244014554204.

Thapa, A., Cohen, J., Guffey, S. & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research*, 83(3), 357-385. doi: 10.3102/0034654313483907

Vaandering, D. (2014). Relational restorative justice pedagogy in educator professional development. *Curriculum Inquiry, 44*(4), 508–530.

Vaandering, D. (2014). Implementing restorative justice practice in schools: What pedagogy reveals. *Journal of Peace Education*, 11(1), 64-80. doi: 10.1080/17400201.2013.794335

Voight, A. & Nation, M. (2016). Practices for improving secondary school climate: A systematic review of the research literature. *American Journal of Community Psychology*, 58(1-2), 174-191. doi: 10.1002/ajcp.12074

Wang, C., Berry, B. & Swearer, S.M. (2013). The critical role of school climate in effective bullying prevention. *Theory Into Practice*, 52(4), 296-302. doi: 10.1080/00405841.2013.829735

Wang, E.L. & Lee, E. (2019). The use of responsive circles in schools: An exploratory study. *Journal of Positive Behavior Interventions*, 21(3), 181-194. doi: 10.1177/1098300718793428

Wright, B.D. & Masters, G.N. (1982). *Rating scale analysis*. Chicago, IL: MESA Press.

Wright, B.D. & Stone, M.H. (1979). *Best test design*. Chicago, IL: MESA Press.

Wolf, K. C., & Kupchik, A. (2017). School suspensions and adverse experiences in adulthood. *Justice Quarterly*, *34*(3), 407–430.

Yin, R.K. (2009). *Case study research: Design and methods.* Thousand Oaks, CA: SAGE.

Zakszeski, B. & Rutherford, L. (2021). Mind the gap: A systematic review of research on restorative practices in schools. *School Psychology Review*, 50(2-3), 371-387.

Zehr, H. (2002). *Little book of restorative justice*. Intercourse, PA: Good Books.

About the American Institutes for Research®

Established in 1946, the American Institutes for Research® (AIR®) is a nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally in the areas of education, health, and the workforce. AIR's work is driven by its mission to generate and use rigorous evidence that contributes to a better, more equitable world. With headquarters in Arlington, Virginia, AIR has offices across the U.S. and abroad. For more information, visit AIR.ORG.

AIR® Headquarters
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
+1.202.403.5000 | **AIR.ORG**

21044_03/23

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.