The Impact of Restorative Practices with Diplomas Now on School Climate and Teachers’ Turnover Intentions: Evidence from a Cluster Multi-Site Randomized Control Trial

Ashley A. Grant, Douglas J. Mac Iver & Martha Abele Mac Iver


To link to this article: https://doi.org/10.1080/19345747.2021.2018745

View supplementary material

Published online: 26 Jan 2022.

Submit your article to this journal

Article views: 318

View related articles

View Crossmark data
The Impact of Restorative Practices with Diplomas Now on School Climate and Teachers’ Turnover Intentions: Evidence from a Cluster Multi-Site Randomized Control Trial

Ashley A. Grant\textsuperscript{a}, Douglas J. Mac Iver\textsuperscript{a} and Martha Abele Mac Iver\textsuperscript{a}

\textsuperscript{a}Center for Social Organization of Schools, Johns Hopkins University School of Education, Baltimore, Maryland, USA

\textbf{ABSTRACT}

High teacher turnover rates and shortages of quality teachers plague many large, urban U.S. school districts disadvantaging their students who often already come from traditionally underserved populations and communities. Restorative Practices is a quickly growing whole school approach to community building and discipline, but little is known about how it impacts teachers and their career decisions. This study specifically investigates whether assignment to Restorative Practices (in combination with Diplomas Now, RP w/DN) improves school climate and increases teachers’ reported intentions to remain at their school? To answer this question, we analyzed data from a multi-site, cluster randomized control trial of in 25 schools from seven large cities across the United States. Our intent-to-treat analyses find that RP w/DN has a significantly positive effect on school climate (ES = .15, .27). Impacts on teachers’ turnover intentions were non-significant, but it significantly reduced perceptions of problematic teacher absenteeism (a related, exploratory outcome). Overall, the findings from this study suggest that Restorative Practices as a policy can improve school climate, but its relationship with teachers’ career decisions is more complicated.

In the United States, fears of an inadequate supply of quality teachers have prompted the study of why teachers leave their classrooms (Ingersoll, 2003; Sutcher et al., 2016). Underfunded and overburdened schools in large urban districts face the largest losses in their teaching forces; less than half of the teachers in these districts remain after five years (Ingersoll, 2004). Individual schools face staggering annual turnover (and therefore replacement) rates at an average of one-fifth of their teaching staff (Holme et al., 2018; Papay et al., 2017).

Most of the research on teacher turnover has examined descriptive patterns, particularly observed variation in turnover based on school working conditions (Nguyen et al., 2020; Simon & Johnson, 2015). Experimental evidence on potential solutions to reduce teacher turnover is rare, with the exception of a few recent studies (Borman & Dowling,
A. A. GRANT ET AL.

2008). Varied teacher hiring and transfer procedures and policies, as well as job opportunity and choice patterns, generally preclude the random assignment of individual teachers to schools. The few evaluations of such policy interventions, including the Chicago Teacher Advancement Program (Glazerman et al., 2010) or pay-for-performance initiatives (Springer et al., 2010) have shown limited, or null, impacts on turnover at large costs. More frequently, longitudinal observational studies of programs such as the North Carolina Bonus Program have shown small impacts on turnover, but remain open to validity threats from teacher selection (Clotfelter et al., 2008).

This study seeks to add to the current research on teacher turnover by evaluating the impact of the SaferSanerSchools schoolwide model of Restorative Practices (RP) when combined with teacher and student supports provided by the Diplomas Now intervention (Corrin et al., 2016). RP is a relatively new school intervention1 aimed at whole school change: reducing punitive disciplinary measures, eliminating disciplinary inequities, and promoting a more positive school environment (Costello et al., 2009; Evans et al., 2013). Although many schools and districts have raced ahead to implement this intervention, evidence of its efficacy largely remains limited to observational and case studies (Darling-Hammond et al., 2020; Song & Swearer, 2016). This study examines how teachers in schools that were randomized to combine Restorative Practices with Diplomas Now (RP w/DN) responded to this two-pronged intervention in their intentions to leave their school and the profession of teaching.

Few studies have examined RP’s impact on teachers (Hurley et al., 2015). Theoretically, RP is designed to improve the whole school environment and positively impact all school community members’ relationships and satisfaction. If RP succeeds at improving school climate, a factor teachers report as important to their decisions to stay, teachers should want to stay at their school and teacher turnover should be reduced. However, some qualitative studies and media reports of opposition to RP have picked up on potential resistance from teachers who claim that RP could instead lead to less organized schools with no consequences for student behavior and more stress for teachers (Dominus, 2016; Lustick, 2017b).

This study tests those hypotheses by examining the direct impact of RP w/DN on teachers’ turnover intentions. We also test whether RP w/DN had an impact on school climate as a potential theoretical explanation of this relationship.

**Literature Review**

**Teacher Turnover: The Problem and Its Causes**

Teacher turnover rates in the United States are high and only growing, particularly in schools serving traditionally underserved student populations. Nationally, 13% of teachers leave the profession each year, with rates of at least 20% in urban, high-poverty schools (Ingersoll et al., 2014). Fewer teachers remain until retirement, resulting in a workforce that is greener (younger) and schools with less teacher pedagogical and organizational expertise. Schools’ loss of human capital additionally requires them to

---

1The Real Justice program, focused on restorative conferencing, was founded in 1994. In 1999, the founders of this program established the broader International Institute of Restorative Practices (IIRP) to train professionals not only in formal restorative conferencing but also in a broad array of informal and preventative restorative practices that build community and trusting, empowered relationships. See below for a more detailed description.
pay to replace teachers who leave every year and means they cannot offer their students high quality learning opportunities afforded by more experienced teachers and more stable schools (Ingersoll & Merrill, 2017; Ronfeldt et al., 2013). Teacher absenteeism, related but distinct from turnover (Mowday et al., 1982), similarly imposes extra financial burden and negatively impacts student achievement (Miller et al., 2008).

Some researchers debate the harmfulness of national rates of turnover, but few argue against the urgency of the higher rates found in most under-resourced and high-poverty urban schools: on average, double that found in other schools (Ingersoll, 2004). Almost half of all public school teacher turnover occurs in a quarter of public schools (Ingersoll, 2004). Much of this turnover occurs between schools as teachers systematically sort away from certain schools; more qualified teachers leave the most challenged schools, creating an unequal distribution of teacher experience and quality (Goldhaber et al., 2018). Although turnover of ineffective teachers is not a negative outcome, the high turnover rate in high poverty schools is not primarily due to this type of turnover.

In analyzing teacher turnover in high poverty schools, researchers have previously used students’ characteristics to explain the higher rates—suggesting students from higher poverty and minority racial backgrounds create more difficult and thus undesirable teaching environments (e.g., Scafidi et al., 2007). More recently, however, researchers have focused on the role of working conditions in high poverty schools, which often experience lower quality leadership, less collegiality, and generally less positive school climates (Simon & Johnson, 2015). Theoretically, teachers in schools with more a more positive and supportive school climate are more likely to feel contented, successful, and productive in their work and feel like further investing their time and energy in their school and the teaching profession (Johnson, 2020; Johnson & Birkeland, 2003). Many studies have shown that teachers tend to leave schools with negative school climates (e.g., Kraft et al., 2016), but none of these studies have been able to incorporate randomized designs for stronger causal evidence.

**School Climate**

The term school climate is broadly defined in the literature as the “patterns of people’s experiences of school life and reflects norms, goals, values, interpersonal relationships, teaching and learning practices, and organizational structures… that support people feeling socially, emotionally and physically safe” (p. 358, from the National School Climate Center, as quoted in Thapa et al., 2013). From an ecological theory approach, the school climate, or experienced environment, is comparable to eco-systems or the meteorological climate—all contain layers of dynamic, unique systems that mutually influence each other (Bronfenbrenner, 1977; Cornell et al., 2016). In the case of schools, the systems contributing to a school’s climate include elements like safety, physical environment, teaching and learning, relationships, and leadership and staff relations, which interact with one another but also uniquely contribute to the school climate and can thus be uniquely impacted by interventions aimed at improving school climate (Thapa et al., 2013).

A more positive school climate is associated with most desired schooling outcomes: higher student achievement, better attendance, less violence and aggression, and more positive social and emotional health (Berkowitz et al., 2017). Teachers in schools with more positive school climates are more likely to be satisfied with their job, show greater
commitment, and are less likely to be absent or leave their school (Guin, 2004; Imants & van Zolen, 1995; Kraft et al., 2016). The pervasive and important consequences of school climate have spurred development of many interventions aimed at improving school climate. School Wide Positive Behavior Incentives and Supports (SWPBIS) is one such intervention that has been shown to improve school climate (Bradshaw et al., 2008; Horner et al., 2009). Opponents of SWPBIS, however, critique its focus of extrinsic rewards and its weak Tier 3 responses to more serious issues (Swain-Bradway et al., 2015; Wilson, 2015). They also worry that SWPBIS does not help wrongdoers understand the impacts of their wrongdoing on others and internalize new norms. Additionally, no studies have tested the effects of PBIS or other school climate interventions on teachers’ commitment or turnover.

**A Restorative Practices (RP) Solution?**

Restorative Practices (Costello et al., 2009) is another intervention, growing in popularity, that seeks to improve school climate and reduce suspension rates and the racial disciplinary gap (Anyon et al., 2016), and which may have the potential to improve teacher retention as well. This intervention, also known as restorative justice practices, is an alternative approach to school discipline, culture, and community building grounded in an ethos of reparation and rehabilitation (Gonzalez, 2012). The more traditional, punitive approach to discipline and justice found in the United States (schools and justice system) emphasizes imposing punishment on wrongdoers. In contrast, a response based in the restorative approach emphasizes a process of dialogue that includes the voices of victims, wrongdoers, their families, and other members of the school community that aims to repair harm and enable repentance, reparation, and reconciliation. This dialogue helps the wrongdoer understand the harm he or she has produced and also helps others understand what the wrongdoer was thinking and feeling at the time of the incident. The dialogue provides an opportunity for the wrongdoer to express remorse, to suggest ways of beginning to repair the harm, and to work collectively with the others involved in the dialogue to reach an agreement on how the wrongdoer can help make things right and how relationships can be restored and the wrongdoer reintegrated into the community (Braithwaite, 1999; Zehr, 2015).

There are many different programs that implement a restorative approach in schools (Evans et al., 2013). These programs share the goal of proactively establishing a caring, supportive, and accountable environment that addresses, rather than ignores, inappropriate behaviors and promotes a fair process of participatory decision-making and student learning. RP programs build this environment through proactive empathy and community building exercises such as proactive circles that regularly encourage students (and teachers) to share goals, thoughts, or feelings with one another. For example, teachers may give students the opportunity during their brief (5–10 min) daily circle time to share with their classmates about someone they consider a role model; or, teachers could even use the circle structure to give students time to reflect on a homework reading.

RP programs sustain this environment through their response to situations where negative actions or behaviors have impacted others and harm has been done to the
school community and/or its members. These responses encourage all members of the school community to reflect on and heal the harm: using affective statements to express how they are impacted by others’ behaviors, asking restorative questions of both wrongdoers and those harmed to increase understanding of who has been harmed and what harm has been done, and giving both wrongdoers and those impacted a say in what needs to happen to make things right. For example, if one student continually disrupted class, students in the class could circle up to each share how the disruption affected them. After listening to their classmates, the disruptive student would then devise a way they could make up for their negative effect on their classmates, such as helping their fellow students with notes to make up for class time lost and prevent future disruptions, which the whole class discusses and then adopts. These responses help members of the school community to develop compassion for and understanding of each other and to build or rebuild supportive relationships. In addition to using affective statements and restorative questions, RP programs often feature small impromptu mediation conferences to address misbehavior between students, responsive peacemaking circles, and formal restorative conferences that seek out appropriate and productive restorative sanctions to repair harm while providing ways for the offenders to be reintegrated into the school community and reclaim their good name (Evans et al., 2013).

Much of the current RP research has focused on the effects of RP on school disciplinary and student outcomes (Darling-Hammond et al., 2020). Observed immediate and direct impacts of RP on school discipline statistics include an 87% drop in out-of-school suspensions in San Antonio, Texas (Armour, 2015); a 52% drop in violent acts in Philadelphia, PA (International Institute for Restorative Practices, 2009); and a 57% drop in discipline referrals in Minnesota (Riestenberg, 2003). Anecdotal and descriptive research results have illustrated that RP generates greater student connectedness, student self-efficacy (in Oakland, CA; Jain et al., 2014), and generally positive perceptions of students and teachers regarding RP (in Minneapolis, MN; McMorris et al., 2013).

Most of this initial evidence, however, does not permit strong causal inferences due to a lack of control or comparison groups (Darling-Hammond et al., 2020). To date, two randomized control trials have shown positive impacts on schools through the reduction in exclusionary discipline (Augustine et al., 2019; Grant et al., 2020), but another study found no significant impacts of assignment to treatment (Acosta et al., 2019). In the study of 44 schools in Pittsburgh, PA (Augustine et al., 2019), assignment to RP significantly reduced suspensions but had no significant effect on student attendance or arrests. Additionally, assignment to RP had a negative impact on teacher performance (as measured by student reports and value-added) and a marginally significant negative effect on student achievement in grades 3–8 (but not in high school grades). Augustine and colleagues posit, from supplementary probing analyses, that these negative findings can be attributed to the quality of implementation of RP (teachers and schools with low usage of RP). An unpublished study (Grant et al., 2020) based on a cluster RCT found a significant impact of RP on one measure of student discipline (suspended three or more days) and a promising ($p < .10$) potential impact on student chronic absenteeism. By contrast, in the Maine study of 13 middle schools, Acosta et al. (2019) found no significant impact of assignment to RP on any outcomes (though there
was a positive relationship between students’ self-reported experience of RP practices and several social emotional outcomes).

**The Promise of RP to Improve School Climate**

Do the positive effects of RP (e.g., reducing exclusionary discipline and bullying) extend to the whole school environment as intended? The whole-school RP model emphasizes that if the goal is to improve the whole school climate, RP must be implemented to include the whole school community, including all students, teachers and staff, and even students’ families. Theoretically, principals should lead RP with staff, promoting relationship development and community throughout the school. In accordance with this theory, Mirsky and Wachtel (2007) found in their case study of alternative schools in Pennsylvania that RP could improve the whole school climate (measured by student behavior and reported prosocial values). Jain et al. (2014) found that 70% of teachers in 24 Oakland schools implementing RP reported that RP seemed to be improving school climate (via a single reported item). Gregory et al. (2016), in an observational study of two schools in their first year implementing RP, similarly noted that more restorative approaches improved student-teacher relationships. However, the apparent evidence that RP improves the overall school climate remains limited to a few studies, has not rigorously measured all aspects of school climate and largely does not meet rigorous evidence standards because of weak research designs (Hurley et al., 2015). In the few published randomized control trials, Augustine et al. (2019) found that teachers at RP schools reported a more positive school climate, but students reported a more negative classroom level climate (and less student engagement, as measured by the Tripod student survey), and Acosta et al. (2019) found no statistically significant differences in school climate between schools assigned to RP and those assigned to the control condition. Looking at an RCT of a holistic SEL intervention incorporating RP, Bonell et al. (2018) found significant decreases in bullying but not in aggression. School climate is an important ecological factor that strongly predicts improved teacher retention, student achievement, and school functioning (Kraft et al., 2016; Thapa et al., 2013). If RP can improve the school climate and working conditions of schools, then this improved work setting could potentially increase teacher retention rates. This study tests this hypothesis empirically.

**RP Within the Diplomas Now Whole-School Reform Approach**

In the current study, we particularly examine the impact of RP in the context of another whole school intervention: Diplomas Now (DN). DN is a whole-school reform model—featuring components provided by Talent Development Secondary, City Year, and Communities in Schools—that aims to reduce secondary school students’ development of early warning indicators of dropout risk in order to help students earn on-time promotion and graduation in high poverty middle and high schools (Corrin et al., 2014, 2016; Sepanik et al., 2015). The DN reform model rests upon four pillars of practices to support whole school transformation: teacher teams and small learning communities, strong curriculum and instruction with professional
development, tiered student supports and interventions guided by an early warning system, and integrated on-site supports involving additional staffing. At DN schools, teachers are organized in grade-level teams in which they share students and receive PD and instructional coaching. Additional staff help to facilitate close monitoring of student early warning indicator data and implementation of timely and tailored interventions for struggling students.

As Corrin et al. (2016) explain in their logic model for DN, the practices in these four pillars are expected to lead to more positive practices and school environments, which in turn are expected to have a positive impact on student attendance, behavior, and course performance. The organization of the school into small learning communities in which interdisciplinary teacher teams share groups of students is expected to yield a more positive relationships and a generally more positive school climate. The provision of professional development and instructional coaching, as well as a curriculum for college readiness, is expected to improve instruction, student engagement, and student course performance. The early warning system to monitor student data and guide timely and targeted interventions is expected to improve student engagement, and particularly attendance, behavior, and course performance. Integrated on-site human capital supports provide the additional person power to ensure that interventions can be implemented effectively, leading to improved student outcomes.

Prior studies of DN impacts have shown a positive impact on teachers’ perception of school climate and on student-reported participation in afterschool activities (both academic and non-academic) and positive relationships with an adult at school who was not teacher (Corrin et al., 2014; Sepanik et al., 2015). In addition, there was a positive impact of DN for middle school students on attending at least 90% of the time (not chronically absent) and on being on-track to graduation (having no warning indicators in attendance, behavior, or course performance; Corrin et al., 2016). DN’s effects on high school students were not significant. Persistent issues with exclusionary discipline and school climate in DN schools led the developer to combine DN with RP, an approach that more directly targets these issues. It was hypothesized that these interventions would integrate well together because (1) both are implemented at the whole-school level; (2) student support staff (e.g., City Year corps members) could help implement both interventions; and 3) RP was consistent with DN’s approach to instruction and the importance or helping students recover from the effects of their prior behavior.

**Purpose of the Current Study**

This study uses data from a cluster randomized control trial to address the primary research questions that follow from the research literature synthesized above. As discussed in more detail below, schools were randomly assigned to implement Restorative Practices with Diplomas Now or to continue business as usual (the control condition). The study seeks to answer the following research questions:

RQ1a) What impact does Restorative Practices in combination with Diplomas Now have on i) teachers’ and ii) students’ perceptions of their school’s climate?

RQ1b) To what extent does this impact vary across different dimensions of school climate (supportive environment, problematic behaviors, and a measure of professional collaboration [teachers] or responsive instructional practices [students])?
RQ2) What impact does Restorative Practices in combination with Diplomas Now have on teachers’ reported intentions to: a) remain at their school, and b) remain in the teaching profession?

RQ3) What impact does Restorative Practices in combination with Diplomas Now have on perceptions of problematic teacher absenteeism?

**Methodology**

**Intervention(s)**

The intervention in this study involved a combination of Diplomas Now (DN, described above) and Restorative Practices (RP). In treatment schools, RP was integrated into the DN model to promote relationship building and provide an alternative approach to school discipline. This study specifically examines the model of RP from the International Institute for Restorative Practices (IIRP), called *SaferSanerSchools Whole School Change Program*. Unlike some RP programs that emphasize certain practices or focus on parts of the school structure, IIRP’s model promotes an approach that involves whole school change. The IIRP model specifies 11 essential elements for full implementation, e.g., affective statements (“personal expressions of feelings in response to specific behaviors”) and restorative staff community (“a staff that models and consistently uses restorative practices with each other to build and maintain healthy adult relationships”; Costello et al., 2009).

Figure 1 provides the logic model for the combined intervention, embedding RP within DN, and shows how the components of the combined intervention and intended outcomes relate to each other. This logic model focuses only on the outcomes relevant to the current study—school climate and teacher turnover measures, excluding the student attendance, suspension/exclusion, and course performance outcomes of the larger DN and RP w/DN studies. As the figure indicates, RP (through its professional development and implementation of essential practices) is expected to improve measures of

![Figure 1. Logic model of combined RP w/DN intervention. Note. PD: professional development; EWS: early warning system.](image)
school climate (reducing problematic student behaviors, increasing perceptions of a supportive environment, and increasing perceptions of a positive and collaborative learning climate for teachers and students). DN’s small learning community organizational structure is also expected to improve school climate measures of positive relationships between students and teachers. The DN pillar concerned with professional development and curriculum aims to increase teacher professional community and responsive instructional practices for students. The DN tiered system of student supports and integrated site supports in human capital to provide supports to students seek to improve student behavior in ways that positively influence school climate. Improved school climate is expected to lead to reductions in problematic teacher absence and to improved teacher retention intentions.

Procedures

Recruitment of Schools

Because of the complexity of the sample, we first describe the design and recruitment of schools for the original DN validation study and then describe the RP sub-study design. Schools were originally recruited for the DN i3 validation study in two waves to begin in Fall 2011 and 2012 (for full information on the DN model and initial experiment see: Corrin et al., 2014, 2016; Sepanik et al., 2015). Recruitment focused on low-performing middle and high schools in large urban districts across the US (in accordance with the specific goals of the DN intervention to keep sixth and ninth-grade students on track to graduation and facilitate school turnaround efforts to support these students). Many of the recruited school districts were already receiving services in some of their schools from Communities in Schools and City Year (in order to reduce the costs that would have been involved in opening and staffing new CIS and CY field offices just for the purposes of the validation study).

Randomization

The original DN Validation Study utilized a cluster randomized control trial that incorporated a block design for assignment to treatment. District leaders nominated schools to participate. Leaders from nominated schools attended awareness sessions to learn more about the DN model and about the randomized evaluation study. Once schools agreed to participate, they were organized by the third-party-evaluator (MDRC) into blocks by school district, school level (middle or high school), and month and year of recruitment. A total of 62 schools agreed to participate in the original DN study from 11 districts (creating 23 blocks by school level, district, and time of randomization). Within each block, all of which had at least two schools, schools were randomly assigned via a lottery to one of two conditions: the DN treatment or “business as usual” control group. Randomization occurred at the school level because DN’s interventions are implemented at the school level. Possible contamination and spillover effects were still possible at the district level (within blocks) if districts advocated for policies and programs similar to DN’s components.
Restorative Practices Sub-Study

Viewing RP as a complementary whole school intervention that could be fruitfully combined with DN, the developers collaborated on a sub-study involving the two reforms. All districts in the DN Validation Study sample were contacted with information about the RP sub-study in 2014. Eight districts agreed for at least some of their randomization blocks to participate in the sub-study, in which DN treatment schools were to implement RP in addition to DN. In total, based on these district level decisions, the sub-study included all schools in 12 of the original 23 randomization blocks. Already randomized treatment schools received the RP intervention professional development resources, while already randomized control schools in these blocks continued “business as usual” (to implement reforms of their choice that did not include supports provided to DN schools in the study). Because participation in the RP sub-study (and addition of RP to DN) was nonrandom (influenced by district decisions), it is not possible to causally assess the impact of RP w/DN compared to DN alone.

Sample of Schools

This study focuses on the 12 blocks (33 schools) from the original DN study that later agreed to participate in the RP sub-study. These 33 schools come from 8 districts, 7 of which are among the top 100 largest districts in size according to number of students and represent the New England, Northeastern, Southeastern, South Central, Midwestern, and Western regions of the United States (Sable et al., 2010). Table 1 provides descriptive characteristics for the sub-study schools in 2011 or 2012 (depending on the time of initial recruitment and randomization).

Attrition and Analytic Sample

Attrition is defined at the school level as schools that either closed or did not respond to this study’s outcome year survey (in Spring 2016). Attrition is defined at the school level because treatment was assigned at the school level and because individual teachers

Table 1. Differences at baseline (prior to randomization) in the school composition characteristics of treatment and control schools in the analytic sample.

<table>
<thead>
<tr>
<th>School characteristic</th>
<th>Treatment</th>
<th>Control</th>
<th>Difference (T-C)</th>
<th>p-Value</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrollment (students)</td>
<td>1,029.4</td>
<td>969.8</td>
<td>59.6</td>
<td>.843</td>
<td>.082</td>
</tr>
<tr>
<td>% FRL</td>
<td>79.0%</td>
<td>78.7%</td>
<td>0.3%</td>
<td>.970</td>
<td>.015</td>
</tr>
<tr>
<td>% Minority</td>
<td>97.1%</td>
<td>94.6%</td>
<td>2.6%</td>
<td>.301</td>
<td>.423</td>
</tr>
<tr>
<td>% SPED</td>
<td>16.9%</td>
<td>20.6%</td>
<td>-3.7%</td>
<td>.399</td>
<td>.346</td>
</tr>
<tr>
<td>% Male</td>
<td>50.8%</td>
<td>54.8%</td>
<td>-3.9%</td>
<td>.426</td>
<td>.327</td>
</tr>
<tr>
<td>Attendance (%)</td>
<td>90.1%</td>
<td>89.8%</td>
<td>0.3%</td>
<td>.922</td>
<td>.038</td>
</tr>
<tr>
<td>Exclusionary discipline days (per student)</td>
<td>2.1</td>
<td>1.8</td>
<td>0.3</td>
<td>.777</td>
<td>.012</td>
</tr>
<tr>
<td>Sample size</td>
<td>13</td>
<td>12</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
</tbody>
</table>

Note. Based on administrative data from the Common Core of Data (CCD) from 2010 to 2011 or 2011 to 2012 based on DN cohort (year of randomization) and MDRC data collection from administrative records. There were no statistically significant differences ($p < .1, df = 23$) between treatment and control groups in the analytic sample.

2Districts most supportive of RP may have responded to participate first, and their decisions about which blocks of schools within the district would participate could have been affected by school leader willingness to participate. Funding was not sufficient for additional blocks of schools to participate in the RP with DN sub-study.

3Also see Supplemental Table A1.
and students are not uniquely identified in data collection and cannot be tracked over time. Supplemental Table A2 reports the response of schools and teachers across blocks and conditions. Overall, six schools did not respond to the 2016 teacher survey (two treatment and four control schools).

The reduced analytic sample for this study includes only “complete” blocks where teachers from at least one control and one treatment school responded to the Spring 2016 survey (which contained the measure of the primary outcome variables). This reduced sample includes teachers in 9 of the 12 blocks and 25 of the 33 schools. The 8 sub-study schools that were excluded from the final analytic sample (due to school closure and/or non-response to the teacher survey producing incomplete or empty blocks) were split evenly among treatment and control groups (4 in each). Within the 9 complete blocks, 686 teachers responded; Table 2 provides descriptive characteristics for the teachers in the analytic sample. (Threats to the validity of the experiment from attrition are further evaluated in the first section of the Results.)

**Intervention Implementation**

Implementation of the SaferSanerSchools Whole School Change Program involved: introductory trainings, school administrator and personnel training, teacher trainings and learning groups, and follow-up supports. Representatives from each school district in the study attended an introductory training session in Spring 2014 (for either 1, 2, or 4-days) given by

### Table 2. Teacher characteristics (spring 2016)—in analytic sample.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Full analytic sample</th>
<th>Treatment</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience—as a teacher</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New (1 year or less)</td>
<td>683</td>
<td>12.5%</td>
<td>11.4%</td>
</tr>
<tr>
<td>Highly experienced (15+ years)</td>
<td>38.5%</td>
<td>36.2%</td>
<td>41.0%</td>
</tr>
<tr>
<td>Experience—in current school</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New (1 year or less)</td>
<td>681</td>
<td>32.3%</td>
<td>33.8%</td>
</tr>
<tr>
<td>Highly experienced (15+ years)</td>
<td>10.6%</td>
<td>9.5%</td>
<td>11.8%</td>
</tr>
<tr>
<td>Educational attainment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelor’s degree</td>
<td>683</td>
<td>33.7%</td>
<td>26.2%</td>
</tr>
<tr>
<td>Master’s degree</td>
<td></td>
<td>50.4%</td>
<td>57.0%</td>
</tr>
<tr>
<td>Specialist certificate/degree</td>
<td>683</td>
<td>13.3%</td>
<td>14.3%</td>
</tr>
<tr>
<td>Doctorate</td>
<td></td>
<td>2.6%</td>
<td>2.6%</td>
</tr>
<tr>
<td>Certification (within state)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regular</td>
<td>669</td>
<td>78.2%</td>
<td>75.0%</td>
</tr>
<tr>
<td>Advanced</td>
<td></td>
<td>9.0%</td>
<td>10.8%</td>
</tr>
<tr>
<td>Probation/Temporary/Other</td>
<td></td>
<td>12.9%</td>
<td>14.2%</td>
</tr>
<tr>
<td>Teacher—job characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-time (1 = full-time)</td>
<td>686</td>
<td>89.2%</td>
<td>85.0%</td>
</tr>
<tr>
<td>Leader (1 = leader)</td>
<td>686</td>
<td>10.5%</td>
<td>10.8%</td>
</tr>
<tr>
<td>Subject</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Math</td>
<td>671</td>
<td>19.9%</td>
<td>18.8%</td>
</tr>
<tr>
<td>English</td>
<td></td>
<td>26.9%</td>
<td>25.2%</td>
</tr>
<tr>
<td>Social studies</td>
<td></td>
<td>13.0%</td>
<td>13.1%</td>
</tr>
<tr>
<td>Science</td>
<td></td>
<td>12.0%</td>
<td>11.7%</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td>28.2%</td>
<td>31.2%</td>
</tr>
<tr>
<td>Grade level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle school (6–8th)</td>
<td>667</td>
<td>72.7%</td>
<td>72.7%</td>
</tr>
<tr>
<td>High school (9–12th)</td>
<td></td>
<td>27.3%</td>
<td>27.3%</td>
</tr>
</tbody>
</table>

*Note. The n shows how many cases were observed for each variable in dataset before imputation.*
IIRP. Intervention staff (from the already in place DN program) and school administrators attended additional trainings on the IIRP model (Introduction to Restorative Practices, Using Circles Effectively, and Facilitating Restorative Conferences). Teachers and staff attended up to four trainings (professional development days) from IIRP and participated in a “start up session” to organize staff professional learning groups (that met twice monthly check in on implementation). Schools and teachers received additional implementation support via up to four days of on-site consultations from IIRP staff, monthly calls with school leadership, and IIRP developed materials (manuals, books, and videos). The control group, defined as “business as usual,” schools were able to implement other school improvement approaches and were given a modest stipend for participating in the data collection. Activities and programs in both treatment and control schools were monitored through surveys of school-based personnel. (See Grant, 2020, for a more detailed implementation study.)

Implementation proved difficult in some schools in this study, with variation in the number of RP training sessions held and program supports received among treatment schools. Of the 17 schools randomized to implement RP w/DN, five schools dropped out of full program supports from IIRP after Year 1 (but these schools did not leave the study: they responded to the 2016 survey and are represented in the analytic sample). Anecdotally, these schools cited local supports as sufficient to maintain their RP program or indicated a change in leadership had led to new priorities for professional development in Year 2. Of the other 12 treatment schools, IIRP rated seven schools as having reached just a baseline level of implementation at the end of Year 1 and targeted these schools for more intensive supports in Year 2.

Real-world implementation challenges affect an intervention’s ability to have the full impact that it might otherwise have had under the extremely rare condition of ideal implementation (O’Donnell, 2008). But, it is essential to know “what works” in the real-world of less than ideal implementation in the underfunded and overburdened schools that populate large urban districts. Thus, we provide an intent-to-treat analysis of the causal impact that assignment to treatment had on school climate and teachers’ turnover intentions. That is, the analysis compares the mean outcomes in the RP w/DN schools with the mean outcomes in the control schools, ignoring the variation in the level of implementation achieved in the various RP w/DN schools. Thus, the analysis provides conservative real-world estimates of the causal impacts of the treatment—as the treatment was actually implemented in this sample—even though implementation in some schools was weaker than program leaders may have hoped.

**Data Collection**

At all schools in the original DN study, teachers, students, principals, and assistant principals were invited to take a paper survey each Spring, from 2012 to 2016. DN staff also completed surveys reporting on the practices being implemented in the treatment schools. Consent procedures (active or passive) varied by district according to each district’s standards. No individually identifying information was collected which helped assure confidentiality. Surveys asked about the extent of use of promoted practices from the DN and RP interventions, school-climate related issues, and teaching practices. Survey measures for the original DN study were developed by the external evaluator in collaboration with the DN
RP-specific survey measures were developed collaboratively by the DN and RP developers. This analysis focuses on the Spring 2016 survey results for teachers and students because full training and implementation of the IIRP model of RP, which began in Fall 2014, requires two school years to complete. Surveys were conducted with students and teachers from grades 6, 9, and 12 (the focal grades of the DN intervention, together with the 12th graders from the original DN cohort). On average, 73% of teachers and 64% of students (in blocks included in the analytic sample) responded to the 2016 survey.

**Measures**

**School Climate**
Our measures of school climate come from 16 teacher survey items and 31 student survey items to provide multiple sources of evidence around this key construct. Teachers’ and students’ reports were examined at the individual level. Our measurement approach for school climate, using a factor analysis, is described in more detail below. In brief, we modeled school climate as a composite of all the items by each respondent (i.e., teachers’ school climate rating overall and students’ school climate rating overall) and as three separate factors. The composites and factors all showed high internal consistency, Cronbach’s α = .88 for all 16 teacher-reported items (scales .83–.89) and Cronbach’s α = .87 for all 31 student-reported items (scales .79–.91).

**Turnover Intentions**
Teachers self-reported their intentions for the coming school year in response to the following question: “Which best describes your future intentions for your professional career?” Teachers’ selections were coded in two ways (based upon customary practices in the teacher turnover literature). First, teachers’ responses were coded into a binary variable of intending to stay at their school: stayers = 1 (“remain in this school”) and school leavers = 0 (all options involving a departure from the school, including transferring schools, districts, sectors, or leaving teaching). Intending to stay at the school was pre-specified as the primary turnover-related outcome in this study. Although these measures of intention are not the same as actual turnover, previous studies have found strong relationships between employee’s job intentions and their satisfaction, commitment, and actual turnover (155 studies reviewed by Tett & Meyer, 2006; Ladd (2011) observed this relationship among teachers in particular. Turnover intentions can also potentially capture teachers who are disengaged and seeking to leave but remain at their jobs due to external constraints (e.g., micro and macro-economic forces). A second turnover-related outcome was pre-specified as an exploratory outcome. For this exploratory outcome, the same responses were coded into another binary variable of intending to stay in the profession: profession leavers = 0 (one option involving a departure from the teaching) and stayers = 1 (all other options).

---

4Many survey items came from previously conducted school surveys by the DN developer that had been used in internal evaluation analyses.
5In middle schools without grade 6, grade 7 was the target group. The student survey sample included an additional block of two schools (for a total of 27 schools) for which data were available.
6A listing of these items is provided in the supplementary materials.
Perceptions of Problematic Teacher Absenteeism

Teachers responded to the question, “To what extent was teacher absenteeism a problem in your school during the 2015–2016 school year?”, on a scale from 1 (not a problem) to 4 (serious problem). This outcome was included as an ancillary outcome to teacher turnover as an imperfect indicator of how many teachers in the school already have “one foot out the door” and/or are so stressed by the school environment, working conditions, students, and leadership that they are having to take sick days and/or “mental health” days (Mowday et al., 1982). Greater (perceived) teacher absenteeism is hypothetically linked with greater teacher disengagement and lower satisfaction.

Assignment to Restorative Practices

Treatment assignment was coded as 1 for teachers in schools assigned to implement Restorative Practices (RP) with Diplomas Now (DN). Teachers at control schools that implemented the “business as usual” condition (did not implement DN or RP) were coded as 0 for this variable.

Adjustment Variables

Although schools were randomly assigned to treatment status, we tested adjustments for some pretreatment demographics to potentially provide more precise estimates. We tested the following groups of adjustment variables for inclusion: teacher characteristics (experience in teaching at the school, certification status, education level, full-time status, and subject taught), school grade level (middle or high school), and school composition (student enrollment, proportion of students enrolled in free and reduced-price lunch programs, and proportion of students from a minority ethnic or racial background).  

Analytic Plan

The analysis was divided into four main steps: analytic setup and sample description, school climate factor analysis, intent-to-treat analysis on perceptions of school climate (RQ 1), and an intent-to-treat analysis on teachers’ turnover intentions (RQ 2) and perceptions of problematic teacher absenteeism (RQ 3).

1) Analytic Setup and Sample Description. We first assessed the internal validity of the experiment by testing statistical assumptions necessary for the final models, including: normality, multicollinearity, and missing data patterns. To assess the integrity of the experimental design we also assessed the evidence for any threats to the internal validity of the experiment. Because this intervention was implemented at the school level and the students and teachers at the schools were anonymous in each year of the study, we cannot investigate attrition of individuals. Table A2 provides a summary of the number of schools in the sample and who responded to the Spring 2016 survey. Additionally, we assessed baseline equivalence in available background variables between  

7The small number of schools in the sample (limited power) and goal of parsimony suggest a model that does not include all of the adjustment variables.
treatment and control schools to test whether randomization resulted in balanced groups (using data from Spring 2011 and 2012, prior to DN or RP implementation).

2) School Climate Factor Analysis. We sought to identify the number of unique factors, or underlying constructs, present among school climate related items from the teacher and student surveys in order to better understand the unique impact of the intervention on different elements of school climate. We conducted two exploratory factor analyses on questions surrounding the components of school climate: the first on the 16 teacher-reported items and the second on the 31 student-reported items. On their constituent surveys, teachers and students responded to how much they agreed with statements describing their school or how prevalent certain practices were at their school. Scales varied from 4 to 7 points, so we standardized responses, allowing the responses to be compared according to their z-score.

We conducted an exploratory factor analysis of five steps for teachers and then for students (Williams et al., 2010). First, we assessed the characteristics of the data (i.e., distributions) and confirmed its suitability for a factor analysis (e.g., correlations, the Kaiser–Meyer–Olkin [KMO] measure of sampling adequacy, and Bartlett’s test of sphericity). Second, we extracted the factors using principal axis factoring (PAF). Third, we determined the number of factors based on the eigenvalues (approaching or greater than one), scree plot, and a parallel analysis (Fabrigar et al., 1999). Fourth, we applied an oblique rotation because the factors were correlated (Ford et al., 1986). We also iteratively tested alternate numbers of factors before selecting the final three factor solution and reporting the results described further in the Results section (and found in Supplemental Tables A3 and A4). We then used responses on the previous year (2015) teacher and student surveys for confirmatory factor analyses (and to estimate model fit indices).

3) Intent-to-Treat Analysis on School Climate. We conducted a two-level multilevel model incorporating block fixed effects for the intent-to-treat analysis, using STATA 14.0 (commands mixed). We theoretically selected a multilevel modeling approach because the data are from teachers clustered within schools and our research questions ask about a school-level intervention (Raudenbush & Bryk, 2002). We confirmed that cluster effects exist within the observed data by examining the intraclass correlation coefficients (ICC) of the analytic variables. We also examined whether the cluster sizes were sufficient for a multilevel analysis.8

We provide the equations for the continuous school climate outcome below. Equation (1) describes the level 1 model for the relationship between perceptions of school climate and teachers’ individual characteristics. Equation (2) describes the level 2 model for the relationship between school-level school climate and their block, treatment status, and student composition. Finally, Equation (3) describes the composite model where model (2) is substituted into model (1) to get one equation which models perceptions of school climate in terms of all the analytic variables. We estimated all regression models using robust standard errors to account for heteroscedasticity. No adjustments were used to account for multiple testing because only one primary

---

8For instance, we could not pursue a three-level model to analyze the clustering within blocks or districts because several of the blocks or districts contain only two participating schools, which is too small for accurate estimation of a model of that type.
outcome was selected for each domain of outcomes.\(^9\)

\[
y_{ij} = \beta_{0j} + \sum \beta_{1i}x_{ij} + e_{ij} \quad (1)
\]

\[
\beta_{0j} = \gamma_{00} + \gamma_{01}T_j + \sum \gamma_{02k}D_k + \sum \gamma_{03j}W_j + u_{0j} \quad (2)
\]

\[
y_{ij} = \gamma_{00} + \sum \beta_{1i}x_{ij} + \gamma_{01}T_j + \sum \gamma_{02k}D_k + \sum \gamma_{03j}W_j + u_{0j} + e_{ij} \quad (3)
\]

In these equations, subscripts refer to teacher \(i\), in school \(j\), which is in block \(k\). Thus, \(y_{ij}\) represents the school climate rating for teacher \(i\), in school \(j\). In Equation (1), \(X_{ij}\) are the set of teacher level covariates for teacher \(i\) in school \(j\). Equation (2) models \(\beta_{0j}\), each school’s average school climate rating. \(T_j\) is the binary indicator for school \(j\) that indicates if the school is in the treatment or control experimental condition. \(D_k\) are the set of binary indicators for the block \(k\) in which a teacher’s school resides. The coefficients can then be interpreted thus: \(\gamma_{00}\) is the school level school climate rating in the control schools in the first block; \(\gamma_{01}\) is the average treatment effect (ATE) on school climate, and \(\gamma_{02k}\) are the effects on school climate of being in each block \(k\). \(W_j\) are the set of school level covariates for each school \(j\) that adjust for potential differences in schools’ student composition. The residual variation is divided in two parts: \(e_{ij}\) is the individual, teacher level random effect (for teacher \(i\) in school \(j\)) and \(u_{0j}\) is the school level random effect for each school \(j\). Both random effects are assumed to have a mean of 0 and variance \(\sigma^2\) (representing the variability within schools) and \(\tau_{00}\) (representing the variability between schools) respectively.

In each set of models, we tested the whole set of adjustment variables described above to help improve the precision of our estimates of the treatment effect. To retain a parsimonious model (and due to the limited sample size) we only retained the covariates in each model which had a significant effect on the outcome. We incorporated block fixed effects to account for the experimental design (random assignment to treatment within blocks). We calculated effect sizes using Glass’s delta, which presents the effect in terms of the standard deviation (SD) units of the control group (Glass et al., 1981; Rosenthal & Rubin, 1986): Glass’s \(\Delta = \frac{\text{Difference}_{\text{weighted}}}{\text{SD}_{\text{control}}}\). We chose this effect size measure due to its ease in interpretation and more importantly because we are comparing the observed effect to the variation observed in the counterfactual, or the situation without this specific intervention.

4) Intent-to-Treat Analysis on turnover intentions. Similar to step two, we regressed assignment to restorative practices on teachers’ reported turnover intentions in order to test whether assignment to treatment had an effect on teachers’ turnover intentions. This regression is similar to the one performed in the previous step but incorporates a logarithmic link function (commands \text{logit} and \text{melogit}) to predict the binary turnover outcomes (e.g., stay in current school = 1 or leave current school = 0). Equation (4) below presents the equation for this model.

\[
\log\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right) = \gamma_{00} + \sum \beta_{1i}x_{ij} + \gamma_{01}T_j + \sum \gamma_{02k}D_k + \sum \gamma_{03j}W_j + u_{0j} \quad (4)
\]

\(\pi_{ij}\) is the proportion of teachers reporting intentions to leave their current school. \(\frac{\pi_0}{1 - \pi_0}\) therefore equals the odds that a teacher will intend to leave their current school. (A similar equation was used to predict the proportion of teachers intending to remain in

\(^9\)We also used the Benjamini-Hochberg procedure to confirm the validity of our tests of all outcomes (Thissen et al., 2002).
The logarithmic linking function is used to account for the non-normal distribution of the binary outcome, which presents the results in terms of log odds. (In reporting the results we translated log odds into the more easily understandable odds ratio.) Additionally, regressions using the logarithmic link function have no individual error term because they predict the actual observed ratio of teachers reporting intentions to leave compared to those who did not (Agresti, 2002).

**Internal Validity Assessments, Analytic Setup, and Model Diagnostics**

**Baseline Equivalency Analysis**
Prior to the main analysis, we first tested the internal validity of the experiment to justify the intent-to-treat analyses that follows. Table 1 compares treatment and control schools in the analytic sample and illustrates our tests of the comparability of schools in the treatment and control groups at baseline, before randomization to condition occurred (in 2010–2011 or 2011–2012, depending upon the block’s randomization date). Tests of differences between the experimental groups in enrollment, percent free and reduced-price lunch, percent minority, percent special education, attendance, and exclusionary discipline days, revealed no significant differences ($p < .05$). Thus, at baseline, before implementation of DN or RP, randomization was effective in ensuring intervention and control groups were well balanced. Baseline data on the survey outcome measures (school climate and teacher turnover intentions) were unfortunately not available.

**Differential Attrition (and Non-Response) Analysis**
We analyzed attrition at the school level, the level of assignment. Six schools did not respond to the Spring 2016 survey and thus are not present in the analytic sample, representing an overall attrition rate of 18.18% (6/33). There was much greater attrition from the control group (25% versus 11.76% in treatment). More relevant for the analysis, we only included schools in blocks where there was at least one treatment and control school because modeling the effects within blocks was essential to the original experimental design and thus to our analysis. Although this results in loss of sample size (and power), it reduces the potential for bias from potentially systematic non-response in certain blocks. In the analytic sample, two more treatment schools are dropped (since the corresponding control school did not respond) for an overall attrition rate of 24.24% (8/33). Attrition was much more comparable with these exclusions: 23.5% (4/17 schools) in treatment and 25% (4/16 schools) in control, for a differential attrition rate of 1.47%. These attrition levels qualify as low attrition according to What Works Clearinghouse standards (Institute of Education Sciences [IES], 2014).

At the same time, the smaller analytic sample of schools raises issues of external validity. A comparison of schools in the full sub-study and schools in the analytic sample indicated that high schools were less likely than middle schools to participate in the survey data collection. Schools that remained in the analytic sample were significantly smaller and had higher rates of students eligible for free/reduced price lunch (factors associated with differences between middle and high schools). It may not be possible to extend generalizations about the impact of RP w/DN very widely.
It can be easily assumed that both teachers and students entered and exited schools in the study over the years these schools were involved in the study. The study did not track individuals and therefore it is difficult to calculate the exact number of “joiners” (post assignment) or to estimate the risk of bias from these joiners (at the student or teacher level). Joiners would only pose a risk if they knew about the intervention and selected into schools because of the intervention. Although a whole school turnaround intervention is high profile, all of the schools in the sample were labeled as in need of turnaround, facing closures, and were therefore likely to be implementing other types of school reform interventions. This assumption is supported by the fact that comparable numbers of teachers in treatment (67.47%) and control (72.96%) reported receiving training on positive behavioral supports (another whole school intervention). Therefore, the treatment of focus in this study may not have stood out enough to differentially draw in new teachers and students.

Another risk to the integrity of assignment would be systematic non-response. If teachers who did not respond to the survey differed systematically from those who did respond, the results could be biased. We did not have access to data that could address this issue.

Variable Analysis
Next, we examined the distributions of main variables for skewed (non-normal) distributions, outliers, multicollinearity, and adequate variability but found no substantial variation from normality (in continuous variables). Additionally, all scales exhibited high internal consistency (Cronbach’s alphas .87 and .94) suggesting the items in each scale had a high consistency with one another. We also assessed covariation between the main variables of interest prior to conducting multivariate analyses.

Describing Schools, Teachers, and Missingness in the Analytic Sample
Tables 1 and 2 describe the characteristics of teacher and schools in the analytic sample. At the school level, all demographic variables were complete. A small number of cases (7%) were missing information on one of the outcome variables and 6% of cases were missing at least one of teacher level covariates. No patterns of missingness were detected via models predicting the likelihood of missing and we assume the values are missing at random.

Results

School Climate Factor Analysis
The exploratory factor analyses results suggested a three-factor solution in both the teacher and student reports, mirroring the grouping of the items as they appeared in the original surveys. For the teacher items, model fit from a confirmatory factor analysis was adequate (Browne & Cudeck, 1993): $\chi^2(149, N=1,288) = 1,156.56, p < .001$.

10The teacher survey did include a self-report measure of whether the respondent was new to the school within the last year. Analysis of survey results did not find a significant difference between treatment and control schools.
RMSEA = .072, CFI = .919. Table A3 shows the factor loadings for the three-factor teacher survey solution, with all items loading at high levels above .6, and with no cross-loadings above .3. The three factors can be described as: supportive environment (3 items, the highest loading item being “There was an atmosphere of trust and mutual respect within the school”), professional learning and collaboration (6 items, “You met with non-teaching professionals… to identify at-risk students and/or plan interventions for those students”), and problematic behaviors (7 items, “To what extent was … students fighting … a problem in your school”).

The student survey items model fit from a confirmatory factor analysis was also adequate (Browne & Cudeck, 1993): $\chi^2(419, N = 5,989) = 6,068.19, p < .001$, RMSEA = .047, CFI = .929. Table A4 shows the factor loadings for the three-factor student survey solution, with most of the items loading at adequate levels above .5, and with no cross-loadings above .3. The accepted solution had a three factor structure: supportive social environment (15 items, “My teachers really listen to what I have to say” as the highest loading item), responsive instructional practices (6 items, “teachers asked me to use critical thinking and reasoning to complete tasks or solve problems”), and problematic behaviors (10 items, similar to the teacher scale of the same name, “How much of a problem is … students bringing weapons like knives and guns to school”).

Additionally, we compared student and teacher responses (school-level averages) on similar items related to school climate. There was a significant correlation ($r = .49$) between the composite school climate factors based on student and teacher ratings. On individual items about the prevalence of school problems, which were identical on both surveys, there were strong correlations ($r = .52–.71$). There was also a moderate correlation between individual teachers’ ratings and the school average rating of their colleagues ($r = .43$). There was almost no correlation between teacher reports of professional learning and collaboration and student reports of responsive instructional practices ($r = -.02$) and there was a positive correlation between teacher and student reports of problematic behaviors ($r = .41$).

**Intent-to-Treat Analysis on School Climate**

Prior to estimating the hypothesized multilevel models for the main analysis, we examined the partitioning of the variance (ICC or intraclass correlation) in each outcome according to each cluster level: school or block. Within the composite school climate variable, as perceived by teachers, 28.4% of the variation resides between individual teachers in the same schools, and 14.2% resides between teachers in the same blocks. Among students, 5.3% of the variation in the outcome resides between students in the same schools, and 8.2% between students in the same blocks. The moderate ICC values among teachers suggest a hierarchical modeling approach would be most appropriate, to account for the nonrandom distribution of the variance. Although the values are lower for students, we also employed a hierarchical model for this outcome to avoid potential Type I error which could still be present at even these low ICC values (Musca et al., 2011).

Table 3 shows the results from final models estimating the effect of assignment to treatment (RP w/DN) on teachers’ and students’ perceptions of school climate overall.
These models, accounting for significant covariates and blocks with dummy variables, were chosen as having the best model fit using the likelihood ratio test (\( p < .01 \); this test compares the explanatory power, model log likelihood, to the prior model). Estimates of the treatment effect were generally consistent across multilevel model specifications. Teachers in schools assigned to RP w/DN reported more positive perceptions of school climate that were .17 of a z-score higher on the average item in the overall composite, compared to teachers in control schools (\( p < .10, ES = .27 \)), adjusting for teachers’ fulltime classroom status, the only significant covariate. Similar patterns emerged among the secondary outcome: students’ reports about school climate. The final model suggests that a student in an RP w/DN school reported that their school’s climate overall is .07 z-score units more positive on the average item than a student in a control school (\( p < .05, ES = .20 \)), adjusting for students’ age, grade, and Hispanic race/ethnicity. These effect sizes are relatively small and suggest that the intervention slightly improved perceptions of school climate.

Additionally, we examined the variation in treatment effect by school climate factors. Table 3 shows the impact of assignment to RP w/DN on outcomes (by respondent).

<table>
<thead>
<tr>
<th>School climate outcomes</th>
<th>Adjusted mean Impact estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment</td>
</tr>
<tr>
<td>Teacher-reported SC composite(^a)</td>
<td>-.05</td>
</tr>
<tr>
<td>Subscales</td>
<td></td>
</tr>
<tr>
<td>Supportive environment(^a,c)</td>
<td>.06</td>
</tr>
<tr>
<td>Professional learning and collaboration(^a,c)</td>
<td>.05</td>
</tr>
<tr>
<td>(Fewer) problematic behaviors(^a)</td>
<td>-.12</td>
</tr>
<tr>
<td>Student-reported SC composite(^a,e,g)</td>
<td>-.22</td>
</tr>
<tr>
<td>Subscales</td>
<td></td>
</tr>
<tr>
<td>Supportive social environment(^a,f)</td>
<td>.12</td>
</tr>
<tr>
<td>Responsive instructional practices(^f)</td>
<td>.07</td>
</tr>
<tr>
<td>(Fewer) problematic behaviors(^e,g)</td>
<td>.18</td>
</tr>
<tr>
<td>Turnover-related outcomes</td>
<td>OR or ES</td>
</tr>
<tr>
<td>At their school(^1,a)</td>
<td>.81</td>
</tr>
<tr>
<td>In the profession(^1,a,b)</td>
<td>1.89</td>
</tr>
<tr>
<td>Problematic teacher absenteeism(^a)</td>
<td>2.12</td>
</tr>
</tbody>
</table>

Note. Primary outcomes are bolded. ATE: average treatment effect; ES: effect size; OR: odds ratio; SC: school climate. School climate scales were created from standardized items to account for the different scales across items. Robust standard errors are shown in parentheses. Results from these final multilevel models include treatment effect, blocks, and model specific significant covariates (fulltime status\(^a\), experienced teacher status\(^b\), and proportion of students from racial/ethnic minoritized backgrounds\(^c\) for teacher-reported outcomes; students’ age\(^d\), grade\(^e\), sex\(^f\), and Hispanic race/ethnicity\(^g\) for student-reported outcomes). Adjusted means were calculated for the average teacher or student in each condition, adjusting for the proportion of the sample on each covariate and the coefficient for each covariate.

\(^1\)For these binary outcomes, treatment effects are expressed in log odds and odds ratios.

\(^\dagger\)\( p < .1; ^* p < .05; ^{**}p < .01\). All significance tests are two-tailed.
effects found on the supportive social environment \( (p < .10, ES = .15) \) and fewer problematic behaviors \( (p < .05, ES = .12) \) subscales.\(^{11}\)

**Intent-to-Treat Analysis on Teachers’ Turnover Intentions**

As with the school climate outcomes, we first examined the partitioning of the variance (ICC or intraclass correlation) in each outcome according to each cluster level: school or block. Small ICC values (proportions of variation) at the cluster level indicated that most of the variation in the turnover outcomes resides between individual teachers in the same school or block rather than due to clustering effects: only 3.1% of the variation in school turnover intentions is due to school level variation.

Table 3 shows the results of the models estimating the effect of assignment to RP w/DN on teachers’ turnover intentions. The effects of assignment to RP w/DN on teachers’ intentions to stay at their school and in the profession were negative but nonsignificant, consistently across model specifications, indicating that RP w/DN assignment may be associated with reductions in teachers’ likelihood of staying in their position and the profession. These small, insignificant impacts on teachers’ intentions suggest that RP w/DN program did NOT have the hoped-for effect of encouraging teachers in high-turnover high-poverty urban schools to make a commitment to stay for an extended tour of duty.\(^{12}\)

However, there is evidence that RP w/DN might encourage teachers to be “present and accounted for” at school more often during the current school year. Specifically, we tested the effect of being in an RP w/DN school on the extent to which teachers perceived teacher absenteeism to be a problem at their school. In the multilevel model incorporating block fixed effects, teachers in RP w/DN schools were less likely to perceive teacher absenteeism among their peers to be a problem at their school than did teachers in control schools \( (p < .05, ES = −0.37) \).

**Discussion**

The potential of RP w/DN to improve school climate is an important finding of this study. While the magnitude of the effect sizes of the RP w/DN intervention on teacher and student perceptions of school climate are relatively modest, they are large enough to suggest that the treatment made a noticeable difference. Specifically, the significantly more positive perceptions of school climate at RP w/DN schools implies that the intervention successfully changed student behavior and cultivated a more supportive social environment. Based on survey responses, treatment schools had more positive student interactions and fewer conflicts such as fights and bullying, which are particularly salient for students’ enjoyment of school and feelings of safety and belonging there.

As we suggest in Figure 1, these findings are consistent with the hypothesized theory of Restorative Justice and RP for schools, which posits that the use of this whole school reform can creating a more positive, welcoming environment through improved

\(^{11}\)Greater statistical significance among student school climate factors with smaller magnitude of effects is likely due to the larger sample size and thus, greater power.

\(^{12}\)We did not pursue a full mediation model (e.g., Structural Equation Modeling) as the effect on teachers’ turnover intentions was not statistically significant.
interactions and relationships (Mirsky, 2007), and with the expected positive effect of DN’s small learning communities on school climate (Levine, 2010). This finding of the positive effects of RP w/DN on school climate is also consistent with prior research from many observational and case studies (e.g., Anyon, 2016; Jain et al., 2014; Lewis, 2009). This finding also builds upon the positive effects found in the main study of this randomized control trial, where researchers found that RP w/DN reduced students’ incidence of suspensions and chronic absenteeism (Grant et al., 2020).

In contrast to the study from Pittsburgh, which found positive effects of RP on teachers’ perceptions of school climate but not on students’ perceptions, we found that RP w/DN had a positive impact on students’ perceptions of school climate but that its impact on teachers’ perceptions did not meet conventional standards of significance. In Pittsburgh, the positive effect of RP on teacher perceptions of school climate was much stronger, and students in Pittsburgh’s RP schools reported a more negative school climate (on their Tripod composite score; ES = .19). In this study, we found that students in RP w/DN schools reported a more positive school climate overall (ES = .15). Further, in the current study, the estimated impact was positive for all three student subscales (with one statistically significant), where the estimated impacts for all 12 student subscales in the Pittsburgh study had negative effects (with 9 statistically significant).

The difference in findings between this study and the one in Pittsburgh could be due not only to differences between an RP and an RP w/DN intervention, but also to the different samples in location (one city, Pittsburgh, versus many cities across the U.S.) or grade level (most of the Pittsburgh study were elementary schools, in contrast to the middle and high schools in this sample). Potentially older students may be able to understand the new and complex concepts behind restorative justice and thus fully participate in RP as originally designed. For example, adolescent age students may be more likely to engage in deeper reflection in response to restorative questions, reflecting on their behavior, its consequences, and appropriate restorative solutions (Schumacher, 2014). In a similar vein, Ingraham et al. (2016) found in one school that successful RP implementation involved adaptation of RP materials and concepts to the elementary developmental levels. However, other studies have found more implementation difficulties in secondary schools (e.g., McCluskey et al., 2008), so the relationship between school level and RP impact requires further investigation.

Regarding the second research question about the impact of the intervention on teachers’ intentions to remain at their school and in the teaching profession, this study had no statistically significant findings. This finding aligns with the in-depth qualitative work of Johnson (2020) who found that specific approaches to discipline did not relate with turnover. Although the intervention had a positive effect on perceptions of school climate, it is possible that improvements in school climate were too small to reduce teacher turnover intentions as hypothesized in our theory of change.

On the other hand, while the impact estimates were not significant, it may be important to acknowledge that they were consistently negative in direction (across model specifications and intentions types). Findings regarding the impact of the intervention on intentions to leave the profession were similar to those on intentions to leave the school (confirmed by the lack of a multinomial trend in the data per the Hausman test). Even though the findings on turnover intention were not significant in the current study, it is possible that...
the experience of multiple reform efforts (such as RP on top of DN) could lead to teacher burnout and higher rates of turnover intention. Because it is common for low-performing schools to be implementing multiple programs simultaneously, it will be important in future studies of RP to take this issue of simultaneous school reforms into consideration.

Interestingly, we found that teachers in RP w/DN schools reported that teacher absenteeism was less of a problem, compared to their peers in control schools. This exploratory finding raises questions about the complex relationship between teacher absenteeism and teachers’ turnover intentions, which has been largely underexplored. One potential explanation for this pattern of findings could be teachers’ experiences of moral injury (and related psychosocial feelings; Levinson, 2015). Teachers working in this set of schools are likely to encounter many contextual and societal injustices, such as the high prevalence of trauma, poverty, and racial segregation. Although RP is set up to help students and teachers become critical of injustice, and teachers may appreciate working in schools with RP, they may also experience an increased moral sensibility, or awareness of these injustices. Teachers facing the immovable injustices mentioned above can then feel a moral injury and powerlessness, which has been shown to predict greater intentions to leave (Sugrue, 2020). Future qualitative studies might further investigate the moral burden placed on teachers and how they cope with it.

The nonsignificant but potentially negative impacts on turnover intentions may also partially reflect the limited time frame of the study: the first two years of RP implementation. Higher turnover intentions during this initial implementation period may reflect the resistance of teachers to a large change in the approach of the school or their resistance to the counter-cultural mindset change of Restorative Justice compared to traditional punitive approaches (Vaandering, 2014). Potentially, teacher turnover (and intentions to turnover) may level off after these initial years when teachers who disagree with the reform have left. Additionally, as teachers witness the positive impacts of RP w/DN (such as improving the school climate) they may be more willing to buy in to this large mindset change and be more likely to stay. Future studies could look into the dynamic relationship between RP implementation and teachers’ satisfaction and turnover over time to examine whether this hypothesis holds.

Finally, these findings add to the nascent literature on the impact of whole school reform on teacher turnover and are some of the first published findings about the effect of reforms that include RP on teachers’ job attitudes and turnover. Most whole school reforms acknowledge the importance of teachers in instruction and implementation but less often evaluate how these reforms impact teachers and the teacher workforce in schools. Given the complexity of teacher turnover, and the competing explanations about how reform initiatives may influence decisions of different types of teachers in different ways, it is crucial for more research to be conducted.

Future studies can seek quantitative and qualitative data from teachers evaluating the impact of whole school reforms on their satisfaction and retention. Regarding RP in particular, future studies could investigate teachers’ experiences of RP, including its psychological impacts and social validity, adding to the literature on teachers’ perceptions of its implementation and true “justice” orientation (Lustick, 2017a; Rainbolt et al., 2019). Additionally, the interesting pattern of findings between problematic absenteeism and turnover intentions call for future studies examining these aspects of teachers’
experiences. Teacher absenteeism is often perceived as a precursor of turnover and signal of burnout (Mowday et al., 1982) that is costly to school budgets and instruction (Bruno, 2002). Descriptively in this study we found only a weak correlation between school-wide problematic teacher absenteeism and individual teachers’ turnover intentions. More recent research is needed to identify how problematic absenteeism is among teachers in underserved schools and how school level and individual teachers’ absenteeism relates to teacher turnover and attempts at whole school reform.

**Limitations**

There are several limitations to this study that deserve extended comment. We first address the issue of external validity. Although the sub-study involved schools in eight large districts across the U.S., the districts and schools that agreed to participate were likely more favorably inclined toward RP than those that did not. This is, however, a limitation of most studies of large-scale reform efforts that require willingness of schools to participate and be randomized. At the same time, the sample includes schools from eight large urban districts across the country that all faced a diverse set of challenges, serve students in high poverty, and who are thus the most frequent target for reforms.

It was not possible to test whether the randomized design of the study succeeded in achieving baseline equivalence in school climate, teacher turnover intentions, and perceptions of problematic teacher absenteeism between the groups of schools, though there was evidence of baseline equivalence on school characteristics.

As with many long-term whole school randomized control trials, there were difficulties in following the entire sample through the extended years (years 4 and 5) which are the focus of this study. The reduced analytic sample, representing nine or ten of the twelve blocks depending upon the outcome measure, appears balanced, but is likely not representative of the original sample, potentially retaining only schools with more interest in reform or greater organizational capacity. Additionally, the reduction of the analytic sample produced power issues that may have limited our ability to detect statistically significant effects, particularly among the teacher sample and prevented us from a full investigation of the variation by site. Future replication in larger samples can address these potential issues.

There is the potential for selection bias in our impact analyses of teachers’ outcomes (due to teachers selecting in and out of schools during the study period) which we cannot adjust for or quantify in these anonymous teacher data. For example, it is possible that teachers who were unhappy with the reforms implemented in study schools—the RP w/DN reform program in treatment schools, and/or the variety of other reforms in control schools—left study schools prior to the time our outcome measures were collected in the final year of this extended study. If so, the nonsignificant impacts of RP w/DN on teachers’ turnover intentions reported here may be misleading. Nonetheless, our results are an illuminating snapshot of the current faculty’s turnover intentions at the end of the second year of RP implementation.

Another limitation involves the turnover intentions variable, which reflects teachers’ prospective plans and is not the same as actual turnover. Teachers’ final decisions to stay or leave (not available in these data) were undoubtedly impacted by the opportunities and constraints they encountered in the job market.
Finally, as we noted throughout, this study cannot separate the effects of RP from the effects of the combined RP w/DN intervention. The design of the RP w/DN sub-study does not permit any strong conclusions about the effects of adding RP to DN, compared to DN alone.

**Conclusion**

The evidence from this study of the positive impact on school climate of the combined Restorative Practices with Diplomas Now intervention should be useful as school leaders make decisions about interventions to improve their school environments and the outcomes for their students. In addition to the impact of the combined interventions on student disciplinary outcomes (Grant et al., 2020), the positive effect on school climate found in this study is important in its own right. Future studies that examine the impacts of implementing RP on its own (without another whole school reform occurring simultaneously) in a large sample of schools will be important for expanding the evidence base for this intervention. It is also important to pursue future investigations of the more complex relationship between RP and teacher attitudes and outcomes (including their career intentions and actual turnover decisions). Improving the school experience for both teachers and students is a critical measure of positive educational change. Preparing future citizens to resolve conflicts in ways that repair harm and restore relationships is a goal worthy of more focused research attention.

**Funding**

This research project was made possible by a grant from The Atlantic Philanthropies [23220, Evaluation of Restorative Practices, Robert Balfanz, Principal Investigator], matching funds from the PepsiCo Foundation, and a grant from USDOE’s Office of Innovation and Improvement [U396B100257: Evaluation Extension, Validating the Talent Development-Diplomas Now Secondary School Turnaround Model, Robert Balfanz, Principal Investigator]. The views expressed in this chapter are those of the authors and do not represent the views of the funders.

**ORCID**

Ashley A. Grant http://orcid.org/0000-0002-4647-723X

**References**


Corrin, W., Sepanik, S., Rosen, R., & Shane, A. (2016). Addressing early warning indicators: Interim impact findings from the Investing in Innovation (i3) evaluation of Diplomas Now. MDRC.


