UNDERSTANDING IMPLEMENTATION OF RESTORATIVE PRACTICES IN LOW INCOME, URBAN HIGH SCHOOLS

A DISSERTATION

SUBMITTED TO THE FACULTY

OF

THE GRADUATE SCHOOL OF APPLIED AND PROFESSIONAL PSYCHOLOGY

OF

RUTGERS,

THE STATE UNIVERSITY OF NEW JERSEY

BY

JOSHUA LAWRENCE KORTH

IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE

OF

DOCTOR OF PSYCHOLOGY

NEW BRUNSWICK, NEW JERSEY MAY 2016

APPROVED: _____
Anne Gregory, Ph.D.

_____
Susan Forman, Ph.D.

DEAN: _____
Stanley Messer, Ph.D.

Copyright 2016 by Joshua Lawrence Korth

Abstract

Whereas much research has led to widespread knowledge about the harmful effects of school disciplinary practices (e.g., suspensions, expulsions), there have been fewer studies about programs that promote alternative forms of school discipline. The current dissertation examined one such program, Restorative Practices (RP), in its early stages of implementation from the perspectives of teachers. RP uses collaborative problem solving and encourages relationship building between students and teachers. Key components of RP include community-building circles in classrooms and restorative conferences that bring together all those involved in a discipline incident to identify ways to repair the harm. Despite school districts investing substantial resources in upfront training days, little is known about the teacher perceptions and factors associated with high or low RP use implemented in the classroom. The dissertation was comprised of two studies. Study 1 included interviews and surveys with 51 teachers in two public high schools. Results showed that teachers varied in their fidelity of RP implementation. Further, multiple regression analyses found days of RP training and positive experiences with RP were predictive of a teacher's total number of RP circles and conferences run. In addition, teachers' RP self-efficacy was predictive of their adherence to implementing the core components of RP. Study 2 examined a cohort of teachers over two school years that continued to use RP ($N = 23$). Results showed that a group of teachers ($n = 7$) were classified as high RP users in both school years. These results held when accounting for minimal RP administrative support, which suggest that that the teachers were "resilient implementers." Implications for future practice are that positive experiences with RP, teachers' beliefs in their own ability to use RP, and upfront training days may be factors that increase teachers' implementation of RP with high fidelity—although future research will need to corroborate these findings. Moreover, as the evidence

and promise for RP continues to grow, more research is necessary to determine why teachers choose to adopt or not to adopt this innovation.

Acknowledgements

This dissertation is the product of many people and their support. I am so thankful for all the guidance I have received on this journey from both family and friends. I could not have done this alone.

The most influential person in this process has been my dissertation chair, Dr. Anne Gregory. From the moment I met her at my GSAPP interview, I knew I wanted to contribute to the important work she was doing. Her dedication to the completion of my dissertation and my growth as a psychologist has gone above and beyond what is necessary of a mentor. She is an exceptional individual who has always pushed me to produce the best work possible. She made the many long commutes to New Jersey from New York City (and back) both engaging and intellectually stimulating. I hope to emulate her warm and genuine approach in my own career. I also want to thank Dr. Susan Forman for her insight and willingness to help me conceptualize this dissertation. She runs a wonderful school psychology program that I am proud to be a part of, and I am equally fortunate to have her participate in my committee. I could not have finished this project without the help of a very dedicated undergraduate student, Aria Hurley. She helped me transcribe teacher interviews that were used for analyses. Then, there were my colleagues and friends that helped me code the teacher interviews this past summer. Gabrielle Roberts, Clarissa Green, and Marguerite Summer—you must have been heaven sent, and I am grateful that you all volunteered your time on this project.

Finally, I want to thank my family. Dad, you always made me feel so proud to be pursuing a doctoral degree, and mom, thank you for always bragging about me to your friends. You have both provided support in more ways than I can write in this section, and I am so lucky to be your son. Jason, you have been a wonderful brother too, and I would not be at this point without you. Finally, I have to thank my fiancée, William “Ted” Caine who took

a major risk by beginning a relationship with me in the first year of graduate school. You provided the motivation I needed on the tough days and shared my excitement when I reached professional and personal goals. Your love and encouragement helped me to reach this milestone. Thank you so much for always being there for me.

TABLE OF CONTENTS

PAGE

ABSTRACT……..….ii

ACKNOWLEDGEMENTS……....iv

LIST OF TABLES…….....viii

INTRODUCTION……...……....1

Urban School Settings…….....2

Teacher Challenges…….....5

Restorative Practices……..11

Implementation Science…….....16

Predictors of Early Implementation……...18

Summary…….....21

STUDY 1……...23

Methods……...…….....24

Participants……....24

Procedures…….....26

Survey Measures…….....27

Semi-Structured, Teacher Interview…….....28

Data Analysis Plan……....32

Results……...32

STUDY 2……...43

Methods……44

Participants……....44

Procedures……44

Measures……45

Data Analysis Plan…….45

Results……46

DISCUSSION……48

Summary……48

Building on Positive Experiences with RP……50

Learning from Positive Experiences……...51

RP Self-Efficacy……53

Administrative Support and the Concept of Resilient Implementers……....54

Study Strengths……..55

Study Limitations……...56

Implications for Future RP Studies and Practice……...57

APPENDICES……...60

REFERENCES……..78

LIST OF TABLES

Table 1 Defining the Eleven RP Essential Elements……...13

Table 2 Descriptive Analysis of Number of Circles and Conferences……..35

Table 3 Use of RP Essential Elements……...35

Table 4 Adherence to Six RP Essential Elements……...35

Table 5 Descriptive Analysis of Teacher Perceptions of RP…….....36

Table 6 Associations Among Teacher Characteristics, Teacher Reports, and RP Use……...41

Table 7 Multiple Regression Results Predicting Total RP Circles/Conferences…….43

Table 8 Multiple Regression Results Predicting Teacher RP Implementation……...…43

Table 9 Paired Samples *T*-Test of Teacher RP Fidelity and Perceptions at Year One and Year Two……..47

Table 10 Classifying Teacher RP Use by Their Total Circles/Conferences Implemented……47

Introduction

About 50% of teachers leave the profession in their first three-to-five years of teaching in urban settings (Elias et al., 2003; Eslinger, 2014; Ingersoll & Smith, 2003). In fact, teachers who exit the profession after only a few years most often report that student misconduct was a major factor in their decision to leave (Borman & Dowling, 2008). As high teacher attrition rates continue to occur in stressed urban settings, so do attrition rates of at-risk students via the school-to-prison pipeline (Kim, 2003; Skiba et al., 2014). Skiba et al. (2014) define the school-to-prison pipeline as a mechanism through which school discipline practices, especially exclusionary actions (e.g., detentions, suspensions, expulsions), lead to decreased probabilities of success and increased probability of student involvement in the juvenile justice system. At first glance, this stark contrast of scenarios for teachers and students leaving urban school settings may seem unrelated. However, studies have shown how the positive relationships between teachers and students can improve student achievement and behavioral engagement in the classroom (Gregory & Korth, in press; Gregory & Ripski, 2008). Conversely, these relationships are negatively impacted through punitive school-wide discipline approaches (APA Zero Tolerance Task Force, 2008; Gregory & Ripski, 2008). While much is known about the deleterious effects of zero tolerance school discipline policies (APA Zero Tolerance Task Force, 2008), little is known about programs that restore relationships and aim to transform disciplinary actions in schools. More specifically, as schools across the nation undergo discipline reform, little is known about why some teachers, especially in low income, urban schools, would choose whether to alter their classroom discipline practice or not when introduced to new approaches within the same schools. In other words, the current studies examine variability in program implementation.

Furthermore, the current studies address gaps in our knowledge through teacher interviews and self-report surveys related to the uptake of new discipline practices delivered

through a program targeting alternative approaches to discipline, restorative practices (RP). To date, very few studies have examined RP in highly stressed, under resourced schools. These schools are often comprised of predominantly low income African-American and Latino students. So often, research on program implementation does not adequately describe the context in which the research is embedded. In order to contextualize how the current program, RP, fits with implementation research, it is useful to review in some detail the stressors and challenges facing many students, teachers, and schools residing in low income, urban neighborhoods. As such, the following review addresses challenges of inner city environments, teacher challenges within schools, and how zero tolerance policies set the tone for school discipline. The review then considers literature related to program implementation and characteristics of teachers who readily adopt new initiatives. Then, the review concludes with the identification of gaps in knowledge about program implementation, specifically related to urban teachers' adoption of RP in their classrooms.

Urban School Settings

Economic stress on students and families. The challenges of inner city environments do not stop at the school doors. Challenges plaguing low-income school settings affect the ecology of the neighborhood schools. In the current studies, the author acknowledges and attempts to address the complexity of urban settings by recognizing the danger of overgeneralizing and disregarding the heterogeneity of experiences within such settings. At the core of the issue, it can be asserted that schools in the United States (U.S.) are not meeting the various needs of inner city students (Smith & Smith, 2006). One challenge in many urban settings is poverty. William Julius Wilson (2003, p. 1106) calls this epidemic the "new urban poverty," a term which operationalizes how a majority of adults in our nation's metropolises are jobless. Poverty then breeds a host of other challenges in these communities. For instance, mostly African-American families as well as Latino families reside in urban

areas plagued by inadequate school quality, inadequate job access, and growing social isolation (Wilson, 2003). Furthermore, some U.S. cities that were once the epicentres of financial opportunity are now deeply segregated, suffer from the persistence of gun violence and gang activity, and are highly impoverished. The neighborhoods in which the current study's high schools are located are similarly characterized by Wilson's (2003) notion of the "new urban poverty." According to U.S. Census Data (United States Census Bureau, 2015) 29.1% of the people in the mid-sized Northeastern city in which the study's high schools are located were living below the poverty line from 2009-2013. In the last five years, the median household income in this city was $33,960 and a minority (22.7%) of residents owned homes (United States Census Bureau, 2015). African-American families comprised the largest racial and ethnic group (52.4%), followed by Latino families (33.8%) and White families (26.3%; United States Census Bureau, 2010).

While the current studies take place in the Northeastern U.S., a city such as Detroit provides another illustrative example of how a metropolis can suffer from new urban poverty (Wilson, 2003). Detroit declared bankruptcy in 2013 while scholars argue that this was due to a city whose economy was dominated by an insecure auto industry, decline of housing and abandonment in the central city, and racial tensions that were particularly present in the workforce (Galster, 2012; McDonald, 2014). The racial conflict and decline in opportunities in Detroit may serve as an analogy for many of our country's inner city environments similar to the one in the current studies. The nationwide trend has been that those with means move away from highly impoverished areas and toward suburban areas (e.g., mostly higher income, White families) while the quality of life for many urban inhabitants continues to decline. With Detroit's loss of population and jobs in the last decade (2000-2010), the city saw a jump in poverty rates, a spike in murder rates, a decline in median family income, and a striking increase in the percentage of single-parent families in the city (e.g., 71% single-parent

families; McDonald, 2014). As Wilson (2003) points out, a confluence of these factors has created poor, segregated communities in our nation's metropolises.

Exposure to violence in neighborhoods. Violence and gang activity in urban communities provides a challenging environment for students and teachers within schools as well. In larger cities, the most common characteristics of gangs, as reported by law enforcement agencies, are that gang members commit crimes together, have a name, claim turf or territory, and hang out together (National Gang Center, 2015). Of concern is that gun violence in gang-related homicides has risen from 73% in 1980 to 92% in 2008 (Cooper & Smith, 2011). According to the U.S. Department of Justice (2012), the National Youth Gang Survey (NYGS) from 2010 had a high response rate (85%) from a representative sample of 2,158 police departments nationwide in large cities, small cities, suburban communities, and rural areas. The NYGS from 2010 estimated that there were 29,400 gangs and 756,000 gang members in the United States that year (U.S. Department of Justice, 2012). According to the NYGS, very large cities (e.g., cities with populations over 100,000, including the Northeastern U.S. city in which the study schools are located) had the highest concentration of gang-related activity. Additionally, in 2010, of more than 700 total homicides in Chicago, Illinois and Los Angeles, California, more than half of those homicides were reported to be gang-related (U.S. Department of Justice, 2012). Gun and gang violence in urban communities is gaining recognition on a national level. The Children's Defense Fund (2013) published a report that cited 2,694 children and teens died from gun violence in 2010. Of those 2,694 children was Hadiya Pendleton, a 15-year-old, who had marched in President Barack Obama's Inauguration and was killed by gunfire less than a mile from the President's former Chicago home; her death was one of three deaths and eight injuries from guns on a single day in Chicago (Child's Defense Fund, 2013). Even though gun violence and gang activity persists in many urban communities, schools themselves are seeing an overall decline

in violence (Mayer, 2010). The National Crime Victimization Survey (NCVS) saw violent crimes declined from 48 students affected per 1000 students in 1992 to 22 students affected per 1000 students in 2004 (Devoe et al., 2005; Mayer, 2010). The current studies are embedded within a community and schools where gun violence and gang activities are a threat once students leave school grounds, as is the norm in many of our nation's metropolises.

Summary of economic stress and exposure to violence in urban school settings. Many of the aforementioned challenges in urban settings are corroborated by empirical evidence about urban schools themselves. Annually, the National Center for Education Statistics (NCES) provides statistics on the condition of education. In 2012, approximately 11.1 million U.S. students aged 5 to 17 lived in families whose income fell below the poverty line (e.g., a family of four making less than $23,283); more specifically, 18% of families in the Northeast met these criteria in 2012 (Kena et al., 2014). Additionally, the percentage of children in poverty nationwide for African-American and Latino families, which comprise of a majority of the student body for the current studies, was considerably higher at 39% and 33% respectively. Across different school locales, urban schools had the largest percentage of students (34%) in 2011-2012 classified as living and attending high poverty schools (e.g., more than 75% of students on free or reduced-price lunch; Kena et al., 2014). In sum, these challenges to urban communities provide the landscape for school systems in need of programs that encourage support and reform.

Teacher Challenges

Teacher burnout and attrition. Current teacher attrition rates of about 50% in the first three to five years of teaching in urban schools are troublesome (Elias et al., 2003; Eslinger, 2014; Ingersoll & Smith, 2003). Besides leaving the profession due to student misconduct (Borman & Dowling, 2008), federal legislation such as No Child Left Behind

(NCLB) has led teachers nationwide to feel pressured to teach "to the test." NCLB, which started under the Bush administration, dictates that schools not making adequate progress on test scores for three consecutive years are considered failing according to nationwide standards (Good et al., 2014). In these failing schools, teaching for test preparation has the potential effect of creating less autonomy in teachers' development of lesson plans for students. Autonomy is a powerful antidote to burnout (Cherniss, 1995). Referencing all of the high stakes testing that is used to evaluate teachers, students, and schools, Elias et al. (2003, p. 304) poignantly states, we need "to prepare students for the tests of life, not for a life of testing." Alongside high stakes testing, pressure to teach more rigorous academics results from the push toward the Common Core curriculum. These factors are often cited as sources of teacher dissatisfaction in media outlets across the nation, and recently, New York's statewide teachers' union withdrew its support of the Common Core curriculum (Baker, 2014). Additional stressors on teachers include rigorous new teacher evaluation systems. Scholars have argued problems with teacher evaluation systems are two-fold in that: a) they are based on inaccurate measures of teacher quality and b) they do not offer supports that ultimately help develop a highly skilled teacher workforce (Bill and Melinda Gates Foundation, 2011; Marzano, 2012). Moreover, teacher evaluation systems are often punitive in nature. These factors at the national and state level are likely contributing to the highest attrition rates the teaching profession has seen in recent years for both younger, newer teachers and older, seasoned veteran teachers, as shown in New York City's Department of Education's six year study following teachers that entered and left the teaching profession (Boyd et al., 2009).

Lack of support from administrators. The aforementioned challenges facing teachers occur outside of each school's control. That said, there are factors within school's inherent structure that are also placing stress on teachers. Interactions between administration

and teachers within schools can be predictive of why teachers stay in schools (Tickle et al., 2011) or why they choose to leave the profession (Boyd et al., 2009; Cherniss, 1995). Administrative support as defined by Borman and Dowling (2008) is how schools assist their teachers amongst a variety of areas that include but are not limited to: instructional methods, curriculum, and student discipline. From this definition, it can be asserted that the relationship between administrators and teachers ultimately contributes to the teacher, student, and school experience. Further, these relationships exist within the school climate and culture, which Forman (2015, p. 27) defines as a school's perception on "the way things are done around here." Conflicting relationships between teachers and administration are not uncommon. A lack of perceived administrative support has been linked to high teacher attrition rates (Boyd et al., 2009). A lack of administrative support and meaningfulness in teachers' work increases the likelihood of burnout (Cherniss, 1995). Conversely, administrative support built on a foundation of trust has been shown to prevent early career burnout for teachers (Cherniss, 1995). Burnout that can lead to attrition is conceptualized by Cherniss (1995) as a process where teachers or others in human service organizations gradually become more concerned with looking out for their own needs and less concerned with the needs of others (e.g., students) that they are serving. However, this particular challenge can be prevented. For instance, in the 2003-2004 School and Staffing Survey, Tickle et al. (2011) found administrative support as the most predictive of teacher satisfaction and teachers' intent to stay in the profession. A particular struggle of urban school environments, according to Elias et al. (2003), is their high turnover rates for teachers, administrators and superintendents for whom an average tenure is two years. With high turnover and attrition in many urban schools, it is challenging to form the teacher-administrative support necessary for a well-functioning school.

Teachers under scrutiny. In addition to national and state-wide legislation as well as a lack of administrative support in some schools, schools in urban areas that do not perform well academically cause disproportional stress on the teachers. As mentioned before, NCLB may stigmatize a school by labelling it as "failing," (Good et al., 2014) or schools are classified as "underperforming" and can be taken over by state control (Dolan, 1992). These labels and intense scrutiny from outsiders create a more stressful environment for the teachers in these schools, who are tasked with the responsibility to remediate. It is ultimately the teacher's responsibility to implement proven strategies to help students that are coming from high poverty households in urban areas as described above (Elias et al., 2003). Moreover, another major challenge to teachers in urban areas is that schools are often asked to undergo reforms which are often implemented under the assumption that there is a simple, short-term fix (Elias et al., 2003). According to Kozol (2005), school reform is meant to target teachers that teach poor students of color in deeply segregated schools, and in these schools, the problems underlying poverty are not being addressed (Elias et al., 2003). Elias and Leverett (2011) acknowledge that teachers in urban school settings have crowded reform schedules, feel that new programming is implemented in a punitive manner, and are working extremely hard in highly pressured conditions. Thus, with so many fixes suggested in failing, urban schools, it appears that the possible solutions and programs mandated to teachers often leaves teachers with mixed messages and unable to decide how to best allocate their teaching efforts.

Negative disciplinary climates. Teachers in schools across the nation receive mixed messages when it comes to school discipline. For instance, teachers could have higher stress in punitive oriented school climates where zero tolerance policies guide discipline. Zero tolerance policies started as an approach to drug enforcement in schools, but these policies are now used to mandate predetermined, punitive consequences for students that break the

rules (APA Zero Tolerance Task Force, 2008; Skiba & Rausch, 2006). The American Psychological Association (APA) Zero Tolerance Task Force (2008) cited that these policies were widespread across the nation. By 1997, at least 79% of schools across the nation had adopted zero tolerance policies in regard to drugs, alcohol, and violence (Boccanfuso & Kuhfield, 2011). In fact, the very high profile forms of school violence that zero tolerance policies target, theft and personal attack, are less problematic in schools currently than low-level incivility that occurs between student and student or between student and teacher (Mayer, 2010). In a report on the overall state of school violence, Mayer (2010) asserted that overall school violence from 1993-2005 has declined, yet pervasive low-level incivility in schools creates what he calls "toxic" learning environments. Widespread incivility has deleterious effects on school climate, school connectedness, and student perceptions of safety (Skiba et al., 2004).

Ultimately, the APA Zero Tolerance Task Force (2008) found that zero tolerance policies have not been shown to improve school climate or safety, and suspension and expulsion are not effective ways of dealing with student misbehavior. In fact, the American Academy of Pediatrics (2013) published a statement calling the effectiveness of exclusionary actions in school discipline, (e.g., out-of-school suspensions, expulsions) "increasingly questionable." As an alternative, the APA Zero Tolerance Task Force (2008) recommends abandoning these practices that lack empirical evidence and disregard children's developmental needs while looking toward "policies that can keep schools safe and preserve the opportunity to learn for all students" (pp. 860). Since the APA Task Force (2008) report, new studies have corroborated their conclusions that punitive discipline methods negatively affect academic outcomes and also increase the likelihood that a student will dropout out of school (Balfanz, Byrnes, & Fox, 2015; Fabelo et al., 2011). In direct reference to the school-to-prison pipeline, Fabelo et al. (2011) found evidence from analysing both educational and

juvenile justice records in Texas that 59% of students disciplined 11 times or more did not graduate from high school. In addition, nearly half of this group of students (those students disciplined 11 times or more) were in contact with the juvenile justice system (Fabelo et al., 2011). Other studies have shown that after accounting for demographics, attendance, and course performance, each additional suspension a student receives decreases the odds of high school graduation by 20% (Balfanz et al., 2015). When the *Breaking School Rules* (Fabelo et al., 2011) report came out, it garnered attention due to its rigorous, longitudinal design and striking conclusions. Starting in the seventh grade, millions of student's educational records and juvenile justice records were made available to researchers for multivariate analyses over the course of six years (Fabelo et al., 2011). Some key findings from this study were that 59.6% of students experienced a suspension or expulsion in the Texas public school system. Of those suspended or expelled, the most at-risk groups were African-American students, especially males, and those classified with a disability (e.g., emotionally disturbed). Another key finding was that after accounting for 83 risk factors including poverty, African-American students still had a 31% higher likelihood of receiving a school discretionary action than identical Latino or White students (Fabelo et al., 2011); this finding provides direct support for the racial discipline gap, a term which scholars operationalize as "minority disproportionality in school discipline" (Skiba et al., 2002, p. 318). Other ethnic minority groups are at risk for school disciplinary action. Finn and Servoss (2013) found in a sample of 10th grade students that Latino males were twice as likely to be issued an out-of-school suspension than their White classmates. Additionally, Losen and Martinez (2013) demonstrated high rates of suspension for African-American females (13%).

Studies have also shown that when students are enrolled in schools predominantly comprised of African-American students, they are more at risk for receiving exclusionary discipline (Gregory, Cornell, & Fan, 2011; Welch & Payne, 2010). For example, using

multivariate analyses of data from 199 Virginia high schools, Gregory, Cornell and Fan (2011) found that in predominantly African-American high schools, White and African-American students were issued more suspensions than schools with proportionally fewer African-American students. Additionally, from a sample of 294 public schools, Welch and Payne (2010) were the first researchers to directly test what they called the "racial threat hypothesis" in school settings; they found that schools with high populations of African-American students were shown to respond harshly and not respond restoratively to student misbehavior. Welch and Payne (2010) conclude that punitive discipline measures are used to control a perceived, unfounded "threat" when African-American students come together in large numbers.

Summary to challenges facing teachers and students in urban schools. In sum, economic stress, exposure to violence, teacher burnout leading to attrition, lack of support from administrators, teachers under scrutiny, and negative disciplinary climates create high stress environments in urban school settings. Ultimately, these school systems are challenging places in which to be a teacher or a student. Taken as a whole, the settings can be characterized by strained teacher and student relationships. Furthermore, these are difficult environments for programs that are aimed at reform. The current studies examine how a program designed to introduce restorative approaches to discipline faired in the early stages of implementation in two high poverty, urban high schools comprised of mostly African-American and Latino students.

Restorative Practices

A new approach to discipline. As a response to the troublesome dynamics that arise between students and school staff throughout the nation's classrooms, there are school-wide programs aimed at reversing the troubling trends (e.g., punitive discipline, zero tolerance policies) mentioned thus far. One such school-wide program is restorative practices (RP),

which aims to transform how students and teachers interact with one another (Costello, Wachtel, & Wachtel, 2010). RP draws its philosophical roots from traditional societies (e.g., the Maori tribe in New Zealand and citizens of Papua New Guinea); these societies place emphasis on mechanisms such as making amends, reconciliation, and compensation that have probably existed throughout human prehistory (Diamond, 2012), which makes these civilizations differ markedly from current justice system practices and traditional school discipline practices (e.g., zero tolerance policies; APA Zero Tolerance Task Force, 2008) in the United States. Nevertheless, restorative justice (RJ), especially for the juvenile justice system has gained some momentum in the U.S. RJ in the juvenile justice system preceded RP in the schools and is helpful in providing a way of approaching criminal justice from a perspective that emphasizes repairing the harm done to people and relationships rather than only punishing offenders. There is promise for improving teacher-student relationships as both RP and RJ allow an opportunity for reconciliation, collaborative problem solving, and perspective taking—key components of building relationships (Schiff, 2013). The 11 Essential Elements of RP are provided in Table 1 and Appendix A. In their SaferSanerSchools program, trainers from the International Institute for Restorative Practices (IIRP) work with school staff (e.g., teachers, administrators, school psychologists) over the course of two years when implementing their whole school change program.

RP program impact theory. In Appendix B, the author outlines a hypothesized program impact theory for RP as adapted from the IIRP program materials in order to understand RP use, stages, and applicable terminology for the current study. Inputs for RP include teachers as the main RP facilitators in schools, school administrators to oversee and implement RP as well, students as clients, key opinion leaders (KOLs) to influence their peers in the organization to use RP, and money from taxpayers to pay for trainers/consultants from the International Institute for Restorative Practices (IIRP). The initial activities of the

outputs include multiple days of training by IIRP trainers/consultants and then booster sessions and IIRP visits throughout the first two years of implementation. The main output activities as suggested by the IIRP are for the staff (e.g., mainly teachers) of a school to run proactive circles on a *weekly* basis in classrooms or after a minor incident (e.g., responsive circles). According to these guidelines of RP output activities, a teacher implementing the program would be expected to run more than 30 proactive circles in his or her classroom throughout the school year. More broadly, staff members are asked to implement RP by using many of the 11 Essential RP Elements in their daily classroom activities with students to encourage them to share their perspective and have their voice heard. In regard to participation, the whole school is meant to be involved in RP as it is marketed as a whole school change process. KOLs and other staff members proficient in RP are asked to join Professional Leadership Groups (PLGs) that may meet on a regular basis to support staff in the implementation of RP. Gregory et al. (2014) identified RP as having a prevention as well as intervention focus in how school staff addresses student misconduct. The current studies explore RP from the perspective of the teacher (main RP facilitators) and their implementation of the program in a highly stressed urban school setting affected by the challenges described above.

Table 1

Defining the Eleven RP Essential Elements

Elements	Definition
1. Affective Statements	Students and staff express feelings
2. Fair Process	Everyone treated respectfully
3. Restorative Questions	Help identify who is harmed
4. Reintegrative Management of Shame	Listening to/acknowledging feelings of shamed person
5. Small Impromptu Conference	Questions/problem solving to resolve low-level incidents
6. Restorative Staff Community	Administration/staff model RP use
7. Proactive Circles	Everyday circles used to build community
8. Restorative Approach with Families	School uses RP with family interactions
9. Responsive Circles	Circles after a minor incident
10. Fundamental Hypothesis Understanding	Humans are happiest when those in authority do things *with* them, not do things *to* them
11. Restorative Conferences	Structured protocol for serious incidents

Note: Adapted from the International Institute of Restorative Practices (IIRP): Safer Saner Schools Program.

Empirical theories supporting RP. RP's Essential Elements are founded on community building, relationships, support, and giving students a voice. Thus, there is an overlap with RP philosophy and authoritative teaching (or parenting) and relational approaches to form these relationships. Authoritative teachers (or parents) offer "responsiveness" and "demandingness" in their relationships (Walker, 2008; Wentzel, 2002) and have also been conceptualized by researchers as those with high degrees of "support" and "structure" in their teaching (Gregory & Weinstein, 2004; Gregory et al., 2010; Gregory & Korth, in press). This balance of being warm and demanding in one's teaching style has been shown to link positively to the level of student engagement in the classroom (Rolland, 2012; Rubie-Davies et al., 2014; Weinstein, 2002). In addition, Gregory and Ripski (2008) examined why teachers who emphasize relationship and community building through this authoritative and relational teaching style in the classroom elicit adolescent cooperation. From their mediational analyses, they found relationship-oriented teachers tended to be experienced by their students as trustworthy, and then trusting students cooperated with school rules and were less likely to receive discipline referrals (Gregory & Ripski, 2008). This study's findings were corroborated by a more recent study by Gregory et al. (2014). Students who reported their teachers used RP elements frequently tended to also feel respected by these teachers (Gregory et al., 2014). Whereas much is known on the beneficial effects of authoritative teaching styles and relational approaches in the classroom, less is known about teachers' readiness to adopt a program relying on a relational approach to discipline in high school classrooms embedded within urban communities.

RP and student outcomes. Currently, there are small-scale studies that have shown promise for RP, yet gaps in our knowledge of RP and the program's effectiveness remain. In a review of single case studies, Lewis (2009) showed that the frequency of punitive school discipline decreased in many U.S. schools over the course of RP implementation. The

positive findings occurred across rural, suburban, and urban schools. Most relevant to the current study, one predominantly African-American school in an urban setting had a reduction in serious incidents and violent acts by 52% in their first year of RP implementation (Lewis, 2009). In another urban school, the number of students with multiple suspensions dropped from 330 students in the 2010-2011 school year to 120 students in 2012-2013 school year, a reduction of 64% (Improving School Climate, 2014).

Few studies have examined how the quality of RP implementation relates to student outcomes. A few exemptions are noteworthy: Gregory and colleagues (2014) examined 29 high school classrooms through 412 student surveys. They found that student perceptions of high teacher RP use was associated with feeling respected and teachers' infrequent use of office discipline referrals (ODRs) for disruption and defiance. A markedly smaller racial discipline gap was observed with teachers that had high RP use rather than low RP use. For instance, teachers perceived as low in RP implementation averaged 9.13 ODRs for African-American and Latino student defiance/disruption versus 1.69 ODRs for White and Asian student defiance/disruption. However, high RP teachers issued 2.92 ODRs to African-American and Latino student defiance/disruption versus 0.77 ODRs to White and Asian student defiance/disruption. Consistent with another study of RP, Gregory et al. (2014) identified variability in the quality of RP implementation (e.g., low versus high RP use) from teachers in the same school (McClusky, 2008). While variability in RP has shown differential effects on student perceptions of teacher relationships and discipline referrals in those classes, little is known about why teachers, from their own vantage point, vary in their implementation of RP. The current studies aim to identify characteristics of teachers and their school setting, to help explain some of the variability in the early stages of RP implementation.

Implementation Science

Implementation models. While the evidence base for RP is growing (Gregory et al., 2014; Improving School Climate, 2014; McClusky, 2008), questions are still numerous about its implementation process, especially in urban school settings. That said, two prevailing models of implementation offer useful frameworks for inquiry. Aarons et al. (2011) describe the Exploration, Preparation, Implementation, and Sustainment (EPIS) framework to elucidate the different stages of school implementation and innovation. Fixsen et al. (2005) have identified six stages to the EPIS framework while Forman and Selman (2011) have elaborated about what occurs at each stage of this process. First, in the *exploration and adoption* phase, one sees if there is a match between the program (e.g., RP) and the school and if there are resources in place for the innovation. Second, *program installation* accounts for how new responsibilities, policies, and procedures will be put into place. Third, *initial implementation* or early implementation is when implementers deal with making a change and breaking from the status quo or choose to conduct business as usual; this early stage of implementation is the focus of the current study. Fourth, *full operation* of a program is when there is administrative support for an innovation and students are receiving the program from staff in the school. Fifth, *innovation* occurs where a new program is made to fit the conditions of the organization and both desirable and undesirable changes are incorporated into a system (e.g., for RP, this is a two to three year process; Gregory et al., 2014). Sixth, *sustainability* is when a program continues for years after the first innovations were made despite turnover of staff and changes in the school environment.

Another model of implementation accepted by scholars focuses on the organizational implementation context (OIC). Aarons et al. (2014) recommend a careful assessment of the OIC and consideration of the perspectives of individual teachers as well as administrators, schools, and district levels when implementing an innovation. In this model of the adoption

of evidence-based programming (EBP), perceived support from administrators is most critical in adoption of innovations from teachers (Rohrbach et al., 1993) and overall student participation in EBP (Kam et al., 2003). Within Global Organizational Constructs, one must consider the school climate, principal leadership, and teacher participation behavior. Within the OIC, the implementation climate and implementation leadership are focal to the behaviors of those to which the innovation is targeted. These factors have a significant effect on the system's attitudes toward EBP. To have positive Implementation Outcomes, acceptability, appropriateness, and fidelity are important to consider throughout the implementation. Finally, to account for Behavioral and Educational Outcomes, authors suggest examining a range of student indicators including student attendance rates, standardized test scores, and discipline data to see if the EBP was effective and adopted by the system. Thus, this framework's theory is comprised of: Global Organizational Constructs, OIC, Implementation Outcomes, and Behavioral and Educational Outcomes. In both frameworks, each step requires support for the innovation and implementers to adopt the change, so there are many junctions at which a school may abandon an innovation and return to the status quo (e.g., business as usual). For instance, in schools, teacher attitudes ultimately impact how the intervention is carried out as planned (Brackett et al., 2012; Lohrmann, 2008) and strong principal leadership has been identified as important to implementation as well (Elias et al., 2006). Amidst these frameworks of innovation and implementation, it is often observed, "bad systems trump good programs" (McCarthy & Kerman, 2010).

Adopter types. Given that the current study takes place in the early stages of an innovation, adopter types are informative for classifying the different perceptions and characteristics of the teachers in the study. For studies at multiple time points, Rogers (2003) has categorized different adopter categories as: Innovators, Early adopters, Early majority, Late majority, and Laggards. However, due to the nature and design of the current study,

adopters will not be classified by these time-sensitive adopter categories. The current study draws on adopter types by focusing on teachers with high versus low use of RP (Gregory et al., 2014).

Little is known about why teachers would be high frequency users or low frequency users of RP in a highly stressed urban setting. Within school organizations, Forman (2015) suggests focusing on "fit" and "support" when implementing a program. Again, new practices and programs need to fit the mission of the school and values of those within the system while they must also garner the support of teachers and administrators, especially the principal; these factors are seen as essential elements in the acceptance and adoption of an innovation. As such, it is important to discern what teacher attitudes and characteristics either lead to adoption or non-adoption of RP and their perceived levels of support from their administration when implementing a program such as RP.

Predictors of Early Implementation

Within the early stages of RP implementation, teachers may choose to adopt (with low frequency or high frequency) or not adopt a program based on a variety of factors. For the current studies, four constructs (self-efficacy, administrative support, feasibility, and positive reinforcement) will be examined to explain why a teacher chooses to adopt an innovation at differing rates or do business as usual. A body of evidence suggests these four constructs may be critical in predicting implementation (Dane & Schneider, 1998; Forman, 2015; Forman, Olin, Hoagwood, Crowe, & Saka, 2009; Tschannen-Moran, Hoy, & Hoy, 1998; Proctor et al., 2011; Weiner, 2009).

Self-efficacy. Self-efficacy is a multi-faceted construct, which has some influence on implementation behavior. Bandura's definition of self-efficacy (1997) describes how an individual's beliefs surrounding their *ability* to organize and implement actions can affect behavior. Tschannen-Moran et al. (1998) extend this definition of self-efficacy to incorporate

how teacher-efficacy not only affects their own teaching behavior but also has lasting effects on student achievement, student motivation, and students' sense of efficacy (Klassen et al., 2011). For the individual teacher, their degree of self-efficacy affects their teaching goals, teaching efforts, and level of aspiration (Tschannen-Moran et al., 1998); those teachers with a greater sense of self-efficacy are those who will be open to new ideas and methods to better meet student needs (Berman et al., 1977; Guskey, 1988; Stein & Wang, 1988). In contrast, teachers that report low self-efficacy often have less satisfaction in their job, more difficulty teaching, and higher levels of stress in the school setting (Klassen et al., 2011). Most relevant to the current studies, it is important to note that teachers with greater self-efficacy are often identified as more resilient (Tschannen-Moran et al., 1998). Additionally, change-efficacy is a form of self-efficacy identified by scholars (Forman, 2015; Weiner, 2009), defined as staff members' (e.g., teachers) beliefs about their capability in implementing a new program. Weiner (2009) has found that change-efficacy depends on teacher perceptions (whether positive or negative) of available resources, implementation task demands, and situational or contextual factors. As of yet, it is unknown whether variability in teacher self-efficacy will affect RP implementation behavior (i.e., high use of RP versus low use of RP). The degree to which teachers feel prepared to implement RP in the current studies will be referred to as *RP self-efficacy*.

Administrative support. Whereas administrative support has been observed by Tickle et al. (2011) as the most predictive factor for teacher retention and has been discussed as a means to prevent early career burnout (Cherniss, 1995), it also impacts program implementation (Forman, 2015). In fact, lack of administrative support has been identified as a mechanism for teacher burnout (Cherniss, 1995) and attrition (Boyd et al., 2009). In its simplest form, Borman and Dowling (2008) have defined administrative support as how schools assist their teachers across domains. When evaluating a program, Forman (2015)

suggests asking whether there is an administrative structure in place to support the intervention. Administrative support is not necessarily stagnant. It is hypothesized that principal support can increase by involving staff members in the planning of an innovation to help promote a positive implementation climate (Aarons et al., 2014; Forman, 2015). Further, the administration plays a substantial role in the implementation process. For instance, if administrators reject innovations, teachers may respond by lessening their commitment to innovations with each subsequent school year (Hatch, 2002). For successful implementation to occur, Forman et al. (2009) included administrative support among a list of essential factors in the implementation process.

Feasibility. Forman (2015) highlights the importance of program feasibility when evaluating outcomes in implementation research. Proctor et al. (2011) define feasibility as how successfully a program can be used in a specific setting. When analysing feasibility, individual providers (e.g., teachers) or organizations are asked about the actual fit, suitability for everyday use, or practicability of a program (Proctor et al., 2011). In order to bridge the gap between research and practice, Forman (2015) suggests exploring program feasibility from the perspective of those individuals undertaking an innovation and providing feedback about the specific barriers that arise when using said innovation. For example, some hindrances to feasibility could be lack of available resources or intensive training requirements (Proctor et al., 2011). For the purposes of the current studies, feasibility from teachers will be measured through analysis of a teacher's perceived level of competing interests or barriers to implementation.

Positive Reinforcement. Skinner (1969) identified the operant model of conditioning. According to Skinner (1969), positive reinforcement is the process involving the presentation of a stimulus or addition to a situation that increases the future likelihood of that response in that situation. A hypothesized mechanism for teachers implementing a

program might be that a self-reported positive experience with said program would increase the likelihood that they might use said program. For implementation, Forman (2015) identified that reinforcement contributes to how potential implementers and stakeholders perceive the value of innovations fitting with their own interests and needs. Additionally, both real and perceived outcomes affect the impact of implementation, so positive reinforcement might explain why a teacher continues to use a program. For instance, an example of this would be that a teacher runs an RP circle that focuses on perspective-taking by asking students what is meant by the analogy, "to walk in another person's shoes." Later, the teacher hears a student making fun of another student. The teacher pulls the student aside and reminds them to think about the other student's feelings from their circle on perspective-taking. The teacher saw a reduction in teasing from this particular student. This experience with the individual student likely reinforced the teacher and motivated the teacher to use RP circles more frequently to teach social-emotional skills. Positive reinforcement has shown effectiveness in promoting positive behaviors with diverse populations (National Research Council, 2004). Positive reinforcement also appears to be an important component of Tier 1 interventions to strengthen cooperative behavior in Positive Behavior Interventions and Supports (PBIS; McIntosh, Ty, & Miller, 2014). There is a strong evidence base for the effects of positive reinforcement on students. More research is needed to learn about the role of positive reinforcement as a correlate of teachers' program implementation.

Summary

As Elias et al. (2003) identified in urban school settings, teachers are being asked to implement new strategies and programs to remediate a host of existing problems. Some of these programs and strategies may be beyond their realm of influence. That said, some teachers are resilient and rise to the challenge of adopting new approaches. In already packed schedules, why do some teachers choose to adopt a program like RP and others decide not to

do so? Individual characteristics, perceptions, and experiences may differentiate teachers' relative uptake of new programming. Specifically, prior research suggests four factors are worthy of additional study: a) teachers' sense of RP self-efficacy, b) perceived administrative support, c) program feasibility, and d) positive reinforcement after trying out a new program strategy. Moreover, the current studies contribute to an understanding of facilitators of program implementation in urban, high poverty schools where teachers and students face the challenges that accompany economic stress, sparse employment opportunities, and unsafe neighborhood conditions (e.g., new urban poverty; Wilson, 2003).

The current dissertation is comprised of two studies: Study 1: Teacher Reports of RP Implementation in Year One and Study 2: Teacher Reports of RP Implementation in Year Two. Using a mixed methods design, Study 1 examined teacher perceptions of RP implementation through interviews and surveys. Teachers were asked to share their experiences with a team of researchers to learn how the first year of RP implementation varied from teacher to teacher. Analyses examined the degree to which administrative support, feasibility, positive experiences, and RP self-efficacy were associated with teachers' RP adherence and use in the early stages of implementation. Following a subsample of Study 1's teachers into the second year of RP implementation, Study 2 examined the degree to which teachers remained frequent users of RP, despite changes in the school administration and a lack of continued RP training.

Taken together, the studies focus on teachers' fidelity of implementation of RP, new programming that aims to transform punitive discipline climates. The research on RP is nascent: RP is associated with reductions in discipline referrals and suspensions in single case designs (Gregory et al., 2014; Lewis, 2009), yet little is known about how and why teachers choose to (or choose not to) adopt this program. The current studies aim to address this gap in knowledge by shedding light on teacher perceptions of RP and their self-reported use of RP

in the early stages of implementation. The methods and findings are presented separately for Study 1 and Study 2, and then results are considered together in the discussion section.

Study 1: Teacher Reports of RP Implementation in Year One

Research Question 1: At the end of year one of school-wide RP implementation, to what degree did RP implementation vary in a cohort of RP trained teachers?

Hypothesis 1:

Through an examination of patterns of self-reported RP adherence and use, it was anticipated that teachers would demonstrate considerable variability in fidelity of implementation. Further, it was hypothesized that a majority of the teachers would report low use of RP while a minority of the teachers would report high use.

Research Question 2: At the end of year one of school-wide RP implementation, did teacher's self-reported perceptions (e.g., teacher's RP self-efficacy, perceived level of RP administrative support, RP feasibility, and positive reinforcement with RP) correspond with differing levels of adoption and implementation?

Hypothesis 2:

It was hypothesized that, as indicated on surveys, teacher reports of a greater sense of RP self-efficacy and perceived administrative support in their utilization of RP would correlate with greater implementation of RP in their classrooms, relative to their colleagues who reported lower levels of RP self-efficacy and RP administrative support. As indicated by their interviews, those teachers who suggested RP was feasible for their classroom (indicating fewer barriers and obstacles) and who specified positive experiences with RP strategies would report high adherence and use of RP in their classrooms. It was anticipated that the results would hold accounting for teacher gender, years of teaching experience, race/ethnicity, and days of RP training.

Methods

Participating schools

Schools in the study were selected given the district had allocated resources for them to receive the two year IIRP SaferSanerSchools RP training program. The district planned to follow the two schools' lead for a large-scale roll out of the programming in subsequent years. Importantly, the current study was initiated at the end of the first year of implementation. Including the initial training and booster sessions, teachers in the participating schools reported attending no days of training (2%), 1 day of training (4%), 2 days of training (22%), 3 days of training (30%), 4 days of training (34%), 5 days of training (6%), or 7+ (2%) days of training from the IIRP. IIRP also provided on-site consultation to teachers and administration and monthly in-depth phone calls to the schools in this study. The IIRP describes the content of training to include: using circles for behavioral and academic purposes, involving staff in the implementation plan, and refining RP skills throughout the school year.

The two schools in the study had various reform needs as identified by a state report. At the time that this report was made public, the two schools were comprised of one single school. Therefore, the following statistics are representative of the two schools in the study. For 2013-2014, these schools were reported to be significantly lagging in academic performance (threshold for this criteria is equal to or below the 19.9th percentile in the state) compared to other schools in the state. In comparison to peer schools (those similar in demographics and situated in urban school settings within the state), the schools were significantly lagging in college and career readiness as well as graduation and post-secondary performance. Enrolment in the school for 2013-2014 was 53.4% Latino, 43.1% African-American, 2.7% White, 0.6% Asian, and 0.2% American Indian. Of the student population, 72.2% of students were classified as Economically Disadvantaged Students, 27% were

Students with a Disability, and 11.8% were Limited English Proficient Students. For this academic year, the graduation rate was 52% at the 7th percentile for peer schools and 3rd percentile for state schools while the dropout rate was 10.2% and at the 6th percentile for peer schools and 1st percentile for state schools. All of these factors align with Wilson's (2003) depiction of new urban poverty and Kozol's (2005) description of highly segregated schools.

Teacher participants

In June of 2014, 76 teachers from these two public high schools were asked to participate in surveys and interviews about their experience of RP. Fifty-one teachers consented to participate in the study, which is a response rate of 67.11%. Participating teachers spanned grade levels (i.e., grade 9 through grade 12) and included teachers from general and special education. In the sample, 25 participants were male (*n* = 25, 49%) and 26 were female (*n* = 26, 51%). Within the sample, 5 participants were aged 18-24 years old (*n* = 5, 10%), 12 participants were aged 25-34 years old (*n* = 12, 24%), 11 participants were aged 35-44 years old (*n* =11, 22%), 11 participants were aged 45-54 years old (*n* = 11, 22%), and 6 were aged 55-64 years old (*n* = 6, 12%); 5 participants did not self-report their age range (*n* = 5, 10%).

The sample was comprised of participants from diverse racial and ethnic backgrounds. For instance, 16 participants identified themselves as Black/African-American (*n* = 16), 12 participants identified themselves as White/Caucasian (*n* = 12), 2 participants identified themselves as Puerto Rican (*n* = 2), 2 participants identified themselves as Cuban (*n* = 2), 8 participants identified themselves as Other Hispanic/Spanish/Latino (*n* = 8), 1 participant identified himself as Korean (*n* = 1), and 1 participant identified herself as Asian Indian (*n* = 1). The survey also allowed participants to specify if they were from other race/ethnicity groups than those provided. One of the Other Hispanic/Spanish/Latino participants specified that she was Dominican while another participant from this category

specified that he was Iranian, Ecuadorian, and Italian. One of the Black/African-American participants specified that she was Sudanese. Additionally, one Asian Indian participant (n = 1), one Native American and Afro-Caribbean participant (n = 1), and one Haitian participant (n = 1) were included in the study. Of the sample (N = 51), 66.7% were Black/African-American/Latino while 33.3% were White/Caucasian/Asian/Other.

Teachers of various subject areas participated. They spanned course content area ranging from English to Spanish and History to Science (e.g., forensic science, environmental science). Some teachers taught Special Education, Inclusion classes, or English Second Language (ESL) classes. Participants' self-reported experience ranged from 1 year to 29 years in their role as a teacher (M = 8.66).

Procedures

Members of the Rutgers research team presented the study and invited teacher participation during a series of individual meetings in June of 2014. The total number of interviews conducted by the Rutgers University research team was 50 while the total number of surveys and demographic information completed was 51. The sample (N = 51) completed consent forms, a teacher survey, interview (except for one consented teacher), and demographic information, which required up to 40 minutes for completion. The interview was audiotaped by a team of either one principal investigator (PI) or one of six graduate students from Rutgers University. Due to technical difficulties, one interview was unable to be recovered; so 49 teacher interviews were used for analyses. Upon completion of this process, participants were compensated with a $25 gift card. Rutgers University Institutional Review Board (IRB) approved the proposed study in December 2013. Then, in May of 2014, the school district granted research approval.

Survey Measures

Teacher self-reported demographic data and RP training status. Teachers were asked to report on their role in the school, their number of years in the role, their current school, if they attended RP trainings (including number of days of training and names of training), their gender, their years of teaching experience, and their race/ethnicity.

RP self-efficacy scale. In this scale, RP self-efficacy corresponds with readiness to adopt RP as an innovation. While Tschannen-Moran and Woolfolk (2001) have done substantial work on capturing and refining teacher self-efficacy in survey form, new self-efficacy items were written for the unique demands on teacher facilitators for the RP intervention. For the current study, teacher RP self-efficacy is measured through 4 items that were created to measure teacher's confidence and preparedness for RP implementation (See Appendix C). Items are based on a four-point likert-scale and ask respondents to rate items ranging from strongly disagree (1) to strongly agree (4) in response to questions such as "I feel confident in using RP elements during my interactions with students." The scale was found to have a good Cronbach's alpha of .87.

Administrative support scale. In this scale, perceived administrative support is measured through 6 items that were created to measure the perceived level of support from the teacher (See Appendix C). While Tickle et al. (2011) used survey items to predict how administrative support related to staff turnover, less is known about how administrative support relates to RP. For this reason, new survey items needed to be administered to hone in on the unique demands of administrative support for teachers using RP. Items are based on a four-point likert-scale and ask respondents to rate items ranging from strongly disagree (1) to strongly agree (4) in response to questions such as "I feel supported by school administrators when I use RP elements." The scale was found to have a good Cronbach's alpha of .85.

Teacher RP implementation. Variability of RP use assessed implementation fidelity. One indicator of fidelity included adherence, defined as the degree to which implementers follow methods and guidelines for delivering the core components of the intervention (Dane & Schneider, 1998). Adherence was measured through the Teacher RP implementation scale. The scale was adapted from the IIRP's survey of teacher knowledge and application of RP elements. It is measured through 6 items that aim to differentiate high use and low adherence using the RP elements (See Appendix C). Items are based on a five-point likert-scale ranging from not at all (1) to always (5) in response to questions such as "I ask students to take specific actions to repair the harm." These 6 questions correspond directly with 4 of the 11 RP Essential Elements in Appendix A, and the scale had a good Cronbach's alpha at .81. The Teacher RP implementation scale will be referred to synonymously with RP adherence and RP Essential Elements in the results and discussion of findings.

Another construct related to implementation fidelity is exposure, defined as number and length of sessions or frequency (Schulte, Easton, & Parker, 2009). The measure of exposure was based on teachers' self-reports of total RP circles/conferences they implemented throughout the school year (hereby called "RP use"). In Appendix C, the responses to the three questions about frequency of circles run (e.g., proactive, responsive, restorative conferences) were aggregated to create the total number of circles and conferences for each respondent.

Semi-Structured, Teacher Interview

Appendix D presents the RP Teacher Coding Manual, which coders used to analyse transcribed teacher interviews. The manual and codes were based on a grounded theory approach (Strauss & Corbin, 1990); the grounded theory approach is one in which identified themes from interviews are used to generate research hypotheses. The author initially read a sample of interviews with Dr. Anne Gregory and tested out the coding process on those

interviews. Once themes were identified, established, and agreed upon, graduate students were trained in how to use the RP Teacher Coding Manual.

Teacher interviews were coded for feasibility and positive reinforcement. In addition to using a grounded theory approach (Strauss & Corbin, 1990), thematic coding of interviews followed the Miles and Huberman (1994) approach to qualitative data; this approach included data reduction, data display, and conclusion drawing. A coding scheme was developed and employed to organize themes for RP implementation. Content analysis was used for data reduction to identify meaning from a vast amount of qualitative data (Patton, 2002).

In order to use the RP Teacher Coding Manual in Appendix D and establish reliability of the constructs being measured, a team of graduate students at Rutgers University attended a two-hour workshop on the manual. In the workshop, coders were asked to rate one teacher transcript on seven constructs with a two- or three-point scale. Each construct was explained thoroughly as was the corresponding rating for the specific teacher's transcribed interview. After this workshop, coders completed 8 more interviews and check-in phone calls were conducted every 2 to 3 interviews that were analyzed to clarify the coding process. Additionally, after this initial reliability period of the 9 coded interviews, the graduate students completed the coding process for the remaining 40 interviews and securely submitted their scores electronically. Overall, excellent agreement amongst coders was shown when ICCs were run (e.g., Competing Practices and Barriers and Positive Experiences with RP both had ICCs of .89).

RP feasibility. Researchers at Rutgers University conducted twenty-minute interviews with teachers in the current study. Appendix E contains the interview protocol. In order to measure feasibility, transcribed responses were coded through thematic analyses. Some of the questions that pulled for teacher's perceptions of RP feasibility included: "How do you build a sense of community in your classroom," "Does RP have the potential to

improve your students' social-emotional learning skills," "Is RP culturally in sync with your students," "Please describe how administrators use RP," "Please describe how other teachers use RP," and "Are there any other practices or programs that give students a 'voice' in solving conflict or repairing harm?" These questions tapped into the feasibility of using RP elements and the overall acceptability (Forman, 2015) of a program like RP to fit teachers' busy reform schedules (Elias et al., 2003; Elias & Leverett, 2011).

Specifically, a team of graduate student coders from Rutgers University rated the teachers' reports of competing practices and barriers on a 3-point scale, which determined the teacher's self-reported lack of feasibility with RP. Coders measured feasibility with the Competing Practices and Barriers code in the coding manual. Feasibility was conceptualized as the lack of or minimization of Competing Practices and Barriers.

RP was not the only initiative introduced to the high schools in the current study, so there was often mention of other programs or practices that could have potentially interfered with RP implementation. For a teacher, there could be no mention of competing practices and barriers or competing practices and barriers were mentioned but minimized in the teacher's implementation of RP (rating of 0). A teacher that showed RP was feasible for her students said:

> "It was just better instruction because everyone was doing the work because everyone was in the circle" (T27).

Additionally, there could be potential obstacles for RP use (rating of 1). An example of a potential obstacle was the teacher that stated:

> "But biology we didn't do much of that and Forensic Science was a really big class so that kind of held me back too because by the time we'd get into the circular shape, that's another 5 or 10 minutes" (T24).

Real or concrete obstacles provided by teachers received a rating of 2. One teacher noted:

> "Because of the different learning levels, I had to take my time for certain things that I fell behind, that I just said, you know what, I really can't. Plus, it was a bit intimidating" (T12).

Competing Practices and Barriers was reverse-coded in regression analyses. Competing Practices and Barriers, had excellent agreement amongst coders as shown by intraclass correlation coefficients (ICC). Competing Practices and Barriers had an ICC of .89.

Positive reinforcement. Instances of positive reinforcement in the interviews tended to arise in the questions: "What kind of support have you received to help you use RP throughout the school year," "Have you had a chance to use RP in your classroom," "What positive changes have you seen in your classroom (from RP)," and "What was your experience of the Restorative Practices (RP) training in preparing you to use RP in your classroom?"

Coders rated Positive Experience with RP on a 3-point scale. If there were no positive experiences clearly mentioned or positive experiences with RP were extremely vague, the teacher would receive a rating of 0. An example of a 0 would be:

> "I tried one responsive circle after I felt like there was an incident and that was a horrible experience for me" (T11).

However, if the teacher had somewhat vague positive sentiments toward RP or a mix of positive and negative experiences with RP, the teacher would receive a 1. For instance, one teacher stated:

> "So, I would say I've had mixed results. I tried it with my other two classes and it was just mediocre" (T47).

In order for a teacher to have received the highest rating of 2, the teacher needed to have provided highly positive RP experiences and a specific example of how RP use improved his or her class. This teacher's response earned a 2:

> "Restorative practices have helped in a sense of giving you another tool to use" (T27).

Positive Experience with RP had excellent agreement amongst coders when ICCs were run. Positive Experience with RP had an ICC of .89.

Data Analysis Plan

Research Question 1: At the end of year one of school-wide RP implementation, to what degree did RP implementation vary in a cohort of RP trained teachers?

Descriptive statistics from the survey data and scales showed means, ranges, and standard deviations. The standard deviation in RP adherence and use indicated the degree to which teachers reported variable implementation of RP.

Research Question 2: At the end of year one of school-wide RP implementation, did teacher's self-reported perceptions (e.g., teacher's RP self-efficacy, perceived level of RP administrative support, RP feasibility, and positive reinforcement with RP) correspond with differing levels of adoption and implementation?

Pearson correlations were run to determine associations amongst teacher demographics, teacher perceptions of RP, and teacher reports of RP adherence and use. Multiple regression analyses were used to identify whether teacher's RP self-efficacy, perceived level of administrative support for RP, feasibility of RP, and positive experiences with RP were significant predictors of a) RP adherence and b) RP use. Covariates were entered into the equations to ascertain if the findings held above and beyond teacher gender, years of experience, ethnicity/race, and RP training days. R-square change was examined as a measure of effect size.

Results

Research Question 1: At the end of year one of school-wide RP implementation, to what degree did RP implementation vary in a cohort of RP trained teachers?

Quantitative Findings

Descriptive Findings. Descriptive statistics for number of circles run (e.g., self-reported frequency of proactive circles, responsive circles, restorative conferences, and the total of all circles and conferences run) can be found in Table 2. These variables reflect the frequency with which the teacher self-reported that he or she used RP circles (e.g., proactive

or responsive) and restorative conferences. Overall, teachers reported somewhat infrequent implementation of circles and conferences at the two schools. Specifically, the total of all RP circles/conferences run (e.g., the combined total of proactive circles, responsive circles, and restorative conferences) ranged from 0 to 50. Across 185 days of school for students, teachers reported an average of 9 RP total circles and conferences in year one of RP implementation (M = 9.04, SD = 11.61). However, two-thirds (66.7%) of teachers reported running 8 or fewer total circles and conferences while only one-third (33.3%) of teachers reported running circles and conferences at the frequency of 10 or more for year one of RP implementation. In addition, some teachers in the study did not run any proactive circles, responsive circles, or restorative conferences (7.8%).

Teacher reports of the RP use varied even more when examining the specific type of circle or conferences implemented. Overall, teachers reported greater use of more prevention-oriented practices (proactive circles) than intervention oriented practices (responsive circles or restorative conferences). Specifically, close to one-third of the sample (29.4%) reported not using proactive circles at all. On average, teachers reported using 4 proactive circles in year one of implementation (M = 4.20). Proactive circles are everyday circles used to build community whereas responsive circles are more high-risk as they are implemented after a minor incident. A greater percentage of teachers (37.3%) reported not running any responsive circles (e.g., circles implemented in reaction to minor conflicts within the classroom). The RP element that required the most training to implement, the restorative conference, was only implemented, on average, less than two times (M = 1.86).

Conversely, Table 3 shows the relatively high levels with which teachers believed that they were adhering to the RP Essential Elements as rated on a likert scale. For instance, each survey item asked whether a teacher was using one of the 6 RP Essential Elements listed in Table 3. As a whole, teachers reported using the RP Essential Elements sometimes or more at

rates of 76.5% to 92.2% on each of the survey items described in Table 3. As a subjective assessment of their RP adherence, this would indicate teachers perceived that they were using RP at relatively high rates. These relatively high levels of using RP elements strikingly differ from the actual number of RP circles and conferences teachers said they led in their classrooms. In addition, Table 4 shows the proportion of the RP elements teachers reported using at least sometimes or more in their classroom. For instance, 52.9% of teachers reported using all 6 of the RP Essential Elements measured by the Teacher RP Implementation scale.

According to the descriptive statistics, there is considerable variability in RP implementation across the teachers in both schools. As hypothesized, a majority of teachers reported low use of RP circles and conferences (below the mean number of circles) while a smaller subset of teachers (33.3%) reported high use (above the mean number of circles). The likert scale on RP implementation elicited a somewhat different response from teachers. Many teachers (52.9%) believed they were adhering to all six of the RP Elements at least sometimes or more in their classrooms. The RP Elements that appeared to be least utilized by teachers were Proactive Circles and Restorative Questions while the most utilized of the RP Elements appeared to be Fair Process.

Table 5 shows descriptive statistics for scales used in multiple regression analyses predicting variability in RP implementation. On average, teachers endorsed that they somewhat agreed with RP self-efficacy items (M = 2.76) and RP administrative support items (M = 2.92). On these scales, teachers responded slightly more positive (more agreement with survey items than disagreement with items) in their perceptions of RP self-efficacy and administrative support with RP.

Table 2

Descriptive Analysis of Number of Circles and Conferences

Variables	*M*	*SD*	*Min*	*Max*	*% No RP Use.*
Total RP Frequency	9.04	11.61	0	50	7.80
Proactive Circles	4.20	5.67	0	25	29.40
Responsive Circles	2.98	4.70	0	20	37.30
Restorative Conferences	1.86	3.87	0	20	58.80

Note. M = mean; *SD* = standard deviation; *% No RP Use* = percentage of teachers reporting never using the RP element.

Table 3

Use of RP Essential Elements

RP Use Statement	High Use	Low Use
I actively encourage students to use affective statements to express how they have been impacted by others' behavior. (Affective Statements)	84.3%	15.7%
I ask the wrongdoer to identify who has been harmed and what harm has been done. (Restorative Questions)	76.5%	23.5%
When addressing misbehavior between students, I structure the conversation using the restorative questions. (Restorative Questions)	82.4%	17.6%
I ask students to take specific actions to repair the harm. (Restorative Questions)	90.2%	9.8%
I use circles to provide opportunities for students to share feelings, ideas, and experiences. (Proactive Circles)	76.5%	23.5%
I actively engage students about rules and ask for their input. (Fair Process)	92.2%	7.8%

Note: Low Use = Teacher responded not at all or rarely; High Use = Teacher responded sometimes, often, or always. The survey items are followed by the RP Element reflected in the item content

Table 4

Adherence to Six RP Elements

Proportion of Elements	*Percentage*
1 out of 6 RP Essential Elements	5.9%
2 out of 6 RP Essential Elements	0.0%
3 out of 6 RP Essential Elements	9.8%
4 out of 6 RP Essential Elements	7.8%
5 out of 6 RP Essential Elements	23.5%
6 out of 6 RP Essential Elements	52.9%

Note: Teachers that responded sometimes or more to the RP Essential Elements as described in Table 3.

Table 5

Descriptive Analysis of Teacher Perceptions of RP

Scales	*M*	*SD*	*Min*	*Max*	*Avg. Response*
RP Self-efficacy	2.76	0.72	1	4	Somewhat agree
RP Administrative Support	2.92	0.66	1.67	4	Somewhat agree
RP Implementation	3.49	0.75	1.33	5	Sometimes

Note. M = mean; *SD* = standard deviation; Avg. Response Range = likert-scale item that corresponded to the mean score when rounded to the nearest whole number.

Research Question 2: At the end of year one of school-wide RP implementation, did teacher's self-reported perceptions (e.g., teacher's RP self-efficacy, perceived level of RP administrative support, RP feasibility, and positive reinforcement with RP) correspond with differing levels of adoption and implementation?

Descriptives of the Qualitative Coding

Positive experiences with RP. During the semi-structured interview, teachers reported on their positive experiences with RP. A teacher with either no mention of positive experiences, extremely vague mention of positive experiences, or mostly negative RP experiences received a rating of 0. Of the 49 interviews transcribed and coded, 8 teachers were captured by this description (n = 8; 16.3%). Here were two teacher responses that received a 0:

> "I mean so, from the get-go, I've heard of other people use it and like for me, I ended up just saying we're not going to do it this year because after hearing people do it, it's like, it seems more like people are doing it because they are trying to appease the administration." (T43, no mention of positive experiences)

> "Right. I tried one responsive circle after I felt like there was an incident and that was a horrible experience for me. I felt like maybe I didn't know what I was doing, and so I kind of put people on the spot and they didn't want to be put on the spot and then I wished that maybe I would have learned more about how to do that before I did that because I think it made things worse." (T11, negative experiences)

Whether it was a negative experience that halted RP implementation or not attempting to use RP, teachers that received a 0 were wholly unable to provide substantial evidence of positive experiences gained from implementing RP.

A teacher with somewhat vague positive RP experiences, general positive sentiments toward RP, or a mix of positive and negative experiences received a rating of 1. A majority of

teachers fell within this rating (*n* = 27; 55.1%). One teacher in particular was an exemplar for a 1. This teacher said:

> "There were a lot of success stories. I'm thinking of one, in particular, the XX teacher who did [circles] a lot and loved them. A few people shied away from it. It was new and like 'what if I'm bad at it.' They were just a little bit overwhelmed by the idea of starting that whole process and knowing what to say. And then there were some people that thought there was a script that went with it and felt like they couldn't do it because they didn't have the script. So mixed emotions across the board from everyone overall." (T24, some general positive experiences and a mix of positive and negative experiences)

Many other teachers reported mixed experiences. Some teachers had certain classes that loved RP circles while other classes expressed verbal distress when RP was implemented.

A teacher that had highly positive RP experiences or provided a specific example of how RP was highly positive for his or her class received a rating of 2. Within the sample, 14 teachers were captured by this description (*n* = 14; 28.6%). Examples of teacher responses for positive experiences with RP are included below:

> "I've had [a] student who gave me problem for months, I called his parents, I tried everything and nothing was working. So um, actually the guy who is being interviewed out there…him and I worked closely together with this student and we sat down and we did a circle and we, he facilitated and asked some questions, and all of us had a dialogue and after that conversation, you know the student, he voiced his concerns. He wanted things a little bit more exciting, a little bit more different, and he told me some of the things that you know maybe we can possibly change in the class, and because of that and without…the students egging him on, when it was just that one-on-one and we were in a circle, I had no problems with him for the rest of the year." (T35, positive experiences with specifics)

> "Yeah. It is the restorative practice just came to reinforce my beliefs so it was much easier for me to implement I think and just talking about my own perspective my own experience but it was very easy for me to implement and to follow through because it's everything I believe." (T1, positive experiences with specifics)

By going into detail, these teachers (*n* = 14) provided the more substantial evidence for positive experiences with RP than a majority of the teachers in both schools (*n* = 35, 71.43%). For teachers that received a 2, some mentioned a change they saw with a student, an experience they had with an IIRP consultant, or a specific change they observed in the classroom or the school from RP use. This level of specificity helped coders distinguish

genuine positive experiences versus general and vague positive sentiments (e.g., "RP is good," "We like RP.")

Competing practices and barriers. Feasibility of RP implementation was reflected when teachers did not mention or minimized barriers or competing practices. To the author of the study, this suggested that these teachers were able to integrate RP into their regular classroom practices. Many teachers (n = 20; 40.8%) indicated a lack of substantial barriers or competing practices. For those teachers, RP seemed feasible since it had been incorporated into their classroom. This teacher did not express any competing practices or barriers:

> "Um, well this year after that incident, I did practice restorative circles more often. They gave us those cards with the questions on them and that was helpful. And then also, um affective statements. Having the students, you know instead of me talking to them, I just let them, not vent completely, just how they are feeling and that also works because a lot of times these kids feel like they do not have a voice." (T35, RP is incorporated into classroom)

This teacher provided a concrete example of a time when RP was incorporated into her classroom. From the interviews, a majority of teachers provided examples of RP use with little or no evidence of interference from other programs, administrators, or time constraints.

A teacher who mentioned other competing practices or felt there were potential obstacles for RP use would receive a rating of 1. From the interviews, 18 teachers mentioned other competing practices (n = 18; 36.7%). These potential barriers and competing practices included:

> "The time definitely, as I'm sure you're aware being a professor, there's a lot of demands of what needs to be done every day and deadlines and 'do it…do it with a smile on your face' (laughter) so it's difficult to get that done when you have a curriculum, when you have test schedules, pullouts, and da-da-da. The list goes on. So it's difficult to do [RP]." (T24, potential competing practices and barriers)

> "There was an incident with one of my girls that I tried to do [a] conference and I was blocked completely by grown-ups and by the XX that was affected by her and by the support and [the administrator] tried his best to get us into the conference with the adults but they did not want to. So that was not addressed." (T1, potential competing practices and barriers)

These teachers were conceptualized as having potential barriers because amidst the challenges provided, it appeared that RP could still be implemented. General references by teachers to time constraints or lack of buy-in from school staff might only potentially block RP implementation efforts.

A teacher that provided real or concrete competing practices or barriers received a rating of 2; it was the strength and magnitude of these competing practices and barriers that made RP a less feasible option. A minority of teachers fit this description for a rating of 2 (*n* = 11; 22.4%), and some examples of their responses are provided below:

> "Mhmm and not even so much with that, it's about you know the administration, if we come in doing circles and all sorts of stuff and not dealing with the content, we get written up. So they, I don't think they really took into consideration, you know, that part even though our administration is like 'oh yeah do practices' and we're doing practices and we get written up and we get fired and all sorts of, you know, things that go on top….so." (T45, real and concrete competing practices and barriers)

> "You're coming into a school that's had you know, that's had a history of violence and stabbings and what you call it, and fights, food fights and gang threats and stuff like that and you're coming in to them and saying well listen, you got to put them in circles. You know, and put them in a circle, it's just something magical and healing about it and it kind of sounds like this new wave kind of like existential garbage and that's the way it's going to come off to them. Them as professionals should understand that's the way it would coming off to them and be already prepared to battle that. It kind of felt like they were like, with the three PD sessions kind of like, yeah well you know, we know some of you guys don't like it but it doesn't really matter because we're kind of here to stay. Here are some questions um, you know. You're in that first stage so you're always going to doubt it but then you're going to see the effects. Well, where are the effects? Eighty percent of the teachers are leaving from, you know they're here, you know the school has improved but there's no evidence to support that it's necessarily because of restorative circles. It may have been from the eighty percent of the staff that was let go." (T32, real and concrete competing practices and barriers)

> "Here though it's you know we have classes with 20 something kids. How are you gonna do a circle?" (T51, real and concrete competing practices and barriers)

While a majority of teachers felt that RP was incorporated into their classroom routine, many real and concrete barriers remained to make it less feasible for RP implementation. The real

threat of losing one's job and mixed messages of whether or not to use RP might have contributed to more variable implementation.

Correlations. Pearson correlations were computed for all control, independent, and dependent variables from the teacher survey and teacher interview data and are reported in Table 5. Pearson correlations are in the expected directions. RP self-efficacy ($r = .60$, $p < .01$), RP administrative support ($r = .50$, $p < .01$), and lack of competing interests ($r = .38$, $p < .01$) were positively associated with teacher's score on the RP Implementation scale. These positive relationships suggest that teachers rating themselves as having higher RP self-efficacy, higher levels of administrative support in RP implementation, and fewer competing interests tended to report implementing RP more frequently. Thus, 3 of the 4 constructs of interest (e.g., RP self-efficacy, administrative support, feasibility) were positively related to teacher's self-reported adherence to the RP Essential Elements. Another key finding was the positive association that showed a relationship between teachers' reported positive experience with RP and total circles/conferences run ($r = .36$, $p < .05$). This finding suggests teachers reporting more positive experiences with RP tended to implement more RP circles/conferences.

A handful of additional significant correlations are worth noting: Teacher race (e.g., being Black or Latino versus Asian, White, or Other) was positively related with the teacher RP implementation scale ($r = .30$, $p = .03$). This positive association suggests that those who identified their race as Latino or Black also reported higher levels of RP implementation in the self-report survey relative to those identifying as Asian, White, or Other. Number of days of training was positively correlated with total circles run ($r = .38$, $p < .01$), RP self-efficacy ($r = .45$, $p < .01$), RP administrative support ($r = .48$, $p < .01$), and level of competing interests ($r = .32$, $p < .05$). These positive relationships suggest that as teachers received more days of training, they exposed their classes to more RP circles/conferences, rated themselves

as more self-efficacious with regard to RP, reported higher levels of administrative support for RP use, and a lack of competing interests that interfered with their RP implementation.

Table 6

Association Among Teacher Characteristics, Teacher Reports, and RP Use

		1	2	3	4	5	6	7	8	9	10
1.	Years in role	____	-0.17	0.09	-0.15	-0.09	-0.06	-0.04	0.12	0.17	-0.09
2.	Gender		____	0.06	0.16	0.06	0.14	0.05	-0.10	0.11	0.07
3.	Race			____	0.26	0.30*	0.13	0.26	0.22	0.20	0.12
4.	Days Training				____	0.21	0.38**	0.45**	0.48**	0.24	0.32**
5.	Teacher RP Implementation					____	0.25	0.60**	0.50**	0.27	0.38**
6.	Total Circles/ Conferences Run						____	0.28	0.05	0.36*	0.20
7.	RP Self-efficacy							_____	0.64**	0.50**	0.55**
8.	RP Administrative Support								_____	0.34*	0.51**
9.	Positive Experience with RP									_____	0.68**
10.	Lack of Competing Interests										_____

Note: $p < .05$, *; $p < .01$, **. Race (1 = Latino or Black; 0 = White, Asian, or Other); Gender (1 = Male; 2 = Female).

Multiple regressions. For the series of regression models, blocks were entered in succession to understand the unique variance of predictors for RP adherence and use above and beyond teacher demographics (e.g., gender, race, and years in role entered in block 1). Table 7 and Table 8 used days in trainings (block 2), then positive experiences with RP, RP administrative support, RP self-efficacy, and lack of competing interests (block 3) as predictors of interest.

Table 7 presents predictors of the number of RP circles and conferences teachers implemented. Model 1 shows that no teacher demographics characteristics were related to the number of implemented circles/conferences. Model 2 shows that when teachers participated in more days of IIRP training, they tended to implement more circles/conferences ($\beta = 0.41$, $p = 0.02$). Training days explained 15% of the variance in the total implemented circles/conferences as reported by teachers. Model 3 shows that training days remained a

significant predictor (β = 0.58, p = 0.003) when entering four other possible contributors to total circle/conference use (positive experiences with RP, administrative support for RP, RP self-efficacy, and lack of competing interests). Yet, only one of these four contributors approached significance, positive experiences with RP (β = 0.41, p = 0.06). Together, these contributors explained a substantial amount of variance (21.1%) in implemented circles/conferences.

Table 8 presents predictors of teacher implementation of RP Essential Elements in their classroom. Model 4 shows that teacher race was related to the Teacher RP Implementation scale (β = 0.37, p = 0.02). Teachers who identified themselves as Black or Latino tended to rate themselves as implementing more RP Essential Elements than their White, Asian, or Other colleagues. In Model 4, teacher demographic characteristics explained 15.1% of the variance on the Teacher RP Implementation scale. Model 5 demonstrated that teacher race (being Black or Latino) remained a significant predictor of Teacher RP Implementation when days of training were included in the regression (β = 0.34, p = 0.04). In Model 6, being Black or Latino was no longer a significant predictor of adherence to RP Essential Elements (β = 0.23, p = 0.12). Yet, when entering four other possible contributors (positive experiences with RP, administrative support for RP, RP self-efficacy, and lack of competing interests), RP self-efficacy emerged as significantly related to the teacher rating on the Teacher RP Implementation scale (β = 0.48, p = 0.03). The more self-efficacious a teacher perceived him or herself with RP, the more likely he or she was to rate himself or herself as adhering to more of the RP Essential Elements on the Teacher RP Implementation scale. The four contributors jointly explained a considerable amount of unique variance of the Teacher RP Implementation scale (27.9%).

Table 7

Multiple Regression Results Predicting Total RP Circles/Conferences

	Predicting variables for Total RP Circles/Conferences		
	Model 1	Model 2	Model 3
	Total Circles/Conferences	Total Circles/Conferences	Total Circles/Conferences
	β	β	β
Gender	0.11	0.06	-0.05
Race	0.11	0.02	-0.03
Years in Role	-0.06	0.01	0.02
Days in Training		0.41*	0.58**
Positive Experience with RP			0.41^{+}
RP Administrative Support			-0.43
Self-efficacy			-0.10
Lack of Competing Interests			0.13
ΔR^2	0.03	0.15*	0.21*

Note: *p <.05, **p <.01, ^{+}p <.07.

Table 8

Multiple Regression Results Predicting Teacher RP Implementation

	Predicting variables for Teacher RP Implementation		
	Model 4	Model 5	Model 6
	Teacher RP Implementation	Teacher RP Implementation	Teacher RP Implementation
	β	β	β
Gender	-0.04	-0.06	0.05
Race	0.37*	0.34*	0.23
Years in Role	-0.16	-0.14	-0.08
Days in Training		0.15	-0.23
Positive Experience with RP			-0.22
RP Administrative Support			0.17
Self-efficacy			0.48*
Lack of Competing Interests			0.27
ΔR^2	0.15	0.02	0.28**

Note: *p <.05, **p <.01

Study 2: Teacher Reports of RP Implementation in Year Two

Research Question 1: Did teacher RP adherence, RP use, RP self-efficacy, and RP administrative support decline from year one to year two of school-wide RP implementation?

Hypothesis 1:

It was hypothesized that all fidelity and self-report measures would decline in the second year of RP implementation. A majority of the teachers in the sample would demonstrate these declines while a minority may increase or maintain their RP implementation efforts.

Research Question 2: Was there a cohort of teachers who retained high RP use across the two years of school-wide RP implementation?

Hypothesis 2:

Amidst an overall trend of decline, it was anticipated that a small cohort of teachers would report high RP use in year one and year two of school-wide RP implementation.

Methods

Participating schools

As year two of RP implementation began, the schools experienced tremendous turmoil. The principals at each high school were transferred or were terminated and most fellow administrators were moved as well. Ties to the RP training group were discontinued under the new administration. Thus, the IIRP did not implement year two consultation days or follow-up training with staff. Moreover, some basic organizational features broke down. For the example, in the first few weeks of the school year, no student schedules were issued for the students and teachers. In addition, many teachers were laid off from their teaching jobs after the June 2014 interviews.

Teacher participants

Given that ties to the RP consultants and training were discontinued in year two of the current study, the number of teachers continuing to implement the program dropped considerably. In fact, of the original sample of teachers that were recruited in year one (N = 51), only 45.10% of this sample of teachers remained in the study for year two (N = 23). Drop out from the study was partially due to the teachers not being employed by the two schools in the second year of RP implementation (n = 10, 35.71%). That said, 12 teachers who remained at both schools opted not to continue their participation in the study (n = 12, 42.86%). The range of reasons was not systematically gathered, yet some teachers mentioned that RP did not appear to be a priority in the school any more. A smaller group of teachers wanted to participate in the study, but dropped out after they were unable to obtain parent/guardian consent from their students in order to remain in the study (n = 6, 21.43%).

Procedures

Members of the Rutgers research team collected data from teachers already consented in Study 1. Study 2 data was collected in the winter and spring of 2015.

Measures

In Study 2, a handful of Study 1 survey scales (Appendix C) were re-administered to continuing teachers. These scales included: the RP Self Efficacy Scale, RP Administrative Support Scale, and Teacher RP Implementation Scale. In addition, the responses to the three questions about frequency of circles/conferences run (e.g., proactive circles, responsive circles, restorative conferences) were aggregated to create a total circle/conference variable for respondents. Similar to Study 1, all scales demonstrated good internal consistency with alphas ranging from .78 to .84.

Data Analysis Plan

Descriptive statistics and *t*-tests were run to examine shifts in RP adherence, RP use, RP administrative support, and RP self-efficacy over the two years of implementation. One teacher respondent was an extreme outlier in year two according to statistical analyses. Thus, this teacher was removed from comparing circles in a paired samples *t*-test so as not to skew mean comparisons. While this teacher's score could not be used to get an accurate measure for mean comparisons, his score could be used to categorize his perception that he was using RP circles and conferences at high rates.

Analysis of covariance was run to covary time 2 RP administrative support and compare two groups of teachers who differed in their circle/conference use in the first year of the program. This analysis was done to see how total circles/conferences run in Study 2 were related to RP administrative support.

Results

Research Question 1: Did teacher RP adherence, RP use, RP self-efficacy, and RP administrative support decline from year one to year two of school-wide RP implementation?

Table 9 shows the results of a paired samples t-test run to compare variables (total circles/conferences run, teacher RP implementation, RP self-efficacy, and administrative support) from year one to year two. Paired samples *t*-tests examine intra-individual changes, meaning they identify whether change occurred for individual teachers from their year one reports to their year two reports. As expected, the *t*-test for total circles/conferences run ($t = 1.38$, $p < .01$) demonstrated the mean total number of circles/conferences differed significantly from year one ($M = 10.64$) to year two ($M = 7.45$) across the sample. On average, teachers reduced the number of circles and conferences they implemented in year two of RP implementation. Unexpectedly, the *t*-test for the teacher RP implementation scale ($t = -2.41$, $p < .05$) showed that adherence to RP Essential Elements increased slightly amongst the sample from year one ($M = 3.53$) to year two ($M = 3.83$). In the second year, teachers, on average, reported they *often* adhered to using RP principles ($M = 3.83$). According to the paired samples *t*-test ($t = -5.91$, $p < .01$), teachers were more self-efficacious in their RP implementation in year two ($M = 3.28$) than in year one ($M = 2.78$). Teachers reported RP administrative support was significantly lower in year two ($M = 2.34$) than in year one ($M = 2.97$; $t = 3.65$, $p < .01$).

Table 9

Paired samples t-test of teacher RP fidelity and perceptions at year one and year two

Variable	Time 1			Time 2				
	M	*SD*	*n*	*M*	*SD*	*n*	*t*	*df*
Total Circles/Conferences Run	10.64	11.03	22	7.45	5.40	22	1.38**	21
Teacher RP Implementation	3.53	.72	23	3.83	.62	23	-2.41*	22
RP Self-efficacy	2.78	.60	23	3.28	.49	23	-5.91**	22
RP Administrative Support	2.97	.71	22	2.34	.75	22	3.65**	21

Note: $p < .05$, *; $p < .01$, **.

Research Question 2: Was there a cohort of teachers who retained high RP use across the two years of school-wide RP implementation?

For the 23 teachers that remained in the study for year two, Table 10 compares total circles/conferences run in year one and total circles/conferences run in year two as a categorization of RP use. If a teacher conducted more than 9 circles, the teacher was classified as having high RP use. This number was derived from the mean of circles run at time point 1 ($M = 9.04$) for the entire sample of teachers in year one ($N = 51$) with no outliers. Of this smaller sample in year two ($N = 23$), those above the mean were classified as high RP users while those below the mean were classified as low RP users. A minority of the sample, ($n = 7$; 30.43%), had high RP use in both years of the study. Many of the teachers ($n = 10$; 43.48%) were classified as having low RP use in both years.

Table 10

Classifying teacher RP use by their total circles/conferences implemented

		High Versus Low For Year 2	
		Low RP Use	High RP Use
High Versus Low For Year 1	Low RP Use	10	2
	High RP Use	4	7

Note: High RP use was classified as a teacher implementing more than 9.04 circles while low RP use was classified as a teacher implementing less than this cut-point.

An ANCOVA was run accounting for RP administrative support in the second year of RP implementation. This was done to ascertain if teachers remained high RP users despite declines in their perception of RP administrative support. Significant differences in RP use held when controlling for RP administrative support, F (1, 20) = 40.98, p <.001. Accounting for RP administrative support, those who were low RP users solely in year two (n = 14) averaged about 4 circles/conferences (M = 4.05). Those teachers who were high RP users in both years (n = 7), averaged almost 13 circles/conferences in year two (M = 12.77) when accounting for RP administrative support (M = 2.35). In other words, no matter whether year one high implementers perceived higher or lower RP administrative support in year two, they still tended to implement numerous circles/conferences. Interestingly, those classified as high RP users in year one and year two reported lower levels on average of RP administrative support (M = 1.93) than the low RP users at time two (M = 2.57). In other words, the high RP users in both years endorsed that they somewhat disagreed with items measuring the support of administration in RP efforts while the low RP users in year two endorsed that they somewhat agreed with the support of administration in RP efforts.

Discussion

Summary

Study 1 and Study 2 examined RP implementation with a sample of teachers working in two low income, predominantly ethnic minority high schools. Together, the studies offer new knowledge about factors associated with RP implementation and patterns in RP use and adherence across two years.

Corroborating prior studies (Gregory et al., 2014; McClusky et al., 2008), Study 1 showed considerable variability in RP adherence and use with teachers who had participated in an average of 3 days of training. Despite considerable district resources invested in training, 66.70% of teachers had low RP use in year one of RP implementation, whereby they

ran zero to eight total circles and conferences across 185 days of the school year. That said, self-reported RP adherence was higher as 52.9% of teachers endorsed that they adhered to all of the RP Essential Elements measured. The slightly differing findings based on two fidelity measures (RP use versus RP adherence) raises questions about the most reliable way to solicit teacher report about their practices. In this study, numerical reporting and likert scale responses solicited different information.

Study 1 also examined four possible correlates of RP implementation (i.e., RP administrative support, RP self-efficacy, positive experiences with RP, and RP feasibility). Multiple regression models showed RP training days (at the $p < .01$ level) and positive experiences with RP (at the $p < .10$ level) were significant predictors of total circles and conferences run (RP use). Teachers who participated in more RP training and those who had a positive experience with RP tended to implement more circles and conferences across the school year. Findings also showed that teachers reporting they had a higher sense of RP self-efficacy tended to use more of the RP Essential Elements. Unexpectedly, the regression results found that RP feasibility and RP administrative support were not significant predictors, despite their significant associations with RP adherence as shown in Pearson correlations. These finding suggest RP administrative support, RP self-efficacy, positive experiences with RP, and RP feasibility remain important constructs to study in future research.

As expected in Study 2, teachers reported a decline in their use of circles/conferences in year two versus year one. This may not be surprising given that reports of RP administrative support also declined from year one to year two and there were considerable changes in administration. Ties to the IIRP were dropped in year two as well. That said, analyses showed several unexpected findings. For individual teachers, RP self-efficacy and teacher RP implementation were significantly higher in year two, relative to their year one

reports. It is unknown why this pattern emerged. Also noteworthy was that despite lower administrative support in year two for RP, a small group of teachers retained high RP use as established by Study 1 (running more than 9 circles/conferences in a school year). This suggests a degree of resiliency amongst a subset of seven teachers in year two of RP implementation.

Building on Positive Experiences with RP

An unexpected finding was that total days of RP training predicted the total circles/conferences facilitated by the teachers. This finding suggests that schools and the IIRP should work together to ensure teachers attend all the training days provided. Yet, solely focusing on *attendance* to the trainings may not be sufficient. There might be value in ascertaining the degree of teacher engagement, knowledge gains, and skill acquisition in the training itself.

Future research might examine teachers' receptivity to and active engagement in the IIRP training, given it likely has implications for later RP implementation. For example, despite having been trained for numerous days to use RP Elements, including using circles effectively, a sizeable percentage of the teachers (29.40%) reported never having implemented proactive, community-building circles in their classrooms. By definition, proactive circles were considered "everyday" components of the RP intervention. Many teachers also attended trainings on leading restorative conferences. Yet, close to two-thirds (58.80%) of the teachers reported never having led such a conference. From these quantifiable discrepancies of actual RP use after attending RP training on the intervention, it appears that there might be some limitations to the trainings. For instance, teachers might need greater implementation supports following training. Especially, as it was observed in these schools that administrative support for RP significantly dropped in year two of implementation. Also, potential knowledge gains from the training might not generalize to

utilization of new practices or changes in teacher behavior. For these reasons, programs like RP cannot solely rely on workshops to create change in schools. Problems arise when resources to programs are not well utilized. As is often the case, school districts invest a majority of resources in upfront trainings, but fewer resources are allocated for continued, supported practice throughout the school year.

Learning from Positive Experiences

Results from interviews highlighted the degree of positive experiences teachers had with RP as well as their perceptions of competing practices and barriers to using RP. Coders were able to reliably rate positive experiences with RP and RP feasibility. It might be the case that the positive experiences reinforced teachers' RP use, thereby increasing it over time. Given the design of Study 1, such a feedback loop could not be detected. However, a female history teacher provided a specific example of a positive experience that led to her continued use of RP. In her interview, she described how a positive experience led to her implementing RP more often:

> "There's one class the only class that, they're about thirteen students, give or take, and I did restorative circles every day in that class, and there was one student in particular who said, 'Ms. [XX],' he was very low, very low. He had an IEP, and he wasn't a high achiever, he wasn't very motivated either, so he's like, 'Ms. [XX], please keep doing the circles because it really works.' And this is surprising coming from, especially from him who's not interested, talkative, not a high achiever, you know, he, in all his classes he was doing pretty badly, and he has an IEP so, to me when he said that, 'Please do the circles,' it was very important that he said that."

In the first section of this quote, this female teacher describes that the value of circles arose from an unexpected source. Many teachers described students with challenging behaviors. Yet, this was one of the few examples when one particular student provided some of the initial momentum for continued RP use. She continued:

> I was like, 'Wow, ok, this is definitely working.' It was not a perfect circle. There were a lot of classroom management issues. And if I learned about restorative circles, like, September 1st, and I knew what the class would turn out to be, I would use them

> from day one, because it set the tone and when I started doing restorative circles beginning when [an IIRP consultant] came in, she's the one who set the, she showed me how to do it…we spent the whole period doing it in class because she asked certain classes if she could do those sort of circles, so from that day on, it was just, the day before spring break, she came in, we did the restorative circle, it was a proactive circle, and she was asking them to use affirmative statements, and they all were basically diagnosing, making a diagnosis of the classroom, what's going right, what's going wrong, they gave me advice, the students."

To add to an already specific positive experience with RP, this teacher has voiced the importance of having an IIRP consultant help her run circles. Not all teachers had this opportunity, but this teacher benefited from both an unexpected request from a challenging student to use circles and direct support for her implementation efforts. She goes on to suggest a possible positive reinforcement cycle:

> "And from that point on, they went more on a hiatus, they went on spring break. They came back, I started using them almost every single day, and she told me to instruct from them, and it was just better instruction because everyone was doing the work because everyone was in the circle, everyone was pretty much forced to do the work because we were all sitting in the circle, we were all doing it together." (T29)

This teacher ultimately demonstrated that multiple positive experiences with RP and its feasibility with her lessons increased the frequency of her total circle use.

Fortunately, a vast majority of teachers (83.67%) reported at least some positive experiences with RP. This finding suggests promise for future RP implementation in highly stressed urban school settings. It remains unknown when these positive experiences need to occur to sustain implementation efforts. When and how to maximize the opportunity for positive experiences remains open to question. Do positive experiences need to occur immediately after training or during the training itself? Do positive experiences need to happen one week or a month after RP is implemented school-wide? Behavioral principles would suggest that early and frequent positive experiences would increase the likelihood of adoption of an innovation. Future research will need to establish a stronger evidence base to

clarify the degree to which positive experiences and positive reinforcement are significant contributors to RP implementation.

Whereas approximately two-fifths of interviewed teachers (*n* = 20; 40.82%) lacked competing practices or barriers to their RP implementation efforts, a majority of teachers (*n* = 29; 59.18%), conveyed at least some perceived barriers. For example, there were real barriers to implementation when competing gang members were in the same class and asked to participate in a circle. The real or perceived threat of losing one's job or getting written up for using RP was enough for other teachers not to use the intervention. A male science teacher explained how the rollout of RP itself contributed to him feeling it was not a feasible practice in his classrooms:

> "It's like they inherently expect the teacher to somehow miraculously pull this off and it's like we're given no guidelines on how to get to this point, it's more like just form a circle and start talking, the only time this would work out, it's like, the only class that I could potentially see this working out is my second period." (T43)

Elias et al. (2003) made the point of how stressful it can be for teachers to fit such programs into their busy reform schedules in urban schools that come under scrutiny. Feasibility encompasses many components for whether or not a teacher decides if an intervention like RP is practical. In Study 1, the correlations showed that lack of competing interests and barriers was positively correlated with administrative support, RP self-efficacy, and positive experiences with RP. This suggests that these constructs deserve further examination to understand the degree to which RP is perceived as feasible for teachers.

RP Self-Efficacy

In Study 1, teacher RP self-efficacy was consistently a significant predictor of teacher RP implementation relative to the other constructs of interest (e.g., RP administrative support). RP Self-efficacy was conceptualized as how confident and equipped a teacher felt in using RP in his or her classroom after training from the IIRP. Contrary to a hypothesized decline from year one to year two, RP self-efficacy significantly increased in year two of the

study. This finding suggests a possible practice effect that occurred between the two years of the study. Bandura's cycle of self-efficacy (1997) would suggest that teacher behaviors of implementing the intervention were contributing to teacher's feeling more able in their implementation efforts and vis-versa. The importance of self-efficacy specific to RP implementation is a key finding because while much is known about ways to increase self-efficacy in teachers in general (Klassen et al., 2011; Tschannen-Moran et al., 1998; Tschannen-Moran & Woolfolk, 2001), less is known about ways to increase self-efficacy in utilizing specific innovations.

Administrative Support and the Concept of Resilient Implementers

In year one, there was a positive association with administrative support and teacher RP implementation (adherence to RP principles). This suggests that the more supported a teacher felt by administration, the more the teacher reported higher levels of adhering to RP principles. Yet, from year one to year two of RP implementation, teacher perceptions of administrative support with the intervention dropped significantly. On average in year one, teachers were responding that they somewhat agreed they received administrative support for RP ($M = 2.97$). In year two, on average, teachers responded that they somewhat disagreed they received administrative support for RP ($M = 2.34$). This might be due to the fact that many teachers were laid off or that there was major turnover in the administration from both schools. The administrative support scale at year two might have captured some of the chaos that was created by significant changes in the school climate.

Despite the attenuation of administrative support in year two, a small subset of teachers ($n = 7$) continued to report high use of RP across the two years of implementation. In a sense, these teachers were "resilient implementers." A teacher demonstrates his or her resilience by assessing the adverse situations and choosing effective ways to cope (Bobek, 2002; Egeland, Carlson, & Sroufe, 1993). Study 2 found that even when controlling for

perceptions of administrative support, a subsample of low RP users in the first year of RP implementation remained distinct from other teachers as high users in the both years of RP implementation. In other words, across the two years, these 7 teachers stood out as persisting in using RP frequently despite the shifts in administrative support. This truly reflects their resilience as implementers when program supports faltered and ended altogether (e.g., no more IIRP consultation or trainings).

Study Strengths

Study 1 and Study 2 had many strengths that increase confidence in their findings. First, Study 1 was a mixed-methods study. Both interviews and surveys were used to measure teacher perceptions, teacher adherence to RP Essential Elements, and teacher RP use. By using both qualitative and quantitative data in analyses, findings were strengthened and "brought to life" through teacher voice. Interviews also allowed for subtle comparisons and differentiation of teachers' views of RP. Second, multiple measures were used to ascertain the extent to which teachers reported RP implementation. It was shown that asking similar questions in different ways provided contrasting results. Teacher responses varied from likert scale items to count data to interviews. Multiple measures of implementation helped elucidate the complexity of teachers' self-reported use of an innovation. Third, the design of Study 2 facilitated following teachers over time. The longitudinal aspect of this study offered new insights despite the considerable sample attrition. Following teachers over time identified possible implementation resiliency in a minority of the sample. Finally, systematic and reliable coding of interviews was done. Facilitated by coder's use of the RP Teacher Coding Manual, high interrater reliability demonstrated that it is possible to systematically differentiate teachers based on their self-reported narratives about implementing a new program.

Study Limitations

A major limitation for both studies was low statistical power. The small samples (N = 51 in year one and N = 23 in year two) decreased the probability of detecting a statistically significant effect if one existed in the sample. Future research on RP should survey and interview more informants/practitioners in schools which are rolling out the program school-wide. While confidentiality and anonymity were stressed, it is possible that teachers felt they could not honestly critique RP or administration in surveys or interviews given the intensive teacher evaluation process in the school at the time.

Additionally, teacher attrition across the two years of the study likely affected the results in Study 2. It would be beneficial to know more about those teachers that left the study and what the specific reasons were for their exit. Because of the small number of participants in Study 2, some comparisons (ANCOVA) were done with unequal groups. Retaining more teachers in the study would have likely strengthened findings. Another limitation was mono-informant bias. Items of RP self-efficacy, teacher RP implementation, and total circles/conferences run were all survey items and could not account for the response bias of a particular teacher. It is possible that some teachers saw themselves more positively in their implementation efforts and then answered all questions positively while the same could be said for teachers that viewed themselves or the intervention more cynically. For instance, in interviews, some teachers were positive but vague and while coding of teacher interviews accounted for this, there could be a positive skew to the survey data that was undetected (e.g., teachers were more positive because a researcher was present). In other words, the study did not account for social desirability effects that might have contributed to teacher reports.

Other possibly important factors to implementation were not measured in this study. For instance, fit, student perception of RP, student report of teacher fidelity, and teacher "grit" are just a few concepts that could contribute to future studies. Fit can refer to the

appropriateness and acceptability of an implementation (Forman, 2015). For instance, Forman (2015) observes that those staff members who perceive the intervention to be consistent with their values will be the most likely to be consistent (e.g., adherence to the innovation) and enthusiastic in their use of an innovation. Future research on RP could examine how there may or may not be a philosophical fit for teachers with RP. Another limitation of this study was that only teachers were included and student voice was not incorporated in the results. Previous work by Gregory et al. (2014) examined student report of teacher RP implementation as a way to measure program implementation and differentiate high versus low RP use. Moreover, an outsider could rate RP use and quality from an instrument such as *RP-Observe* (Gregory et al., 2015). While not measured in this study, teacher grit could prove important for RP and implementation science. Robertson-Kraft and Duckworth (2014) found evidence that grit positively impacts teacher performance and was a robust predictor for teacher effectiveness and retention in the profession. Grit is the tendency of an individual to sustain persistence, perseverance, and passion for long-term goals (Robertson-Kraft & Duckworth, 2014). Given that grit is more stable than self-efficacy, it might be more robust and reliable as a predictor of RP implementation. Teachers with high levels of grit have been shown to be more effective in high stress, low income school districts (Robertson-Kraft & Duckworth, 2014).

Implications for Future RP Studies and Practice

Practitioners may choose to use a program like RP based on what they hear from colleagues or by researching interventions with an evidence-base. Forman (2015) points out how evidence-based interventions yield positive outcomes for clients, yet their use in school settings still remains low (Ennett et al., 2003). The current studies add to the evidence-base for RP, an intervention that has shown initial promise in reducing the racial discipline gap (Gregory et al., 2014). In schools today, when implementation efforts falter or are not

implemented as recommended, a program tends to be viewed as ineffective. For these reasons, RP need to be implemented with high fidelity to truly determine how effective it is as a program.

Researchers, practitioners, and those either receiving or conducting trainings may have questions about whether these trainings actually relate to later to shifts in practices. In many cases, these shifts in practices might not be detected for weeks, months, or years. Study 1 found that more days of training from the fall of year one related to higher teacher reports of RP adherence in June of year one. This suggests upfront training days do matter for later practice.

Research on the necessity of implementation supports and the deleterious effects associated with a lack of administrative support might suggest that if RP was no longer supported in the school, teachers would not keep up their practice. Yet, the subset of resilient implementers suggests some teachers latch on to an innovation, and on their own, they persevere. While a group of resilient implementers was identified, future studies and practice of RP would benefit from knowing what made them continue using the innovation. Also, there were two teachers that increased their number of circles and conferences from year one to year two. Future research could be done to understand what triggered this change in those teachers from low to high use, despite a lack of additional training and school level supports. Even in the initial stages of RP implementation, it might be helpful to learn what works and does not work from high RP users and those that have expressed positive experiences.

Little prior research has been done on reinforcing experiences of innovations. Study 1 sets the stage for future research on positive reinforcement by showing a link between teachers' reports of positive experiences and continued practice. There was a marked difference in implementation efforts by those teachers that provided specific examples of a class, student, or consultant that contributed to their perception of RP as positive for their

classroom and school. This highlights the need for early, reinforcing experiences to motivate teachers to persist in using an innovation like RP.

Factors such as days of training, positive experiences, and self-efficacy with RP proved useful in understanding associations with RP use and RP adherence. While administrative support and frequency of RP use declined in year two, RP self-efficacy and adherence to RP Essential Elements significantly increased against these odds. While this study added to the literature about the importance of factors related to the frequency of using an innovation, additional implementation supports are likely needed to ensure RP is implemented with quality.

In sum, RP implementation had its array of challenges in the two urban schools in the current studies. Findings suggest, however, that some teachers, against the odds, incorporate it into their practice. That RP "weathered the storm" of changing administrations suggests it may hold promise for eventually shifting teacher discipline practices away from zero tolerance approaches to behavior and toward a relationship-building and problem-solving approach.

Appendix A

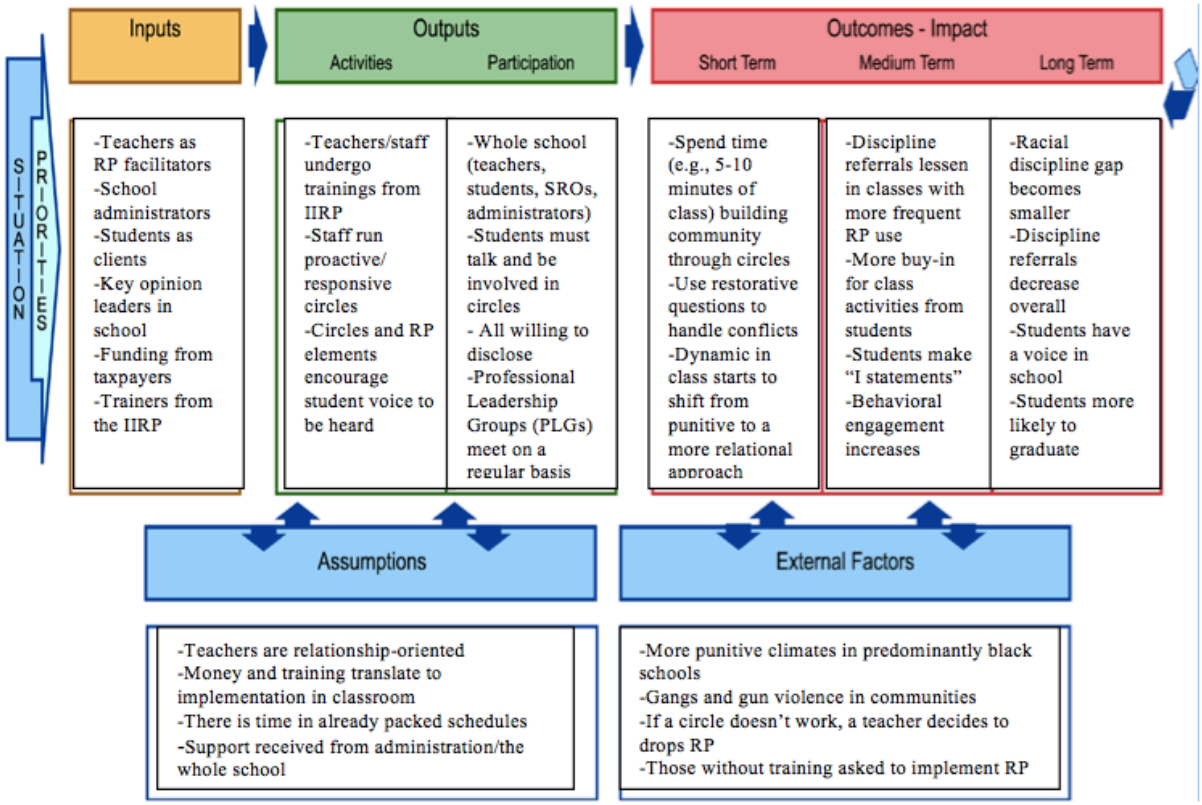
Defining the Eleven RP Essential Elements

Elements	Definition
1. Affective Statements	Students and staff express feelings
2. Fair Process	Everyone treated respectfully
3. Restorative Questions	Help identify who is harmed
4. Reintegrative Management of Shame	Listening to/acknowledging feelings of shamed person
5. Small Impromptu Conference	Questions/problem solving to resolve low-level incidents
6. Restorative Staff Community	Administration/staff model RP use
7. Proactive Circles	Everyday circles used to build community
8. Restorative Approach with Families	School uses RP with family interactions
9. Responsive Circles	Circles after a minor incident
10. Fundamental Hypothesis Understanding	Humans are happiest when those in authority do things *with* them, not do things *to* them
11. Restorative Conferences	Structured protocol for serious incidents

Note: Adapted from the International Institute of Restorative Practices (IIRP): Safer Saner Schools Program.

Appendix B

Logic Model for RP



Note: The outcomes of RP are measured in the short-term (immediate and early stages of implementation), medium-term (after one year of RP implementation), and long-term (two years or more of RP implementation.) The first intended outcome is that teachers and students spend time together (e.g., 5-10 minutes of class-time) building community through circles and use of RP elements. The teacher in the short term has a script of questions and tools to use when conflicts arise instead of resorting to an immediate exclusionary discipline referral. Additionally, the IIRP claims that shifts in the dynamic of the class start immediately by implementing RP; so, a paradigm shift from punitive discipline to a more relational approach is meant to be an outcome of the program in the short-term. In the medium-term, discipline referrals lessen in classes with more frequent RP use. The intended medium term outcomes also include more buy-in and observed engagement in class activities, and students will use more affective statements in their communication with each other and school staff (e.g., "I feel respected when the class listens to my needs," "I get upset when there are fights in my classroom"). The IIRP says that it takes two years to effectively implement RP, so this is used to gauge the beginning of the intended long-term outcomes of the program. In the long-term, the racial discipline gap is meant to become smaller through use of the program. Discipline referrals are meant to continue to decrease overall. An intended long-term outcome is that students have a voice in school and that they can generalize these skills to the community. Additionally, students are more likely to be invested in school and graduate.

Appendix C

Restorative Practices Implementation Survey

RP self-efficacy scale

Please indicate the degree to which the following statements are more or less true for you.

Strongly disagree = 1, Somewhat disagree = 2, Somewhat agree = 3, Strongly agree = 4

1. After training from the IIRP, I am able to integrate RP into my interactions with students.
2. I feel **confident** in using RP elements during my interactions with students.
3. I feel prepared to run **proactive** and/or **responsive** circles.
4. I feel prepared to run **restorative conferences**.

Administrative support scale

Please indicate the degree to which the following statements are more or less true for you.

Strongly disagree = 1, Somewhat disagree = 2, Somewhat agree = 3, Strongly agree = 4

1. There is a designated person in my school to consult with about RP.
2. I have had opportunities to discuss RP with other teachers and/or staff.
3. I know who to go to in my school if I need to refer students for a restorative conference.
4. It's clear to me that the use of RP is a mandate in this school.
5. RP is a top priority in my school.
6. I feel supported by school administrators when I use RP elements.

Teacher RP implementation scale

Please indicate how often you do the following:

Not at all = 1, Rarely = 2, Sometimes = 3, Often = 4, Always = 5

1. I actively encourage students to use affective statements to express how they have been impacted by others' behavior.

2. I ask the wrongdoer to identify who has been harmed and what harm has been done.
3. When addressing misbehavior between students, I structure the conversation using the restorative questions.
4. I ask students to take specific actions to repair the harm.
5. I use circles to provide opportunities for students to share feelings, ideas, and experiences.
6. I actively engage students about rules and ask for their input.

Total circles/conferences run

At this point in the school year, please indicate how many of the following you have led or co-facilitated:

All answers are numerical (e.g., 0 = no circles run, 10 = 10 circles run)

1. Proactive RP Circles
2. Responsive RP Circles
3. Restorative RP Conferences

Appendix D

RP TEACHER CODING MANUAL

Joshua Korth, Psy.M. and Anne Gregory, Ph.D.

Rutgers, The State University of New Jersey

Directions on how to use the manual: Some codes in this manual will be based on a three point scale (e.g., 0, 1, 2) while others will be dichotomous (e.g., 0 for Absence, 1 for Presence). Please highlight sections of dialogue and then, make a notation about which code the quote justifies. For instance, with the code *relational approach*, you would highlight a section of dialogue, which indicates the presence, absence, or neutrality of a relational approach in response to the question, "How do you build a sense of community in your classroom?" Then, you might may a note next to this section of dialogue saying "relational" or a possible number you would assign (e.g., relational 1?, relational 2, etc.) By referring to the criteria as well as potential examples of a 0, 1, or 2, you would then rate the teacher on this 3-point-scale based on highlighted sections of the interview as your justification. Make sure each highlighted section justifies your ultimate coding decision. Codes are global which means that you should average responses from the entire interview. Always use the coding manual in conjunction with the interview to properly fill out a coding sheet. In the case that a question was not asked eliciting dialogue for a certain code, you would indicate this by assigning a 99. This is to be used in extreme circumstances where no evidence can be found for the code.

Code 1: Relational Approach

- A relational approach to teaching is one in which community building and relationships between teachers and students are a priority. In order to determine which teachers have this approach versus a more traditional academic approach (e.g., teaching to the test), teachers were asked how they build a sense of community. When

using the coding scheme below, please also keep in mind the priority hierarchy which can help you distinguish between a 0, 1, or 2.

- This code is typically based on the first question of the interview, but you may use the entire interview for your rating if applicable.

Priority Hierarchy

A teacher that clearly emphasizes building relationships in their response will receive the highest rating (2). While examples of community building can be present for all teachers, any indication that this is not a priority moves the teacher down to a 1 as the highest potential code.

Coding Scheme

0- Minimal evidence of a relational approach or vague response to community building question.

1- Somewhat relational approach in the classroom. Evidence of some focus or a general focus on community/relationships in the classroom.

2- Highly indicative of a relational approach in the classroom. Teacher values community and his or her relationship with students as a priority. This could be achieved by a specific example of how the teacher builds community in response to the interviewer's question.

99- Not enough information provided.

Examples of 0's

"And when they come to me in my classroom, being it is a discipline room, then they are in there for a reason. It is not my fault that they are there. They must have done something wrong for them to be there." *Minimal evidence of a relational approach* (e.g., blaming students for being in this teacher's room, discipline trumps community)

Examples of 1's

"But I try not to be umm punitive all the time so that's where it, restorative may come into place because you want to have rewards for when to students are doing something positive, even if that's like an acknowledgement or you know something…" *Somewhat relational approach*

"Um, by first establishing uh modeling appropriate behavior, uh talking about what appropriate behavior is for our environment." *Emphasis on rules more than personal relationships*

Examples of 2's

"Something very off topic, or like even a movie that's coming up, or just a bit, a little bit of like a 'did that really happen?' but you know, just getting off topic and just having a laugh with the kids." *Clear value placed on relationships with students*

"Getting to know names, getting to know students on a personal level, um I know in general for their birthdays I would give them birthday presents and things like that and they would be surprised…" *Strong emphasis on relational approach*

"I think that I personally engage which each one of the students, and that's worked for me." *Strong emphasis on relational approach*

Code 2: Positive Experience with RP

- This code reflects the overall degree to which a teacher mentions using RP and having positive results while using these practices. This code can include statements that a teacher makes about RP school-wide or other teachers/administrators use of RP. Many teachers will mention challenges or problems with implementation, but this will be accounted for in another code, competing practices. In addition, please focus **globally** (e.g., experiences from many examples, not just the **best** example of positive RP experience) on dialogue indicating positive RP experiences before deciding on a particular rating.
- *Some questions that might indicate the global positive experience with RP include:* "Have you had a chance to use RP in your classroom," "If yes, what positive changes have you seen in your classroom," "Which RP principles (e.g., affective statements,

fair process, reintegrative management of shame) seem to be most useful in your classroom or school," "More generally, I want to learn about the use of RP in your school," "Please describe how: Administrators use RP; Other teachers use RP," and "Has RP created a difference in the school climate."

Coding Scheme

0- No clear mention of positive RP experiences, *extremely* vague mention of positive RP experiences, AND/OR mostly negative RP experiences. *Remember do not focus on overall RP challenges (e.g., school issues, competing programs/practices, barriers), only a specific mention of a negative RP experience can be incorporated in the coding.*

1- Somewhat vague positive RP experiences. The teacher may speak in generalities about positive experiences AND/OR *somewhat* vaguely about positive RP experiences on multiple occasions. In this case, the teacher may also mention some positive experiences in addition to some negative experiences, which makes him or her seem unsure of how positive the RP experience has been in the classroom.

2- Highly positive RP experiences. The teacher indicates a specific example in detail about how usage of RP was highly positive for his or her class AND/OR the teacher mentions multiple positive experiences he or she has had with RP.

99- Not enough information provided

Examples of 0's

"I tried one responsive circle after I felt like there was an incident and that was a horrible experience for me." *Mostly negative RP experience,* *see asterisk in example of 2's for more guidance

Examples of 1's

"…But the circles, we like them" *Somewhat vague positive comment*

"These practices are good for our students." *Somewhat vague positive comment*

"We have seen some positive effects. I would say yes, but I am not clear how much…" *Somewhat positive and speaks in generalities*

Examples of 2's

"By using the circles, the sense of community comes together, begins to come together" *Highly positive and more specific positive comment,* *must be accompanied with more positive comments like this for a 2.

"The good thing about RP is that it is really self-reflective." *Specific positive experience from RP use*

"Wow, ok, this [RP] is definitely working." *Positive experience from RP use,* *In a hierarchy of coding decisions, look for these statements over just comments saying "RP is good," "circles are good," etc.

Code 3: Frequency of RP use

- While frequency of RP use will also be measured quantitatively, this code accounts for the individual teacher's report of how often they have been using RP in their classroom.
- *Some questions that might indicate the frequency of RP use include: "*Have you had a chance to use RP in your classroom," "Which RP principles (e.g., affective statements, fair process, reintegrative management of shame) seem to be most useful in your classroom or school," "Has RP created a difference in the school climate," "Provide a specific situation that improved due to RP."

Coding Scheme

0- Teacher mentions they did not use RP in their classroom.

1- Teacher used RP somewhat frequently in their classroom (e.g., they mention using circles less than 8 times total).

2- The teacher mentions using RP highly frequently in their classroom (e.g., daily, weekly, monthly).

99- Not enough information provided

Examples of 0's

"A lot of us have been trying it but not going beyond like the trying portion, like mastering it kind of you know.

Like seeing how it works out but not really making it a routine." *Did not really use RP frequently*

Example of 1's

"But I did them a number of times, like I would say, three to five times throughout the year." *Somewhat frequent use and less than 8*, *vagueness in the estimate, not as frequent as a user receiving a 2

Examples of 2's

"I really like it right after we've been away for a week of even sometimes like 3 day weekends and what I do is when we come back it's a check-in." *More frequent use,* *evidence of potential monthly or regular use

"I started using them almost every single day." *Highly frequent user,* *definitive example of a 2.

Code 4: Competing Practices and Barriers

- Often in schools that have been asked to undergo reforms, multiple programs or initiatives will be asked of teachers in that particular calendar year. RP was not the only initiative introduced to the high schools in the current study, so there is often mention of other programs or practices that could potentially interfere with RP implementation.
- *Some questions that might indicate competing practices include:* "Are there any other practices or programs the give students a 'voice' in solving conflict or repairing harm," "More generally, I want to learn about the use of RP in your school," and "Has RP created a difference in the school climate?"

Coding Scheme

0- No mention of competing practices OR mention of competing practices but RP is emphasized for this teacher. For these teachers, RP seems to be a program/practice that has been incorporated into their classroom.

1- The teacher mentions one or two other competing practices and programs which are *potential* obstacles for RP use in his or her classroom.

2- The teacher mentions *real* or *concrete* competing practices AND/OR barriers AND/OR one specific competing practice/barrier that is very strong in magnitude and might be repeated by the teacher in the interview. The strength and magnitude of these competing practices and barriers make it appear that there is no opportunity for the teacher to incorporate RP in his or her teaching (e.g., no time with packed schedule).

99- Not enough information provided

Example of 0's

"It was just better instruction because everyone was doing the work because everyone was in the circle." *RP is incorporated into teaching, no mention of competing practices*

"Yeah its a lot being kinda thrown at you. Yeah but just seeing how it actually did work, I mean I don't want to

just toss it to the side." *Forging ahead with RP use*

Example of 1's

"The other model they're bringing in for us is that Kagen thing….they're big on this cooperative thing right now." *Mention of competing practice*, *To receive a 1 there must be some evidence of a potential barrier or conflict with RP use

Example of 2's

"The challenge is time." *Barrier, competing practices*

"With this being a renewed school, so many initiatives came our way. So it's hard to prioritize and you don't…" *Minimal opportunity to incorporate RP into schedule*

"Yeah, this school, well this school is not designed actually for, for big circles at this time I feel. For us to have a class or, or a period dedicated to a circle." *Barrier to using RP*

Code 5: Implementation Supports

- For a program such as RP to be successful, it needs to have support from the administration as well as the teachers. Depending on the teacher, he or she may have experienced a varying level of support in their implementation efforts. In addition, support from IIRP consultants may be factored into this code.
- *Some questions that might indicate the level of implementation supports are*: "What kind of *support* have you received to help you use RP *throughout* the school year" and "Is there a person you can go to with questions about RP?"

Coding Scheme

0- The teacher does not experience support in his or her RP use. In other words, there may be minimal support.

1- The teacher experiences a mix of support from teachers and administration. There may be some indications of support from at least one of the following: IIRP consultants, other teachers, or administration. The teacher may also indicate instances of feeling unsupported by those within the schools. In other words, there may be a few supports for this teacher.

2- The teacher provides a specific example or source of support, which he or she has received in their RP use along with a positive response to the interviewer's probe on support. In other words, there is good enough support to lead to continued practice.

99- Not enough information provided

Examples of 0's

"Support? Hmm… minimal." *Minimal support*

"I mean I didn't have any direct, I feel like I didn't have any direct support." *Does not experience support*

Examples of 1's

One possible support (e.g., IIRP consultants, other teachers, or administration) is mentioned but the teacher indicates not feeling completely supported in their use of RP.

Examples of 2's

"Because we have PLC's every Tuesday and with the PLC's that's when we come together with our administrator Mr. E and we discuss um the different restorative practices, all kind of different things that we may have to implement based on our professional development classes that we have taken so far." *Specific example of support that would lead to improved practice*

"So it's good to see that I'm not alone in this." *Positive response to interviewer probe*

"Oh our vice principal, he's our go to…" *Specific source of support*

Code 6: Cultural Match

- This code indicates whether a teacher believes RP is culturally in synch with his or her students or not.
- This code is based on a 2-point scale (e.g., 0-disagree, 1-agree).
- *The question that indicates a cultural match is:* "Is RP culturally in sync with your students, for example in relation to their race, ethnicity, gender, immigrant status?"

Coding Scheme

0- The teacher disagrees with this statement or provides an example of why RP is not culturally in sync with students.

1- The teacher agrees with this statement or provides an example of why RP is culturally in sync with students.

99- Not enough information provided

Examples of 0's

"No. You know, when I have my classroom, you can even tell they are from different countries, because they don't bring the issue like, "Oh, I'm Dominican, I'm American," no, they don't do that. They don't." *The teacher disagrees with statement that RP is culturally in sync with students.*

Examples of 1's

"Yeah, I think it works." *Agrees with statement that RP is culturally in sync with students.*

Code 7: Desire for Continued RP Use in School

- This code indicates if a teacher has made a statement on hopes, wishes, or desires for the continuation of RP in their school. It might indicate some of the challenges with current RP use and their desire for improvements in future implementation.
- This code is based on a 2-point scale (e.g., 0-absence, 1-presence).
- *Some questions that indicate a desire for continued RP use include:* "How could community building improve in this school?" and "Hopes for RP use?"

Coding Scheme

0- Absence of a statement endorsing a teacher's desire for continued RP use in the school.

1- Presence of a statement endorsing a teacher's desire for continued RP use in the school.

99- Not enough information provided

Examples of 1's

"I'm hoping that a program like this could become, could become very successful. Because it will impact our population in a positive way." *Desire for continued RP use in the school*

Date: _____/_____/_____ Teacher number: _____ Name of coder: _____

RP Teacher Coding
Relational approach – (0 = minimal evidence/vague response; 1 = somewhat relational approach; 2 = relational approach, highly value community and students; 99 = not enough information)
Rating: ____ Justification/comments:
Positive experience with RP – (0 = negative experiences, no mention of positive, or extremely vague; 1 = positive but general or somewhat vague or mix of negative and positive; 2 = positive with specifics or multiple positive experiences; 99 = not enough information)
Rating: _____ Justification/comments:
Frequency of RP use – (0 = no use of circles; 1 = less than 8; 2 = daily, weekly, or monthly use; 99 = not enough information)
Rating: _____ Justification/comments:
Competing practices and barriers – (0 = no mention or RP is incorporated into class; 1 = potential competing practices/barriers to using RP; 2 = real or concrete competing practices/barriers to using RP; 99 = not enough information)
Rating: _____ Justification/comments:
Implementation supports – (0 = minimal supports; 1 = a few supports; 2 = good supports leading to improved practice; 99 = not enough information)
Rating: _____ Justification/comments:

Cultural match – (0 = disagreed; 1 = agreed; 99 = question not asked)
Rating: _____ Justification/comments:
Desire for continued RP use in school – (0 = absence; 1 = presence; 99 = not enough information)
Rating: _____ Justification/comments:

Appendix E

Semi-Structured Interview

The interview will be audio-recorded, but whatever you say will only be reported in the aggregate with no identifying information. All proper names will be removed from transcripts. Please be open and honest. Feel free to share all sides of your experience, as there are no right or wrong answers. Do you have any questions before we begin?

- Bolded questions were required and other questions were prompts for further discussion surrounding implementation.

1. **How do you build a sense of community in your classroom?**
   - **a.** What are some strengths of your classroom?
   - **b.** How could community building improve in this school?

2. **Have you had a chance to use RP in your classroom?**
   - **a. If yes, what positive changes have you seen in your classroom?**
     - **i.** Challenges?
     - **ii.** Which RP principles (e.g., affective statements, fair process, reintegrative management of shame) seem to be most useful in your classroom or school? Least useful?
   - **b. If no, describe obstacles that got in the way.**
     - **i.** Think back to the initial introduction of RP, did it seem like something you were able to do?
     - **ii.** Was the material accessible?

3. **Does RP have the potential to improve your students' social-emotional learning skills** (e.g., perspective taking, empathy, problem solving)?

4. **What was your experience of the Restorative Practices (RP) training in preparing you to use RP in your classroom?**
   - **a.** Any prior training similar to RP?

5. **What kind of *support* have you received to help you use RP *throughout* the school year?**
   - **a. Are you involved in a Professional Learning Group (PLG)?**
     - **i. If yes,** please describe some of your experiences in the PLG.
   - **b. Is there a person you can go to with questions about RP?**

6. **More generally, I want to learn about the use of RP in your school. Please describe how:**
   - **a. Administrators** use RP
   - **b. Other teachers** use RP
   - **c.** Hopes for RP use?
   - **d.** Fears?

7. **Has RP created a difference in the school climate?**
   - **a.** Provide a specific situation that improved to due to RP
   - **b.** How was it restorative?
   - **c.** Reduced in- or out-of-school suspensions?

8. **Is RP culturally in sync with your students, for example in relation to their race, ethnicity, gender, immigrant status?**
   - **a.** *(Other examples are Latino, African American, ELL status)*
   - **b.** Align with their upbringing?

9. **Are there any other practices or programs the give students a "voice" in solving conflict or repairing harm?**
   - **a.** Other forums or mediation?

References

Aarons, G. A., Ehrhart, M. G., Farahnak, L. R., & Sklar, M. (2014). Aligning leadership across systems and organizations to develop a strategic climate for evidence-based practice implementation. *Annual Review Of Public Health*, *35,* 255-274.

Aarons, G. A., Hurlburt, M., & Horwitz, S. M. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*, 4–23.

American Academy of Pediatrics. (2013). Policy statement: Out-of-school suspension and expulsion. *Pediatrics*, *131*, 1000–1007.

American Psychological Association Zero Tolerance Task Force. (2008). Are zero tolerance policies effective in the schools? An evidentiary review and recommendations. *American Psychologist*, *63*, 852–862.

Baker, A. (2014, February 16). Common core curriculum now has critics on the left. *The New York Times*. Retrieved from http://www.nytimes.com/2014/02/17/nyregion/new-york-early-champion-of-common-core-standards-joins-critics.html?_r=0

Balfanz, R., Byrnes, V., & Fox, J. (2015). Sent home and put off-track: The antecedents, disproportionalities, and consequences of being suspended in the ninth grade. In D. Losen (Ed.), *Closing the school discipline gap*. New York, NY: Teachers College Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Berman, P., McLaughlin, M., Bass, G., Pauly, E., & Zellman, G. (1977). *Federal programs supporting educational change: Vol. VII. Factors affecting implementation and continuation*. Santa Monica, CA: RAND.

Bill and Melinda Gates Foundation. (2011). *Learning about teaching: Initial findings from the measures of effective teaching project*. Bellevue, WA: Author. Retrieved from www.gatesfoundation.org/college- ready-education/Documents/preliminary- findings-research-paper.pdf

Bobek, B. L. (2002). Teacher resiliency: A key to career longevity. *Clearing House, 75,* 202.

Boccanfuso, C. & Kuhfield, M. 2011. *Multiple responses, promising results: Evidence-based nonpunitive alternatives to zero tolerance*. Washington, D.C.: Child Trends.

Borman, G. D., & Dowling, N. M. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research, 78*, 376-409.

Boyd, D., Grossman, P., Ing, M., Lankford, H., & Wyckoff, J. (2009). The influence of school administration on teacher retention decisions. Retrieved January 8, 2015 from http://www.teacherpolicyresearch.org/portals/1/pdfs/TeacherRetentionAdministrators22May2009.pdf.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2009) Who Leaves? Teacher attrition and student achievement. National Center for Analysis of Longitudinal Data in Education Research. Working Paper 23. Retrieved January 8, 2015 from http://www.teacherpolicyresearch.org/portals/1/pdfs/Do%20Effective%20Teachers%20Leave.pdf.

Brackett, M. A., Reyes, M. R., Rivers, S. E., Elbertson, N. A., & Salovey, P. (2012). Assessing teachers' beliefs about social and emotional learning. *Journal of Psychoeducational Assessment, 30,* 219-236.

Children's Defense Fund. (2013). *Protect children not guns 2013*. Retrieved from http://www.childrensdefense.org/library/protect-children-not-guns/protect-children-not-guns-2013.html

Cooper, A. & Smith, E. (2011). *Homicide trends in the United States, 1980-2008.* Retrieved from http://www.bjs.gov/content/pub/pdf/htus8008.pdf.

Costello, B., Wachtel, J., & Wachtel, T. (2010). *Restorative circles in schools: Building community and enhancing learning*. Bethlehem, PA: International Institute for Restorative Practices.

Cherniss, C. (1995). *Beyond burnout: Helping teachers, nurses, therapists, and lawyers recover from stress and disillusionment*. London: Routledge.

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, *18*, 23-45.

DeVoe, J. F., Peter, K., Noonan, M., Snyder, T. D., & Baum, K. (2005). *Indicators of school crime and safety: 2005*. U.S. Departments of Education and Justice. Washington, DC: U.S. Government Printing Office.

Diamond, J. (2012). *The world until yesterday: What can we learn from traditional societies?* New York, NY: Penguin Books.

Dolan, M. (1992). State takeover of a local district in New Jersey: A case study. New Brunswick, NJ: Consortium for Policy Research in Education.

Egeland, B., E. Carlson, and L. A. Sroufe. (1993). Resilience as process. *Development and Psychopathology, 5,* 17–28.

Elias, M. J., Klein, D. A., DeLuca, A., Smith, D. F., Fattal, L. F., Bento, A., & Leverett, L. (2006). Reflections on the laws of life. *Principal Leadership, 6*, 39-43.

Elias, M. J., & Leverett, L. (2011). Consultation to urban schools for improvements in academics and behavior: No alibis. No excuses. No exceptions. *Journal of Educational and Psychological Consultation, 21*, 28-45.

Elias, M. J., Zins, J. E., Graczyk, P. A. & Weissberg, R. P. (2003). Implementation, sustainability, and scaling-up of social-emotional and academic innovations in public schools. *School Psychology Review, 32,* 303-319.

Eslinger, J. C. (2014). Navigating between a rock and a hard place: Lessons from an urban school teacher. *Education & Urban Society*, *46*(2), 209-233.

Ennett, S. T., Ringwalt, C. L., Thorne, J., Rohrbach, L. A., Vincus, A., Simons- Rudolph, A., & Jones, S. (2003). A comparison of current practice in school-based substance use prevention programs with meta-analysis findings. *Prevention Science*, *4*, 1–14.

Fabelo, T., Thompson, M. D., Plotkin, M., Carmichael, D., Marchbanks, M. P., III, & Booth, E. A. (2011). *Breaking schools' rules: A statewide study of how school discipline relates to students' success and juvenile justice involvement*. New York, NY: Council of State Governments. Retrieved from http://justicecenter.csg.org/ resources/juveniles

Finn, J D., & Servoss, T. J. (2013). Misbehavior, suspensions, and security measures in high school: Racial/ethnic and gender differences. *Journal of Applied Research on Children: Informing Policy for Children at Risk, 5,* 11. Retrieved from http://digitalcommons.library.tmc.edu/childrenatrisk/vol5/iss2/11/?utm_source=digitalcommons.library.tmc.edu%2Fchildrenatrisk%2Fvol5%2Fiss2%2F11&utm_medium=PDF&utm_campaign=PDFCoverPages.

Fixsen, D. L., Naoom, S. F., Blasé, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature.* Tampa, FL: University of South Florida.

Forman, S. G. (2015). *Implementation of mental health programs in schools: A change agent's guide.* Washington, D.C.: American Psychological Association.

Forman, S. G., Olin, S., Hoagwood, K., Crowe, M., & Saka, N. (2009). Evidence-based intervention in schools: Developers' views of implementation barriers and facilitators. *School Mental Health, 1*, 26–36.

Forman, S. G., & Selman, J. S. (2011). Systems-based service delivery in school psychology. In Bray, M.A. and Kehle, T. J. (Eds.) *Oxford Handbook of School Psychology.* New York, NY: Oxford University Press.

Galster, G. (2012) *Driving Detroit*. Philadelphia, PA: University of Pennsylvania Press.

Glisson, C., & James, L. R. (2002). The cross-level effects of culture and climate in human service teams. *Journal of Organizational Behavior*, *23*, 767–794.

Good, A. B., Burch, P., Stewart, M. S., Acosta, R., & Heinrich, C. (2014). Instruction matters: Lessons from a mixed-method evaluation of out-of-school time tutoring under no child left behind. *Teachers College Record*, *116*, 1-34.

Gregory, A., Clawson, K., Davis, A., & Gerewitz, J. (2015). The promise of restorative practices to transform teacher-student relationships and achieve equity in school discipline. *Journal of Educational and Psychological Consultation*, *25*, 1-29.

Gregory, A., Cornell, D., & Fan, X. (2011). The relationship of school structure and support to suspension rates for black and white high school students. *American Educational Research Journal*, *48*, 904–934.

Gregory, A., Cornell, D., Fan, X., Sheras, P. L, Shih, T., & Huang, F. (2010). Authoritative school discipline: High school practices associated with lower student bullying and victimization. *Journal of Educational Psychology, 102,* 483-496.

Gregory, A. & Korth, J. (in press). Teacher-student relationships and behavioral engagement in the classroom. In K. Wentzel & G. Ramani (Eds.) *Handbook of Social Influences on Social-Emotional, Motivation, and Cognitive Outcomes in School Contexts.* New York: New York: Taylor and Francis.

Gregory, A. & Ripski, M. B. (2008). Adolescent trust in teachers: Implications for behavior in the high school classroom. *School Psychology Review, 37*, 338-353.

Gregory, A. & Weinstein, R. S. (2004). Connection and regulation at home and in school: Predicting growth in achievement for adolescents. *Journal of Adolescent Research*, *4,* 405-427.

Guskey, T. R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education, 4*, 6369.

Hatch, T. (2002). When improvement programs collide. *Phi Delta Kapplan, 83,* 626.

Ingersoll, R. M. & Smith, T. M. (2003). The wrong solution to the teacher shortage. *Keeping Good Teachers, 60*, 30-33.

Improving School Climate. (2014). Bethlehem, PA: International Institute for Restorative Practices. Retrieved from http://www.iirp.edu/pdf/IIRP-Improving-School-Climate.pdf.

Kam, C. M., Greenberg, M. T., & Walls, C. T. (2003). Examining the role of implementation quality in school-based prevention using the PATHS curriculum. *Prevention Science, 4*, 55–63.

Kena, G., Aud, S., Johnson, F., Wang, X., Zhang, J., Rathbun, A., Wilkinson-Flicker, S., and Kristapovich, P. (2014). *The Condition of Education 2014*. U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved from http://nces.ed.gov/pubsearch.

Kim, C. Y. (2003). Procedures for public law remediation in school-to-prison pipeline litigation: Lessons learned from *Antoine v. Winner School District*. *New York Law School Law Review*, *54*, 955.

Klassen, R. M., Tze, V. C., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998-2009: Signs of progress or unfulfilled promise?. *Educational Psychology Review, 23,* 21-43.

Klein, K. J., & Sorra, J. S. (1996). The challenge of innovation implementation. *Academy of Management Review*, *21*, 1055–1080.

Kozol, J. (2005). Apartheid in America? *Phi Delta Kappa, 87*, 264-275.

Lewis, S. (2009). Improving school climate: Findings from schools implementing restorative practices. Bethlehem, PA: International Institute of Restorative Practices. Retrieved from http://www.iirp.edu/pdf/IIRP-Improving-School-Climate.pdf

Lohrmann, D. K. (2008). A complementary ecological model of the coordinated school heath program. *Public Health Reports, 123*, 695-703.

Losen, D. J. (2011). *Good Discipline: Legislation for Education Reform.* Boulder, CO: National Education Policy Center. Retrieved from http://nepc.colorado.edu/publication/discipline-policies-legislation.

Losen, D.J., & Martinez, T.E. (2013). *Out of School & Off Track: The Overuse of Suspensions in American Middle and High Schools.* The Center for Civil Rights Remedies. Retrieved from http://civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/out-of-school-and-off-track-the-overuse-of-suspensions-in-american-middle-and-high-schools/Exec_Sum_OutofSchool_OffTrack_UCLA.pdf.

Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership*, *70*, 14-19.

Mayer, M. (2010). Structural analysis of 1995-2005 school crime supplement datasets: Influencing students' fear, anxiety, and avoidant behaviors. *Journal of School Violence*, *9*, 37-55.

McCarthy, P. & Kerman, B. (2010). Inside the belly of the beast: How bad systems trump good programs. *Administration & Policy In Mental Health & Mental Health Services Research, 37*, 167-172.

McCluskey, G., Lloyd, G., Kane, J., Stead, J., Riddell, S., & Weedon, E. (2008). Can restorative practices in schools make a difference? *Educational Review*, *60*, 405–417.

McDonald, J. F. (2014). What happened to and in Detroit? *Urban Studies (Sage Publications, Ltd.)*, *51*, 3309-3329.

McIntosh, K., Ty, S., & Miller, L. (2014). Effects of school-wide positive behavioral interventions and supports on internalizing problems: Current evidence and future directions. *Journal of Positive Behavior Interventions, 16*, 209-218.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis.* Thousand Oaks, CA: SAGE.

National Research Council. (2004). Engaging schools: Fostering high school students' motivation to learn. Washington, DC: National Academy Press.

U.S. Department of Justice (2012) *National Youth Gang Survey Analysis.* Retrieved from http://www.nationalgangcenter.gov/Survey-Analysis.

No Child Left Behind Act of 2001, 20 U.S.C. § 1111 (2002).

Patton, M.Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., . . . Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*, 65–76.

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

Rohrbach, L.A., Graham, J.W., & Hansen, W.B. (1993). Diffusion of a school-based substance abuse prevention program: Predictors of program implementation. *Preventive Medicine*, *22*, 237–260.

Rolland, R. G. (2012). Synthesizing the evidence on classroom goal structures in middle and secondary schools: A meta-analysis and narrative review. *Review of Educational Research 82*, 396-435.

Rubie-Davies, C. M., Weinstein, R. S. Huang, F. L., Gregory, A., Cowan. P., & Cowan, C. (2014). Successive teacher expectation effects across the early school years. *Journal of Applied Developmental Psychology, 35,* 181-191.

Schiff, M. (2013). *Dignity, disparity, and desistance: Effective restorative justice strategies to plug the "school-to-prison pipeline."* Paper presented at the Center for Civil Rights Remedies National Conference. Closing the School to Research Gap: Research to Remedies. Washington, DC.

Schulte, A. C., Easton, J. E., & Parker, J. (2009). Advances in treatment integrity research: Multidisciplinary perspectives on conceptualization, measurement, and enhancement of treatment integrity. *School Psychology Review, 38*, 460-475.

Skiba, R. J., Arredondo, M. I., & Williams, N. T. (2014). More than a metaphor: The contribution of exclusionary discipline to a school-to-prison pipeline. *Equity & Excellence In Education*, *47*, 546-564.

Skiba, R., Michael, R. S., Nardo, A. C., & Peterson, R. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *Urban Review, 34,* 317-342.

Skiba, R. J., & Rausch, M. K. (2006). Zero tolerance, suspension, and expulsion: Questions of equity and effectiveness. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 1063–1089). Mahwah, NJ: Erlbaum.

Skiba, R., Simmons, A. B., Peterson, R., McKelvey, J., Forde, S., & Gallini, S. (2004). Beyond guns, drugs and gangs: The structure of student perceptions of school safety. *Journal Of School Violence*, *3*, 149-171.

Skinner, B. F. (1969). *Contingencies of reinforcement: A theoretical analysis.* New York, NY: Appleton-Century-Crofts.

Smith, D. L., & Smith, B. J. (2006). Perceptions of violence: The views of teachers who left urban schools. *High School Journal*, *89*, 34-42.

Stein, M. K., & Wang, M. C. (1988). Teacher development and school improvement: The process of teacher change. *Teaching and Teacher Education, 4,* 171-187.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* Newbury, CA: Sage.

Tickle, B. R., Chang, M., & Kim, S. (2011). Administrative support and its mediating effect on US public school teachers. *Teaching and Teacher Education, 27*, 342-349.

Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68,* 202.

Tschannen-Moran, M. & Woolfolk, A. (2001). Teacher efficacy: Capturing an elusive concept. *Teaching and Teacher Education, 17*, 783-805.

U.S. Census Bureau. (2015). Retrieved from http://www.census.gov/.

Walker, J. M. (2008). Looking at teacher practices through the lens of parenting style. *Journal of Experimental Education, 76*, 218–40.

Welch, K., & Payne, A. (2010). Racial threat and punitive school discipline. *Social Problems, 57*, 25–48.

Weiner, B. J. (2009). A theory of organizational readiness for change. *Implementation Science*, *4*, 67.

Weinstein, R. S. (2002). *Reaching Higher: The power of expectations in schooling.* Cambridge: Harvard University Press.

Wentzel, K. R. (2002). Are effective teachers like good parents? Interpersonal predictors of school adjustment in early adolescence. *Child Development, 73*, 287-301

Wilson, W. J. (2003). Race, class and urban poverty: A rejoinder. *Ethnic & Racial Studies*, *26*, 1096-1114.