10-23-2020

Assessing the Effects of Restorative Practices on Teacher Practices in Elementary Classrooms

Cheng Fisher
cheng.fisher@jefferson.kyschools.us

Assessing the Effects of Restorative Practices on Teacher Practices in Elementary Classrooms

By

Cheng Z. Fisher

A DISSERTATION

submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy
in Education and Social Change

Annsley Frazier Thornton School of Education
Bellarmine University

October 2020

ABSTRACT

Title: Assessing the Effects of Restorative Practices on Teacher Practices in Elementary Classrooms

Author: Cheng Z. Fisher

Dissertation Chair: Grant S. Smith, Ph.D.

Dissertation Committee Members: Florence C. Chang, Ph.D.
William A. Wells, Ed.D.

Restorative practices, endorsed by the U.S. Department of Education (2014) as an alternative discipline approach and a continuum of proactive strategies with the aim of community development through relationship building, have been broadly adopted in American schools. However, few empirical studies have rigorously examined the effects of restorative practices on teacher practices through direct observations. This quasi-experimental designed study tested the hypothesis that restorative practices experience would increase positive interactions between teachers and students. Measurably, it would improve positive teacher practices and positive student behavior. The study analyzed data from a school-randomized evaluation with two rounds of observations in a single school year. An observation tool was developed for data collection of all measures. Observational data of 140 elementary teachers and their students from nine restorative schools were used to examine the impact of restorative practices on teacher practices and student behavior. Participants were grouped into four condition levels, indicating different dosage levels of restorative practice experience ranked by the training year.

The results of the Mann-Whitney test did not suggest significant differences in teacher practices and student behavior between the intervention and comparison groups. Repeated measures ANOVA indicated that teachers with more experience of restorative practices

progressed significantly in non-specific praise over time. Correlation analysis suggested a significant positive and moderate relationship between teacher practices and student behavior. Specifically, positive student behavior was significantly associated with restorative language and opportunities to respond, but not non-specific praise. Overall, the results of current data indicated that the experience of restorative practices did not make substantial impacts on teacher practices and student behavior. However, the experience increased teachers' awareness of using basic positive practices. The study concluded with limitations and implications of findings and recommendations for practice and future research.

ACKNOWLEDGEMENTS

This dissertation would never have come to fruition without the professional and personal support of many individuals. First, I want to express my great appreciation to my chair and advisor, Dr. Grant Smith. Thank you for believing in my abilities. I would not have achieved the accomplishments in this Ph.D. program without your professional guidance and scholarly discussion. I want to acknowledge Dr. Florence Chang for constructive feedback and continuous support for my intellectual development in the field of research and evaluation. I also want to thank Dr. Will Wells for the time and effort you have put into helping me grow throughout this dissertation process.

Furthermore, this study would not have become a reality without the support and assistance of Ms. Saundra Hensel and Dr. Leslie Taylor. Thank you for allowing me to continue studying the evaluation project. Hopefully, my study will make a small contribution to the improvement of this district initiative.

Finally, it is with my deepest gratitude that I must thank my family. My biggest thanks and apology to my husband Aaron, and my son Matthew. Thank you for taking care of everything at home so that I can focus on my study. And thank you for being my loyal audience reading my paper and listening to my speech even though it was incredibly dull for a 7-year old. Last but not least, I want to thank my parents in China for teaching me everything I need to survive and thrive in any life circumstances. I hope you are proud of whom I have become.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

In 2014, the U.S. Department of Education (USDOE) issued Guiding Principles to assist schools and districts around the nation in continuously improving school climate and discipline policies. Among the promising programs and practices endorsed by this guiding document, restorative practices (also known as restorative justice practices) was highly recommended as an alternative discipline approach and a set of proactive strategies to build student social-emotional capacity (U.S. Department of Education, 2014).

As indicated by USDOE (2014), restorative justice practices are the extension of restorative justice. The term "Restorative Practices" is frequently used by the International Institute for Restorative Practices (IIRP) to expand the use of restorative justice practices in regular school settings. Restorative practices aim to build healthy school communities and repair harm through various practices (Costello, Wachel, & Watchel, 2009).

The USDOE defines restorative practices as a continuum of practices used in schools, ranging from short informal conversations to formal conferences with multiple stakeholders' involvement. IIRP further specifies the continuum into three levels, in a total of 11 elements of restorative practices, including School-wide (affective statements, restorative questions, small impromptu conferences, restorative staff community, fundamental hypothesis), broad-based (proactive circles, responsive circles, fair process, re-integrative management of shame, and restorative approach with families), and targeted (restorative conferences) (International Institute for Restorative Practices [IIRP], 2010).

The development of restorative practices was rooted in the affect theory (Acosta, Chinman, Ebener, Malone, Phillips, & Wilks, 2019; Costello et al., 2009). Attributed to the psychologist Silvan Tomkins, affect theory suggests that quality mental health can be developed

by maximizing positive affects and minimizing negative affects. More importantly, people should appropriately express their emotions with others as much as possible to build positive relationships and connect within the community (Tomkins, 1962, 1963, 1991; Nathanson, 1992; Costello et al., 2009).

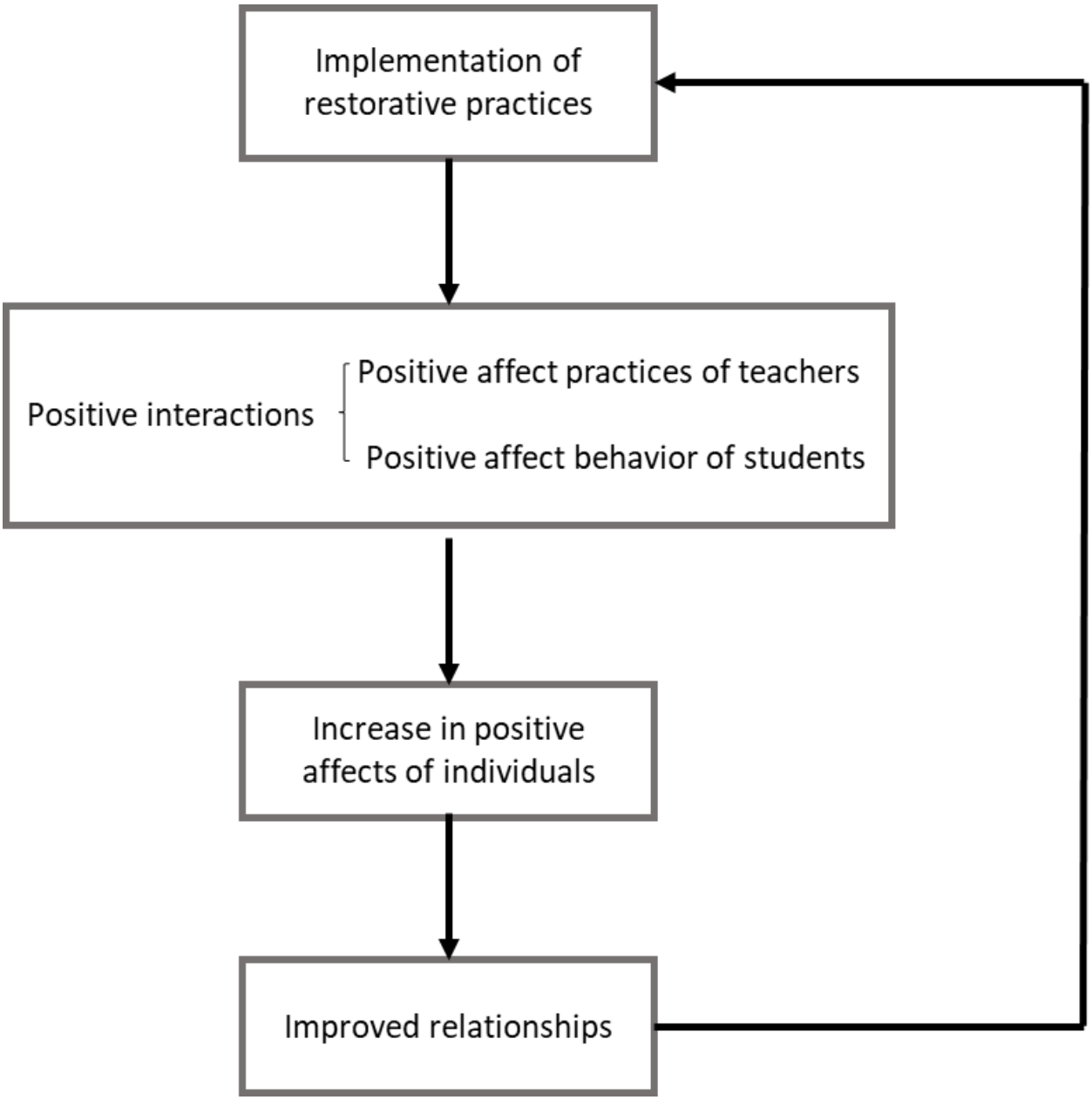
Tomkins (1962, 1963) has identified and categorized the nine biological affects into three psychological expresses, including positive affects (enjoyment/joy, interest/excitement), neutral (surprise/startle), and negative affects (shame/humiliation, distress/anguish, disgust, fear/terror, anger/rage, and dissmell). The theoretical foundation of restorative practices is to maximize positive affects and to minimize negative affects through free expressions by applying the practices.

A conceptual framework portrayed in Figure 1.1 is developed by the researcher to explore the mechanism of restorative practices to maximize positive affects and improve relationships through positive interactions between teachers and students. This framework is the theoretical base to transform the construct of affect or relationship into observable measures of interactions that further break down into teacher practices and student behavior. This conceptual framework is further elaborated in chapter two after an in-depth literature review of affect theory.

Some literature has suggested promising educational outcomes of restorative practices in decreasing school discipline incidents, improving academic achievement, and narrowing racial disproportionality gap (Oakland Unified School District, 2014; Lewis, 2009). However, very few peer-reviewed studies with rigorous empirical or quasi-design have investigated the effectiveness of restorative practices in regular school settings (Acosta et al., 2019; Green et al., 2019). To date, quality empirical evidence regarding the effectiveness of restorative practices is meager and inconclusive. More scholarly studies are desired to enhance a shared understanding of this topic.

Figure 1.1

A Conceptual Framework Illustrating the Mechanism of Relationship Building through Restorative Practices



Acosta and colleagues (2019) conducted the first randomized control trial study examining the outcomes of restorative practices across a two-year implementation period. Their study aimed to assess the effects of restorative practices on school connectedness, bullying victimization, and youth development outcomes on middle school students.

Funded by the U.S. Department of Justice, Green, Willging, Zamarin, Dehaiman, and Ruiloba (2019) are currently conducting a five-year evaluation of restorative practices with a cluster-randomized design. The results of the study are expected to be disseminated in 2020. This longitudinal study adapts implementation science approaches to monitor the program progress and assess the effectiveness of restorative practices. Prior to these empirical studies, a quasi-experimental designed study (Gregory, Clawson, Davis, & Gerewitz, 2016) examined the impact of restorative practices on the relationships between high school minority students and their teachers.

The findings regarding the effectiveness of restorative practices on behavioral and youth developmental outcomes are inconclusive (Acosta et al., 2019; Gregory et al., 2016). Most analyses have primarily relied on student or teacher self-reported data (Acosta et al., 2019; Green et al., 2019; Gregory et al., 2016). Researchers have discussed the need for observational data to deepen the understanding of restorative practices at the classroom-level (Acostal et al., 2019). It has been commonly agreed that observation from an outside evaluator is the "gold standard" for evaluating an educational program (Forman et al., 2013; Gregory et al., 2016). However, most restorative practices in classrooms are impromptu, which presents a challenge for observers to capture the practices during a limited period. So far, no study has used direct classroom observational data to examine the outcomes of restorative practices. As a result, there is a lack of established classroom observation tools to evaluate restorative practices.

Most literature has focused on measuring indirect long-term intervention outcomes of restorative practices. These behavioral outcomes are measured by administrative disciplinary data, graduation rates, academic achievement, or perceptions of school climate. However, it is problematic to determine long-term outcomes without discerning the fidelity of program implementation and the stages of the implementation at the time of the study. If a program has not been fully implemented, the effectiveness of the program cannot be determined. The implementation of restorative practices is a complex non-linear process. It takes multiple years to reach the full implementation stage. Researchers have discussed this limitation (Green et al., 2016).

Nevertheless, little research has evaluated the short-term outcomes of restorative practices, such as teacher practices and student behavior. Also, the majority of the studies have investigated restorative practices in middle and high school settings. We have limited knowledge about the impact of restorative practices in elementary schools.

Statement of the Problem

While there is a general agreement that restorative practices make positive impacts on student outcomes, empirical evidence of its effectiveness at the classroom level is lacking (Acosta et al., 2019; Gregory et al., 2016; Costello et al., 2009; USDOE, 2014). As restorative practices become more and more popular in the school system, relevant research faces many challenges to provide meaningful inferences to guide continuous improvement efforts.

First, the field needs well-designed empirical studies to assess the outcomes of restorative practices. It is unclear how restorative practices impact teacher practices and student behavior at the classroom level. Second, there is a lack of research using direct observations. Existing research primarily depends on self-reported survey data or administrative behavioral data (Acosta et al., 2019; Green et al., 2019; Gregory et al., 2016). Therefore, an investigation into the

effects of restorative practices using direct classroom observational data is needed, especially at the elementary school level.

Purpose of the Study

The primary objective of this study is to evaluate the impact of restorative practices on teacher practices and student behavior in elementary classrooms through direct observations. This quasi-experimental designed study draws on the existing data from a stratified school-level randomized evaluation of restorative practices to examine if teachers with more experience in restorative practices show more use of positive affect practices and make more progress in positive practices between times.

Significance of the Study

Several key points highlight the significance of the study. First, the study fills the knowledge gap about the implementation of restorative practices at the practice (classroom) level with the objective and direct observational data collected from an external researcher. Most studies have focused on examining long-term outcomes, which would be difficult to detect without the complete implementation of the intervention. The current study explores how the implementation experience directly changes behavior or daily practices.

Another significant contribution of the study is to provide a classroom observation tool allowing future research with the use of observational data to investigate the outcomes of restorative practices. Due to the affective nature, these practices are difficult to be observed in a short visit. The newly developed observation tool has replaced the prior district instrument since many practices, such as circles, affective statements, and small impromptu conversations, could not be recorded during the planned observation period. Although the new tool is not near perfect, it allows capturing some essences of the practices in an ordinary instructional period to

encompass further improvement. Finally, the study reveals more knowledge about restorative practices in the elementary school setting, which has not been explored much in the literature.

Research Questions

The conceptual framework for the study builds on the affect theory that serves as the theoretical foundation of restorative practices. This framework (figure 1.1) reveals the transformation of relationship building through implementing the intervention. Restorative practices provide a platform for positive interactions, which results in positive affects and relationships between participants. The data offer leveled groups based on teachers' experience with restorative practices. Examining the differences between and within these groups over time helps us understand the impact of restorative practices on teacher practices. Four research questions guide the investigation of the current study.

Research Question 1: Are there differences in teacher practices and student behavior between classrooms where teachers have been trained in restorative practices and classrooms where teachers have not been trained in restorative practices?

Research Question 2: Are there differences in teacher practices and student behavior among classrooms with teachers in different intervention condition levels (trained 3 years ago/ 2 years ago/ 1 year ago/ not trained)?

Research Question 3: Do teacher practices and student behavior change overtime differently between the four condition levels?

Research Question 4: What are the relationships between teacher practices and student behavior?

The first question provides an overview of the differences between the trained group and the untrained group. It explores whether the formal implementation of restorative practices makes a difference. The next question extends this inquiry. The researcher is interested in

understanding whether more experience in restorative practices yields more positive outcomes in teacher practices and student behavior. The second and third questions look into the between-subjects and within-subjects effects. The last question explores the correlations between teacher practices and student behavior.

The study takes advantage of using a comparison group which is a group of teachers who work at the nine restorative elementary schools but have never received official training on restorative practices. Therefore, these participants may be familiar with the concept of restorative practices and may also apply them in their classrooms. The point is that they have less restorative experience than the intervention group.

Assumptions

The current study assumes that restorative practices contribute to the changes in teacher practices and student behavior. In other words, the changes in teacher practices and student behavior are the outcomes of the implementation of restorative practices. The restorative program is one of the multiple initiatives in these selected schools. For this study, we assume a causal relationship between the participants' experiences of restorative practices and their behavioral changes.

Another critical assumption of the study is that the study schools implement restorative practices with an adequate level of fidelity. The district data department developed an evaluation plan in 2017 to guide the evaluation process. However, at this time, the district has not conducted a process evaluation to measure the fidelity level of the program implementation. This plan also did not include a logic model of the program evaluation. To help the understanding of study design, Appendix A illustrates the underlying logic of the current study. This study could not evaluate the program implementation from input to output indicated in Appendix A. It is an assumption that the program is implemented with acceptable fidelity as planned.

Although the current research does not assess the fidelity of implementation due to many limitations, it should be emphasized that process evaluation is essential and critical before examining the outcomes of a program or a practice. This regard is further elaborated in the limitation session at the end of the study.

Key Terms

Restorative justice: A problem solving approach to crime that focuses on restoration or repairing the harm done by the crime and criminal to the extent possible, and involves the victim(s), offender(s) and the community in an active relationship with statutory agencies in developing a resolution. (United Nations 2003, p.28)

Restorative practices / Restorative justice practices: non-punitive disciplinary responses that focus on repairing harm done to relationships and people, developing solutions by engaging all persons affected by a harm, and accountability. A variety of restorative practices can be used in schools, ranging from brief on-the-spot responses to student behavior in the classroom to community conferencing involving multiple parties. The goals of restorative justice intervention in schools are to address the harm committed and enhance responsibility and accountability, build relationships and community, and teach students empathy and problem solving skills that can help prevent the occurrence of inappropriate behavior in the future. (USDOE, 2014, p.24)

Affect: Any experience of feeling or emotion…Often described in terms of positive affect or negative affect, both mood and emotion are considered affective states. Along with cognition and conation, affect is one of the three traditionally identified components of the mind. (APA, retrieved from https://dictionary.apa.org/affect)

Affect theory: The idea that feelings and emotions are the primary motives for human behavior, with people desiring to maximize their positive feelings and minimize their negative

ones. Within the theory, affects are considered to be innate and universal responses that create consciousness and direct cognition. Eight primary affects are postulated: positive ones of excitement and enjoyment; the negative ones of distress, fear, shame, disgust, and anger; and the relatively neutral one of interest. Despite their biological nature and triggering mechanisms, primary affects are subject to significant social modification and social causation. (APA, retrieved from https://dictionary.apa.org/affect-theory)

Fidelity of implementation: The implementation of a practice or program as intended by the researchers or developer. (Innovative Resources for Instructional Success Center [IRISCenter], retrieved from https://iris.peabody.vanderbilt.edu/module/ebp_02/cresource/q1/p01/)

Inter-rater reliability: The extent to which independent raters produce similar ratings in judging the same abilities or characteristics in the same target person or object. (APA, retrieved from https://dictionary.apa.org/interrater-reliability)

CHAPTER TWO: LITERATURE REVIEW

Few quantitative studies with rigorous empirical design have investigated the effects of restorative practices. Most research related to this topic has been qualitative or mixed-method design. This review of literature starts with tracing the origin of restorative practices. Various definitions and components of restorative practices are explored.

We then scrutinize the theoretical framework of restorative practices, which constitutes the design of the current study. A conceptual framework is developed after the review of the theories to guide the investigation. The literature of student-teacher interactions provides evidence associated with the topic and supports the conceptual framework and instrument design.

This literature review focuses on studies conducted by independent researchers in general educational settings. It briefly discusses three comprehensive evaluations, which provide the overall context of the field. A variety of definitions and implementation processes are related to restorative practices indicated in these studies.

Finally, the review concentrates on three empirical studies focusing on measures, implementation components, findings, and limitations. There is also a lack of clear standards on the implementation process. For instance, some schools only receive the introduction training, while some are provided with on-going training sessions, coaching, or consultation throughout the implementation stages. The way that an intervention is implemented can profoundly affect the results of research. Therefore, it is critical to understand the effects of the intervention with consideration of its implementation process.

From Restorative Justice to Restorative Practices

Restorative practices evolved from restorative justice, which is rooted in the criminal justice system. In 1975, two adults were convicted of vandalism in a local neighborhood in Canada. In the court, a probation officer suggested a Christian response as an alternative for a fine and probation. It was the first time that restorative justice was applied in a criminal justice system. Inspired by the application, the first Victim-Offender Reconciliation Project (VORP) was created in Ontario. The primary goal of VORP was to reach reconciliation between victims and offenders through communication, sharing feelings, and reducing damage. Within a decade, similar projects emerged in many countries in Europe and North America, such as Germany, England, Austria, and United States (Marshall & Merry, 1990, Hopkins & Masters, 2003).

The 1989 Children, Young Persons, and Their Families Act in New Zealand triggered reform in the youth justice system. As a result of this reform, a set of principles was established to guide practices and decision-making regarding how to treat offenders in youth justice cases (Maxwell & Morris, 1993). In the following year, the first Family Group Conferencing was conducted in New Zealand. This mediation empowered the young person, the families, and the victims in the process of decision-making. At the end of each session, a facilitator would leave the room and allow participants to collaborate and develop an action plan. This agreement addressed concerns and provided obtainable actions to support offenders and victims further. This reform intended to reduce youth crime in the Maori population and to include family and community into the decision-making process (Hospkins & Masters, 2003).

In New Zealand, Family Group Conferencing (FGC) allows young offenders to develop a plan and convince the court for another opportunity. On average, 5000 FGCs are performed each year in New Zealand. Courts formally consider FGC plans for decision-making. Maxwell and Morris's (1993) data indicated that only about 5% of conferences concluded without an FGC

plan. The majority of the plans were approved and accepted by courts. Many successful stories of implementing FGC are also shared from other countries (Jackson, 1998; Marsh & Crow, 1998). Researchers claimed that many victims and offenders had welcomed this mediation. The approach helped reduce the anxiety caused by the crime (Umbreit & Roberts,1996).

In light of the movement of FGC, restorative conferencing emerged in New South Wales, Australia. The police in Wagga Wagga initiated the first restorative conference in 1993. Braithwaite (1989), in the theory of "reintegrative shaming," emphasized the importance of offenders receiving personal and community supports. In a restorative conference, a neutral facilitator uses a scripted format and questions to guide the process. Supporters from families and communities are invited to attend the meeting. The positive atmosphere is created to help participants explore their feelings toward the harm caused by the incident or crime. Many countries have been implementing this model and resulted in positive outcomes. It has become popular and successful among schools (Hoskins & Maters, 2003).

In the United Kingdom, the 1998 Crime and Disorder Act in England and Wales did not constitute restorative justice into the youth justice process. However, an alternative model was highly encouraged to apply to young offenders who committed the first and second minor offenses. Adequate funding from the Youth Justice Board has been invested in training the youth offending team members for the restorative conferencing (Hoskins & Masters, 2003). In some areas of England and Wales, this mediation became prevalent. However, despite the positive feedback provided by victims who participated in this process, only a small percentage of youth criminal cases involve the Restorative Justice process (Dignan, 2002; Hoskins & Masters, 2003). Instead of reluctance from victims, one of the reasons for low participation of this process was the complexity of the practices (Dignan 2002; Holdaway et al., 2001; Newburn et al., 2001).

In the year following the 1998 Crime and Disorder Act, the Youth Justice and Criminal Evidence Act in 1999 successfully introduced the restorative justice process to the youth justice system. Under this act, the majority of first offenders who plead guilty are given a referral order. They are ordered to attend the Youth Offender Panel designed for the restorative process like Family Group Conferencing. This process involves the offender, families, victims, supports, community members, and a trained facilitator. Members of the panel receive a week-long training. Elements of reparation are concluded and recorded in a contract agreed by the panel. The review panel evaluates the progress of the young offender to determine criminal conviction. A study indicated that among the young first offenders who participate in a Youth Offender Panel, about three-quarters of them achieved their contracts. Less than 25% of them committed another crime. Seventy-five percent of participants expressed satisfaction with beneficial outcomes (Newburn et al., 2002).

Howard Zehr (1990), an American criminologist, was one of the pioneers of the restorative justice movement. He reiterated that the restorative justice process should focus on relational inquiry in discussing what happened, who has been harmed, and what needs to be done to repair the harm. He developed a list of guiding questions to facilitate the inquiry process. IIRP further adapted these questions as restorative questions as one of the essential elements of restorative practices (Costello et al., 2009).

In the early 21st century, restorative justice was introduced to educational settings. It soon became popular in schools around the world, such as Australia, New Zealand, Canada, the USA, UK, and other European countries (McCluskey et al., 2008). Different countries adopted restorative justice in various forms to meet the context and needs of their educational system (Miers, 2001). Many schools around the world have been implementing restorative practices as an alternative response to zero-tolerance discipline policy, as punitive disciplinary approaches

have been criticized for its ineffectiveness to improve student behavior and school climate (American Psychological Association Zero Tolerance Task Force, 2008).

Definitions of Restorative Practices

Restorative practices originated in restorative justice. Restorative justice emphasizes repairing harm between individuals and within the community. When a behavior violation occurs, restorative justice uses formal conferences to engage relevant parties in a problem-solving process. Restorative practices follow the same principles of restorative justice, applied in an educational setting. Moving beyond restorative justice, the continuum of restorative practices includes proactive and responsive elements. (Costello et al., 2009; USDOE, 2014).

Even though restorative justice has been broadly adopted in schools around the world, there is little consensus on the definition, active elements, or implementation standards. The lack of a clear definition of restorative practices presents many challenges for researchers to assess the effectiveness of the overall program.

International Institute for Restorative Practices (IIRP) is an international institute that aims to improve the social community and human relationships through restorative practices. They define restorative practices as "an emerging social science that studies how to strengthen relationships between individuals as well as social connections within communities" (IIRP website, n.d.). The institution names 11 essential elements of restorative practices and groups them into three levels: school-wide (affective statements, restorative questions, small impromptu conversation, fair process, reintegrate shame, staff community, fundamental hypothesis), broad-based (proactive circles, responsive circles, family approach), and targeted (restorative conferences) (IIRP, 2010) (Table 2.1).

Table 2.1

11 Essential Elements of Restorative Practices and Definitions

		Essential Element	Definition
School-Wide	1	Affective Statements	Students and staff use "I" statements to express their feeling towards a behavior event. The purpose of the statements is to make individuals aware of the positive or negative impact of their behavior.
	2	Restorative Questions	Standard questions for both individuals, causing harm and affected by harm. The questions focus on what happened, who has been affected, and what should do to make things right, etc.
	3	Small Impromptu Conferences	An informal and short conversation takes place right after a low-level incident to prevent the escalation. It may involve the use of affective statements and restorative questions.
	4	Fair Process	Three primary components include engaging individuals in decision-making, explaining the reasons for the decision, clarifying expectations.
	5	Reintegrative Management of Shame	Help wrongdoers understand the violation of their behavior rather than personal characteristics. Separate deed from the doer and avoid stigmatized shame.
	6	Restorative Staff Community	Use restorative practices for conflict resolution and community building among school staff.
	7	Fundamental Hypothesis Understandings	Enhance the understanding of the primes of restorative practices: high expectations and high support. Individuals with authority are expected to do things "with" people, not "to," "not," or "for" them.
Broad-Based	8	Proactive Circles	The purpose of the circles is to build community. They take place regularly (80% of the time). They can also be used for academics.
	9	Responsive Circles	The purpose of these circles is to repair harm after behavioral incidents. Participants are engaged in the reflection and problem-solving process.
	10	Restorative Approach with Families	Use restorative practices to engage families in meaningful conversations and problem-solving process.
Targeted	11	Restorative Conferences	A formal conference is used to respond to severe behavior events. It may involve other community members, parents, and school officials.

Restorative practices provide strategies to prevent and intervene in student behavioral infractions in school settings. As a whole-school approach, school practitioners are encouraged to exercise a continuum of strategies for both proactive and responsive purposes. Along the continuum from left to right, there are affective statements, small impromptu conversations, circles, and conferencing. These practices aim at community building and relationship restoration. Schools are encouraged to focus on the proactive component at 80% of the overall practices to prevent student behavioral violations (Costello et al., 2009).

In collaboration with other educational organizations, Schott Foundation (2014) published a guide for educators to exercise restorative practices in schools. In this guide, restorative practices are defined as "processes that proactively build healthy relationships and a sense of community to prevent and address conflict and wrongdoing (p. 2)." They specify nine types of restorative practices: community conferencing, community service, peer juries, circle process, resolution programs, peer mediation, informal restorative practices, and social-emotional learning.

Oakland Unified School District (OUSD, 2014) has been implementing restorative practices for more than a decade. The school district prefers the term restorative justice. The restorative justice framework at OUSD provides a 3-tier school-wide approach: Tier 1 component focuses on building relationships and fostering a positive school climate; Tier 2 approaches emphasize the use of non-punitive responses to conflicts; Tier 3 of the framework provides one-on-one individual supports. The district defines its restorative justice framework as "taking a community-building approach that addresses the root causes of a student disruptive/conflict behavior through listening, accountability, and healing (OUSD, 2014, p. 3)." The implementation of restorative justice at OUSD includes circles, mediation, restorative conversations, family groups, and community conferences.

In early 2000, the Youth Justice Board of England and Wales commissioned the most extensive independent evaluation of restorative practices. The implementation in the national restorative practices program included active listening, restorative inquiry, circle time, mediation, and restorative justice conferences (Bitel, 2005). The restorative justice approaches in schools in the United Kingdom, as highlighted in the evaluation report, "encompass a range of initiatives that operate along the continuum of the gravity of rule-breaking or harm done (p. 10)." Common practices in the country were circle time, peer mediation, the 'no blame' approach, and restorative conferencing. Some schools included conflict resolution education as a part of the citizenship curriculum. The aspects of restorative practices in the nation focused on listening, communication techniques, anger management, and a sense of responsibility (2005).

Another evaluation project in the UK investigated restorative practices in three Sottish Councils (Kane et al., 2005). In this report, restorative practices were defined as "restoring good relationships when there has been conflict or harm and developing school ethos, policies and procedures to reduce the possibility of such conflict and harm arising (P. 6)." Each school was encouraged to adopt the restorative approaches to meet the school's context. A range of practices was recommended, including ethos building, curriculum focus on relationships, restorative language, restorative inquiry, restorative conversations, mediations, circles, restorative meetings, and formal conferences.

Various terms are used to refer to restorative practices, such as restorative approaches, restorative justice practices, or restorative interventions. Their definitions have commonly accentuated the value of building relationships and repairing the harm through meaningful engagement of stakeholders. However, the examples above have illustrated a broad range of definitions and adaptations of the practices at schools. For this study, we use restorative practices

as a general term for restorative strategies as an extension of restorative justice in educational settings.

In 2014, USDOE and the U.S. Department of Justice (DOJ) issued a school discipline guidance package to assist the school improvement effort of enhancing school culture and climate around the nation. This document defines restorative practices as non-punitive disciplinary approaches focusing on repairing harmed relationships and engaging all parties in the problem-solving process. The goals of restorative practices in schools are to address responsibility and accountability, to prevent future inappropriate behavior through building relationships and teaching empathy and problem-solving skills.

Theoretical Framework of Restorative Practices

Charles Darwin's books, On the Origin of Species (1859) and The Expression of the Emotions in Man and Animals (1872), built the theoretical foundation for almost all the modern scientific theories of emotional behavior (Ludwig & Welch, 2019). Darwin (1872) suggested that all mammals, including humans, are born with a particular set of emotions. Our emotions support communication and preparatory function. They are essential for surviving and motivation. He suggested that the cortical system regulates heart rates and emotions of man and animals. Darwin believed that these social instincts are biologically hardwired and inherited in humans.

Influenced by Darwin's work, two scientists, William James and Carl Lange (1922), independently came up with a similar theory. The James-Lange theory proposed that the experience of emotions was the brain responding to the stimuli or information through the nervous system. The autonomic nervous system controls heart rate, blood pressure, physiologic conditions, and responses (1922). James' theory put viscera at the center of emotions. James believed that visceral reactions demonstrate patterns of certain emotions.

Researchers challenged the James-Lange theory. Cannon (1927) suggested that visceral reactions do not have observable responses to specified emotional behavior. He argued that visceral responses could be observed among the various emotions or no emotion. For instance, people may feel a fast heartbeat and sweating when they experience fear or anger. It was evident that visceral responses cannot thoroughly select certain emotions.

In the book, Descent of Man (1871), Darwin reviewed the empirical evidence on emotions and summarized emotions into three general classes. In his 2nd edition of Expression of Emotions (1890), Darwin provided a list of 34 emotions and classified them into 8 clusters.

James (1890) indicated that feelings and emotions could be distinguished and isolated through the stream of consciousness. Consciousness is continuous. It constitutes cognition and constantly motivates the actions of knowing, thinking, and saying. A feeling or an emotion is an object of consciousness, also an object of its cognition.

After nearly 50 years of investigation, Sylvan Tomkins concluded that humans have two distinguishable lives: an affective life and a cognitive life. Tomkins first introduced affect theory in his book Affect Imagery Consciousness (1962). He defined affect as the biological portion of emotion.

> Affects are sets of muscles and glandular responses located in the face and also widely distributed through the body, which generates sensory feedback, which is either inherently “acceptable” or “unacceptable.” These organized sets of responses are triggered at subcortical centers where specific “programs” for each distinct affect are stored. These programs are innately endowed and have been genetically inherited. (p. 243)

Evolving from James’s cognition and Darwin’s evolution of emotion, Tomkins (1962, 1963) asserted that emotions are the primary motivators of human beings (Abramson, 2015;

Dharwadker, 2015; Gregg & Seigworth, 2010). Affects, stimuli of emotions, make a profound impact on human behavior. According to Tomkins (1962, 1963), humans possess nine innate affects. Each one of them uniquely provides a natural response for us to act and survive. The affect theory organizes these affects into three groups: six negatives (dissmell, disgust, fear, anger, distress/sadness, and shame), one neutral (surprise), and two positives (interest and joy).

Restorative practices are rooted in the psychology of affect (Acosta et al., 2019; Costello et al., 2009; McCluskey et al., 2008). These practices allow participants to be emotional during the process of building community and repairing harm. To develop positive relationships, people need to feel positive about each other and the community (Abramson, 2015). To build a healthy community, the psychology of affect suggests community members do three things (Costello et al., 2009): (1) maximize positive affects; (2) minimize negative affects; (3) express emotions freely. Restorative practices provide a continuum of strategies to facilitate interactions and to achieve these goals. The psychology of affect explains the mechanism of how these restorative strategies improve connectedness among individuals (Higgin, 1987).

The current study examines the positive affect practices of teachers and the positive affect behavior of students as the outcomes of implementing restorative practices. Affect theory provides a theoretical ground for this investigation. Restorative practices focus on improving relationships among students and adults in schools (Costello et al., 2009). Positive interactions between students and teachers in classrooms are desirable outcomes of restorative practices.

With this understanding, the researcher developed a conceptual framework to guide the investigation of the study. Restorative practices facilitate positive interactions between teachers and students. Positive interactions lead to positive affects of individuals, which improves relationships between teachers and students. These interactions can be coded into observable and measurable behavior occurrences: use of affective statements and praises, asking and answering

questions, and responding to requests. Finally, the feedback loop suggests that positive relationships also enhance the implementation of restorative practices and the overall systematic process (Figure 1.1).

Teacher Practices and Student-Teacher Interaction

Literature has suggested that student-teacher interaction is a potent indicator of student behavior and academic outcomes (Birch & Ladd, 1997; Fowler, Banks, Anhalt, Der, & Kalis, 2008; Hamre & Pianta, 2001). Skinner and Belmont (1993) used a model of motivation examining the effects of teacher practices and student engagement. Fourteen teachers for grade 3-5 self-reported their interactions with each student in their classrooms. A total of 144 students reported the perceptions of their interactions with the teachers. Data were collected in the fall and spring during the school year.

Correlational and path analyses revealed that teacher involvement with students had the most influential impact on the student perceptions of their teachers. Teacher involvement with individual students is the essential characteristic of student-teacher interactions for elementary students. Student instructional engagement was significantly related to the teachers' behavior. Positive involvement of teachers with their students was strongly associated with positive emotional engagement reported by the students. This relationship was found to be true for the negative interactions between teachers and students as well. The lack of student engagement or disengagement was also positively related to negative attention from the teachers. Growing research has confirmed the reciprocal effects of teacher and student interactions in classrooms (Skinner & Belmont, 1993).

Sutherland and Wehby (2001) reviewed six empirical studies examining the effects of use opportunities to respond (OTR) to academic requests on student behavioral and academic outcomes. The review concluded that the increased use of OTR led to improved student

academic outcomes and decrease in off-task and disruptive behavior (Camine, 1976; Skinner, Belflore, Mace, Williams-Wilson, & John, 1997; Skinner, Ford, & Yunker, 1991; Skinner & Shapiro, 1989; Skinner, Smith, & McLean, 1994; West, & Sloane, 1986).

Sutherland and his colleagues (Sutherland, Wehby, & Copeland, 2000) examined the effect of teacher use of behavior-specific praises on student on-task behavior. The results suggested that student on-task behavior increased as teachers increased the use of behavior-specific praise. Numerous studies echo this finding on the effect of behavior-specific praise on student on-task behavior (Broden, Bruce, Mitchell, Carter, & Hall, 1970; Ferguson & Houghton, 1992; Hall, Lund, & Jackson, 1968; Hall, Panyan, Rabon, & Broden, 1968).

On the other side, negative interaction patterns between students and teachers intensify inappropriate student behavior (Gunter & Coutinho, 1997; Gunter, Denny, Jack, Shores, & Nelson, 1993; Gunter et al., 1994). Gunter and colleagues discovered that negative interactions between teachers and students with challenging behavior were seven times more likely to occur than positive interactions. Furthermore, students with chronic behavior challenges are more likely to develop negative relationships with their teachers (Ladd & Burgess, 1999). The bidirectional nature of student-teacher interactions indicates that the way a teacher positively or negatively interacts with students can profoundly influence student behavioral responses and vice versa (Doumen et al., 2008).

The ratio between positive and negative feedback a student receiving from a teacher is another reliable indicator of student behavioral outcomes (Jenkins, Floress, & Reinke, 2015). The recommended ratio of positive versus negative feedback of a teacher ranges from 3:1(Shores, Gunter, & Jack, 1993; Wong & Wong, 1998) to 5:1 (Sugai & Horner, 2005).

Positive and negative interactions of teachers with their students strongly associate with student classroom behavior. Pas and colleagues (2015), based on their observations in the

classrooms, found that teachers using more positive recognitions of student behavior resulted in less disruptive student behavior. The ratio of positive to negative student-teacher interactions is critical information to improve teacher practices (Reinke et al., 2016).

Overall, research has suggested many effective teacher practices linking to positive student behavior patterns, such as clearly defined classroom expectations (Rosenberg, 1986; Sprick, 2009), high rate of OTR (Haydon, Mancil, & Van Loan, 2009; Sutherland, Alder, & Gunter, 2003), behavior-specific praise (Ferguson, & Houghton, 1992), pre-correction (Colvin, Sugai, & Patching, 1993; Stormont & Reinke, 2009), and specific behavior correction (McAllister et al., 1969).

Furthermore, researchers have recognized the importance of measuring student-teacher interactions to inform and improve teaching practices. In terms of measuring interactions, rating scales by self-reported surveys, and direct observations are common methods (Reinke, Herman, & Newcomer, 2016). While rating scales are useful for briefly identifying the gap in individual perceptions and the need for interventions, researchers criticize that the method lacks sensitivity and specificity in the direct assessment of teacher practices and student behavior in classrooms (Yoder & Symons, 2010). On the contrary, direct observations allow researchers to repeatedly and more objectively measure the changes in practices over time (Chafouleas, Riley-Tillman, & Christ, 2009; Yoder & Symons, 2010).

The literature review on student-teacher interaction also guided the instrument design of the study in several aspects. First, the restorative practices observation instrument classifies teacher practices and student behavior into positive and negative sessions. Teacher behavior intending to produce positive student affects is coded in an item of positive affect practices. Accordingly, negative teacher behavior is recorded in an item of negative affect practices. Likewise, student behavior is coded and recorded in positive and negative behavior items

(Appendix C, p. 1). The observation protocol (Appendix D) provides more details on data collection.

Besides, research has highlighted the importance of the ratio between positive and negative student-teacher interactions associating with student behavior outcomes (Jenkins, Floress, & Reinke, 2015). Therefore, in the current study, the counts of the occurrences of observed positive and negative behavior are converted into percentages (Appendix C).

Moreover, behavior-specific praise, non-behavior-specific praise, and OTR, along with observable elements of restorative practices, are categorized into the items of teacher positive affect practices. For instance, the restorative language combines behavior-specific praise, affective statements, small impromptu conversations, and positive physical affects (Appendix E).

In the past district evaluations, many restorative practices, such as affective statements, small impromptu conversations, restorative questions, or circles, rarely occur during the scheduled short observation period. It would be challenging to detect any meaningful change if the behavior is barely observed. Combining these practices that use specific non-instructional language (verbal and nonverbal) to promote positive affects increases the likelihood of identifying changes in individual behavior over time.

Evaluations of Restorative Practices

In the 2000s, restorative justice was introduced to educational settings. It soon became popular in schools around the world (McCluskey et al., 2008). With little knowledge of the effects of restorative practice, the Youth Justice Board of England and Wales (Bitel, 2005) commissioned a comprehensive evaluation in the United Kingdom to assess a pilot initiative in 6 primary and 20 secondary schools in the region. The Restorative Justice in Schools Program aimed to reduce bullying, robbery, victimization, and exclusions. Multiple data collection methods were applied throughout the three-year study period, including student and staff

surveys, interviews for students and staff, and administrative data for school demographics and performance. The intervention group was compared with a comparison group on multiple measures.

Staff in the program schools attended introductory training on restorative practices. The training content included active listening, restorative inquiry, circle time, peer mediation, and restorative justice conference. The rest of the implementation primarily relied on internal administrative support from each school (Bitel, 2005).

The study concluded that "restorative justice is not a panacea for the problems in schools" (Bitel, 2005, p.65). The study found insignificant effects on student attitudes and little impact of conferences on exclusions. The staff from program schools reported an improved school environment, while the non-program schools demonstrated a higher reduction in student behavioral incidences. Nevertheless, the researchers stressed, with the correct implementation, the approaches could improve the school learning environment and youth development. The researchers also highlighted the full commitment of the head teachers was the single most paramount factor for program success.

Numbers of studies have focused on examining the responsive components of restorative practices, mainly formal conferencing. Indicated in their findings, although conferencing, as an alternative discipline approach, provides satisfaction for the victim participants, there is little evidence suggesting that this method has an impact on the overall school environment (Bitel, 2005; Blood, 2005).

In 2004, commissioned by the Scottish Executive Education Department (SEED), a team from the University of Edinburgh and Glasgow conducted a two-year collaborative evaluation examining a restorative practices initiative in 18 schools in three local authorities. Unlike the England and Wales evaluation, the schools in this study have the complete autonomy to adapt the

practices and develop goals to meet school contexts and needs. Researchers used a wide range of data collection methods to investigate the influence of the initiative. Qualitative and quantitative data were collected through observations, interviews, focus groups, documentary analysis, and surveys. The outcomes of this initiative, the achievements of the schools, were categorized into four indicators: significant achievement across school, significant achievement in places, early states but evidence of progress, and other priorities dominate (Kane et al., 2007, p. 12).

The evaluation concluded that half of the schools demonstrated strong evidence of improved school community. Staff training is the central component of the program. Visible modeling and support from internal and external experts were vital to successful implementation (Kane et al., 2007).

The Oakland Unified School District (OUSD) is one of the pioneers in the United States that adopted restorative justice in the school district. In 2014, the school district prepared an evaluation report for the Office of Civil Rights and USDOE. This report recorded the successes of its ten years implementation of restorative practices. The district integrated restorative practices with other tiered district initiatives, such as Positive Behavior Interventions and Supports (PBIS) and social-emotional learning, to increase school capacity and program sustainability (OUSD, 2014). Quantitative and qualitative data included student and staff surveys, behavioral data from district databases, and interviews. Self-reported survey data inferred that restorative practices had positive effects on reducing disruptive behavior, resolving conflicts, and improving relationships. The study suggested that restorative practices contributed to significant reductions in suspension rate and racial discipline gap. The report indicated a positive impact of restorative practices on attendance, reading level, and graduation rates comparing between restorative and non-restorative schools.

Empirical Studies on Restorative Practices

Congress passed the Every Student Succeeds Act (ESSA) in December 2015. Under ESSA, schools are required to invest federal funding and improvement efforts on evidence-based practices and interventions. This federal law further categorizes evidence-based interventions into four levels: strong evidence, moderate evidence, promising evidence, and demonstrates a rationale.

Level 1 Strong Evidence must have at least one well-designed and well-implemented experimental study (e.g., a randomized control trial) showing a statistically significant and favorable effect of the intervention on student outcomes. Level 2 Moderate Evidence must have at least one well-designed and well-implemented quasi-experimental study (e.g., matched control group). Level 3 Promising Evidence must have at least one well-designed and well-implemented correlational study with statistical control for selection bias. Finally, level 4 Demonstrates a Rationale should have a well-specified logic model from research or evaluation, and continuing efforts to examine the effects of the intervention (USDOE, 2016). Undoubtedly, there is an urgent demand for rigorous empirical and quasi-experimental studies on restorative practices to meet the requirements of high-level evidence.

Acosta and colleagues (2019) conducted the first study using a randomized controlled trial (RCT) design to examine the outcomes of restorative practices. The study assessed the effects of restorative practices on school connectedness, bullying victimization, and youth development outcomes on middle school students.

The researchers used a self-reported survey to collect student perception data on school climate (school), peer relationship (peer), and youth developmental outcomes (student). Baseline and post-program data were collected from a total of 2771 middle school students to assess the changes in student perceptions due to the intervention. Three-level measures (school, peer, and

student) were compared between the treatment and control groups. The treatment group consisted of seven randomly selected middle schools in the state of Maine. The control group included six matching schools (one school was dropped out of the study in the first year of the implementation) (Acosta et al., 2019).

Multi-layers of interventions were implemented in the treatment schools, including training, monthly consultation by phone, participatory learning groups, site visits. IIRP coaches provided these services or interventions (2019).

After two years of investigation, Acosta and her colleagues (2019) did not detect any significant differences in intended outcome measures (school connectedness, youth development, and bullying) between the treatment and control groups. On a positive note, the study found that students who reported a better experience with restorative also reported more positive outcomes (connectedness, peer relationships, and overall school climate). One limitation of the study was that more than 90% of the participants were white making the results difficult to generalize to a large population across the nation.

One of the primary aims of restorative practices is to enhance interactions between students and adults in the school community. Gregory and her colleagues (2016) studied how restorative practices affected the relationships between minority students and their teachers. They used student and teacher surveys to explore whether the implementation level of restorative practices correlated with student perceptions of being respected by their teachers and whether there were any racial differences in these relationships. They also investigated the effects of restorative practices on discipline referrals and the difference between ethnic groups. In this study, IIRP was the primary provider for whole-staff training, site visits, coaching, and consultation.

Gregory et al.'s (2016) study did not suggest any significant racial differences in restorative practices experience between multiple ethnic student groups, such as Latino, African American, Asian, and White. One surprising finding was that the perceptions of the implementation experience from students and their teachers were not statistically correlated. Student-reported data, but not teacher-reported, suggested that a higher level of implementation of restorative practices significantly predicted fewer discipline referrals. The researchers stressed an increase in the fidelity of program implementation was imperative for improving student-teacher relationships.

Green and colleagues (2019) received funding from the U.S. Department of Justice for a five-year longitudinal study to examine the effectiveness of restorative practices on behavioral and academic outcomes, such as suspensions, bullying, truancy, GPA, safety, and teacher support. The study intends to provide a comprehensive understanding of the challenges and potential of the implementation of restorative practices. A cluster-randomized control trial design is employed. Data collection involves web surveys, school administrative data, document analysis, and interviews. Indicated by their timetable, the results of the study will be disseminated by the end of 2020.

Green et al. (2019) adopted strategies from implementation science to ensure the fidelity of implementation. The treatment schools utilized the Dynamic Adaptation Process, a multidimensional approach based on implementation science, to facilitate an iterative data-driven decision-making process within the school implementation teams. Their research protocol explicitly defined the two-tiers and multiple-stages program implementation and fidelity measures to monitor the implementation process effectively.

Limitations and Future Research

Most studies on this subject have used self-reported data collection methods, either a teacher-reported or a student-reported survey, to provide information on both the independent and dependent variables. It implies s bias towards significant results. For instance, the inconsistency between different data sources indicated the bias or challenges of using self-reported data in evaluating the program outcomes (Gregory et al., 2016).

Furthermore, most research in the field using self-reported surveys is conducted in secondary school settings (Acosta et al., 2019; Green et al., 2019; Gregory et al., 2016). Students in middle and high schools provide more valid perception data compared to elementary students due to the maturity level. As a result, limited literature has examined the impact of restorative practices in elementary school settings.

Moreover, few studies have used quantitative observational data to explore the impact of restorative practices. Most available literature has measured the program outcomes, such as behavioral outcomes and school environment change. Observations from external experts have been considered a gold standard for evaluating a program (Forman et al., 2013). Researchers have been calling for future studies using observational sources to assess the impact of restorative practices at a classroom level (Acosta et al., 2019; Gregory et al., 2016).

It is also difficult to generalize the findings of these studies because the components of implementing restorative practices vary widely. Many studies have measured the effects at one or two-time points across a couple of years. The full implementation of restorative practices may take many years. It is nearly impossible to detect long-term impacts if a program has not been entirely implemented (Gregory et al., 2016).

Green et al. (2019) apply implementation science to monitor implementation quality strategically. However, the assumption of using implementation science is that the program or

practice is evidence-based. Implementation science emphasizes two indivisible components to achieve research suggested outcomes or effects. First, the program must be proven effective. Second, the program must be implemented effectively, as suggested by research (Fixsen, Naoom, Blasé, Fridman, & Wallace, 2005). An evidence-based program or practice requires at least two empirical studies with randomized group designs conducted by different researchers assessing the outcomes of a program or a practice (Fixsen, Blasé, Naoom, & Wallace, 2009). Based on this definition, restorative practices, at this time, cannot be categorized as an evidence-based practice. Green and her colleagues (2019) have not provided any discussion regarding this assumption.

Unfortunately, at this time, very few empirical studies about restorative practices have used randomized control trial design. The majority of the literature has focused on behavioral and academic outcomes based on the data collected through surveys, administrative databases, and interviews. Most of the investigations took place in middle and high schools. Researchers have identified the need for using observational data to assess the outcomes of restorative practices, and enhance the understanding of the practices on classroom levels (Acosta et al., 2019; Gregory et al., 2016).

CHAPTER THREE: METHODOLOGY

Participants and Context

This study was conducted in a large urban public school district in the Southeast. The district serves nearly 100,000 students and employs more than 6,000 teachers. The majority of the student population are non-white, including about 35% African-American. The demographics in the 19-20 school year show that three-fifths of the students are eligible for the Free-Reduced Lunch (FRL) program. More than 100 languages are spoken by students in the district.

Thirty schools in the district have been implementing restorative practices. In the past three years, the district gradually added an additional 10 schools to implement restorative practices as a whole-school approach (10 schools in 17-18, 20 schools in 18-19, and 30 schools in 19-20). Recently, the district has decided to speed up the district-wide implementation, which will add 20 restorative schools each year. With this movement, it is critical to have a profound understanding of the current implementation status of the program to inform implementation strategies rather than entirely focusing on long-term outcomes.

The district annually compares the discipline outcomes between restorative schools and district averages, such as numbers of suspensions, days of suspension, and referrals. However, these data are not adequate to guide further implementation efforts since it takes many years for schools to fully implement the practices. Before the full implementation, it would be difficult to measure the long-term effects. For the 19-20 evaluation of restorative practices, the district decided to use stratified random school selection and a new observation instrument to address the growing needs for future implementation.

Nine restorative elementary schools were selected for the study. According to Title I status, eight are high-poverty schools with more than 80% of students eligible for FRL. The average student enrollment of each school is 409. The percentage of white students ranges from 7% to 49%. For the African-American student population, it ranges from 16% to 85% (Table 3.1).

Table 3.1

Student Enrollment by Race and FRL (school-level)

School[a]	Title I [b] (Y/N)	White (%)	Black (%)	ELL [c] (%)	FRL [d] (%)
1	Y	14	67	12	86
2	Y	29	33	33	79
3	Y	12	35	41	85
4	Y	28	26	40	79
5	Y	36	30	27	81
6	Y	12	60	19	86
7	Y	7	80	14	88
8	Y	8	85	0	88
9	N	49	16	17	44

Note. **a**. Schools are not listed in a specific order. School numbers are not assigned school codes; **b.** Schools receive federal Title I funds are considered as high-poverty schools; **c**. ELL=English Language Learner; **d**. FRL=Free-Reduced Lunch

The overall teacher demographics by gender and race indicate that most teachers are white females (see Table 3.2). The student to teacher ratio ranges from 12:1 to 17:1. This calculation includes teachers in all subjects and positions, such as K-5, special areas, ECE, and resource

teachers. Therefore, the student-to-teacher ratios in the observed classrooms were slightly higher than the report.

Table 3.2

Teachers by Gender and Race (school-level)

School	Female (%)	White (%)	Black (%)	Student-Teacher Ratio
1	80	67	33	12:1
2	83	77	23	12:1
3	83	86	14	12:1
4	77	90	10	15:1
5	91	85	12	14:1
6	81	78	22	12:1
7	77	63	33	15:1
8	86	79	21	17:1
9	84	75	3	13:1

The participants of the study were 140 classroom teachers and their students in the nine randomly selected restorative schools. The sampling section explains the sampling procedure. The demographics of individual teachers and students were not recorded in order to protect their privacy.

The researcher of the study is a district resource teacher in the behavior department. She is also a licensed trainer for restorative practices through IIRP. All observational data for analysis were coded and recorded by the researcher as the only observer. The data used for

assessing interrater reliabilities were collected by the researcher and another district resource teacher in the same department, who is also a licensed trainer for restorative practices.

Sampling Methodology

In fall 2019, the district started an evaluation of restorative practices requested by the behavior department. At the school level, nine restorative schools were selected through stratified random sampling. Sixteen elementary schools in the district have received whole-school training on restorative practices. Three elementary schools received whole-school training in the summer of 2017, six elementary schools were trained for restorative practices in the summer of 2018, and seven elementary schools became restorative schools in the summer of 2019.

The first round of randomization took place at the school level. Three schools were randomly selected from each training group. For the first group, there are only three elementary schools trained in restorative practices three years ago. They automatically become cohort 1. Cohort 2 was comprised of three randomly selected elementary schools among the six trained schools two years ago. Finally, cohort 3 included three schools randomly selected from the seven elementary schools being trained in the following summer. A total of nine elementary schools participated in the district evaluation, three schools for each cohort receiving restorative practices training at separated school years. The random function in Microsoft Excel was used for random selections.

In collaboration with the department coordinator and the district data department, the researcher developed an RP Classroom Observation Tool (Appendix C & D), and an Observation Protocol (Appendix E). This observation tool was used to collect data from all classrooms in the nine randomly selected elementary schools for the district evaluation. The researcher, as a

district resource teacher in the behavior department, was tasked for data collection in these elementary schools.

The district evaluation initially planned for three rounds of observations. However, the global COVID-19 pandemic resulted in statewide school closing starting in March 2020. The third round of data collection could not take place (Appendix B).

Table 3.3

Numbers of Observed Classrooms

School[a]	# of Observed Classrooms		
	Round 1	Round 2	Both
1	10	14	10
2	11	11	9
3	13	19	12
4	11	13	8
5	11	16	10
6	14	15	10
7	10	13	10
8	13	14	12
9	10	12	9
Total	103	127	90

Note. a. Schools are not listed in a specific order. School numbers are not assigned school codes.

The observational data were initially collected for the district evaluation of restorative practices (Table 3.3). Since this data collection was ordered by the district to evaluate its

program, no consent was required at that point. District IRB approval was obtained for the use of the same existing data for the current study.

Table 3.4

Criteria for Random Selection of Participants

Group	Condition Level	Criterion[a]
Intervention groups (1)	3	Teachers received restorative practices training three years ago.
	2	Teachers received restorative practices training two years ago.
	1	Teachers received restorative practices training one year ago.
Comparison group (0)	0	Teachers have never received restorative practices training.

Note. a. The training status of teachers are based on the records of the district RP training list.

Criteria in Table 3.4 were applied to identify qualified participants. A training list from the behavior department recorded the dates when school employees received the initial whole-school training in restorative practices. The initial training year was used to determine teacher training status or condition level. Participants were grouped into four condition levels. The second stage of randomization for participants was initially proposed based on the assumption of adequate observations in each group for random selection. However, unexpected school closing due to Covid-19 resulted in a smaller sample. The random selection of participants could not be performed as proposed. Chapter Four describes the actual sampling procedure.

An independent-samples t-test was selected to investigate the differences in teacher practices and student behavior between the intervention and comparison groups. A priori power analysis using G*Power 3.1.9.7, a software calculating statistical power, indicated a sample size of 156 to detect a medium effect size of .40 with a minimum statistical power of .80.

Furthermore, a mixed-design ANOVA was used to examine the interaction of time and condition level. A priori power analysis suggested a sample size of 76 with a statistical power of .96 to detect a medium effect size of .25.

Table 3.5

Power Analyses for Independent t-Test and Repeated Measures ANOVA

Research Question	**Statistical Analysis**	**Recommended Sample Size and Power**	**Available Observations**
1	Independent t-test	Sample size group 1: 78 Sample size group 2: 78 Effect size d: 0.40 Power (1-β): 0.80	Round 1 (103) & Round 2 (127) observations
2 & 3	Repeated measures ANOVA	Sample size: 76 (19 / Condition Level) Effect size f: 0.25 Power (1-β): 0.95	Classrooms with repeated observations (90)

Intervention

The implementation of restorative practices in the school district included three major components: a two-day whole-school training, monthly consultations, and weekly support from district resource teachers.

The two-day restorative practices training was provided to all school staff during the summer break. Day-one training introduced the basic premise and theories related to restorative practices. The content covered six out of the 11 essential elements of restorative practices (IIRP, 2010), including affective statements, restorative questions, small impromptu conferences, fair process, reintegrative management of shame, and fundamental hypothesis understandings. Day-two training focused on the effective use of restorative circles. Participants had opportunities to learn the types of proactive and responsive circles, develop circle lesson plans, practice circle facilitation, and participate in different circles. Overall, the two-day training covered the majority of the essential practices except for formal restorative conferences.

Six schools from Cohort 1 and 2 received the training from IIRP under the contract between the district and the institution. The contract ended in the 18-19 school year. Therefore, the three schools of Cohort 3 were trained by district resource teachers who were all licensed trainers for restorative practices authorized by IIRP.

Restorative practices schools receive monthly consultation from IIRP and district resource teachers. The consultation service targeted different concentration areas of the practices based on the school's need or the district's recommendation. IIRP coaches provided the consultation for the first two years. District resource teachers continued the service in year three. Therefore, Cohort 1 schools received two years of consultation visits from IIRP and one year from the resource teachers. Cohort 2 schools received one year of consultation from IIRP and one year from the resource teachers. Cohort 3 schools only received consultation from the

resource teachers. The activities during these consultations varied greatly, such as classroom observations, Professional Learning Community (PLC) meetings, individual coaching, and administrative meetings.

Finally, district resource teachers provided weekly support to all restorative practices schools. Again, the support varied among schools, from on-demand training to teacher coaching to feedback to administrators.

The sampling process identified and grouped participants according to the initial training year. However, it was unknown whether these teachers directly participated in monthly consultation or weekly district support. This study views the overall experience of restorative practices as the intervention. Different periods of these experiences indicated different dosages of the intervention.

All participants have received some level of intervention. They all experienced restorative practices to some degree. The experience of restorative practices as the intervention was leveled into groups to examine whether there were differences in the outcomes between different intervention levels. The training was a standard measure to classify the experience of the participants. Therefore, teacher training status was used to identify intervention groups.

Teachers in the comparison group (coded as 0) had never received official training about restorative practices. However, they might have some experience in school. They might exercise restorative practices as well. Teachers in the intervention group (coded as 1) received the initial school-wide restorative practices training and were considered formally implementing the intervention.

Teachers were further grouped into one of the four condition levels: trained three years ago (coded as 3), trained two years ago (coded as 2), trained one year ago (coded as 1), and not trained (coded as 0). For instance, the intervention for level 3 group was training and three years

of in-school experience. The intervention for level 0 group was in-school restorative practice experience (Table 3.4). It was hypothesized that teachers with different levels of experience of restorative practices would have different outcomes in teacher practices and student outcomes. There were four dosages of the intervention to test the progressive differences.

Instrument Development

The current study only used the data collected through the first page of the observation instrument in Appendix B. This portion of the instrument aims to capture the interaction between the teachers and students. To translate the interaction into observable and measurable variables, the researcher looked into two sources: the theoretical foundation of restorative practices, and a well-established observational instrument measuring interaction.

First, the theoretical base of this instrument development is the affect theory that constitutes the theoretical ground of restorative practices (Costello et al., 2009; Acosta et al., 2019; McCluskey et al., 2008). As discussed in previous chapters, affect theory categorizes human affects into two primary groups: positive and negative. The line between positive and negative affects is the neutral affect, surprise/startle. According to IIRP, the psychology of affect suggests that individuals maximize positive affects and minimize negative affects to build relationships within a community. The current instrument classifies the teacher practices and student behavior into two broad categories: positive affect and negative affect (Appendix C).

Few observation instruments in the field have measured direct interaction in classroom settings. Dyadic Parent-Child Interaction Coding System – Third Edition (DPICS 3rd ED) was developed for Parent-Child Interaction Therapy (PCIT) (Eyberg, Nelson, Duke, & Boggs, 2005). This evidence-based program (Shriver & Allen, 2008; Eyberg, Nelson, & Boggs, 2008), including its coding system, has been adapted for Teacher-Child Interaction Training (TCIT)

program in early childhood setting (Tiano & McNeil, 2006; Lyon, Gershenson, Farahmand, Thaxter, Behling, & Budd, 2009; Gershenson, Lyon, & Budd, 2010).

Aspects of DPICS support the current instrument design. For instance, some coded adults' behavior in DPICS includes direct command, labeled and unlabeled praise, reflective statements, question, and negative talk. The current instrument includes restorative language (affective statements, behavior-specific praise, small conversation for building relationships), non-specific praise, Opportunities to Respond (OTR), and negative language and physical affects. Definitions for independent variables of the study are provided in the instrument protocol (Appendix E).

For child behavior, DPICS suggests compliance or noncompliance, answer to questions / no answer to questions, yelling, and destructive behavior. The current instrument includes initiating interaction (Lebuffe & Naglieri, 1999a), answering questions (positively or negatively), responding to requests (positively or negatively), distractive voice level, or movement (Appendix C).

Almost all observations in the TCIT studies were coded in 10-second intervals (Tiano & McNeil, 2006; Lyon et al., 2009; Gershenson et al., 2010). This method requires extensive training for numbers of observers. Hence, it was not adopted in this research. For the current study, the researcher, the only observer, counted the occurrences of each coded behavior, which were later converted into percentages to standardize the metrics across classrooms. The percentages of behavior occurrences were used for statistical analyses.

To check the inter-rater reliability, the researcher and a co-worker observed 16 classrooms together. The instrument demonstrated acceptable inter-rater reliability (Appendix F). The first page of Appendix F was based on six elementary classroom observations conducted in December 2019. Koo and Li (2016) provided a guideline for interpretation of intraclass

correlation: below .50 – poor; between .50 and .75 – moderate; between .75 and .90 – good; above .90 – excellent. According to this guideline, the current instrument demonstrated moderate reliability for the teacher observation (ICC = .546) and excellent reliability for the student observation (ICC=. 950). As requested by the district, the same two raters conducted ten additional observations in secondary classrooms in January. The 16 observations provided good reliability for the teacher observation (ICC=.755) and excellent reliability for the student observation (ICC= .934).

Internal consistency reliability was not examined prior to the official data collection due to the small sample size. After the IRB approval, internal consistency was assessed for two subscales and measured with Cronbach's alpha: positive teacher-student interaction and negative teacher-student interaction. The subscale of positive teacher-student interaction was composed of 11 items from positive teacher practices and positive student behavior. The subscale of negative teacher-student interaction included seven items from negative teacher practices and negative student behavior. The commonly cited acceptable range of a Cronbach's alpha value is .70 or above (Nunnally & Bernstein, 1994). Based on the sample of 230 observations, the subscale scores of negative interaction revealed acceptable reliability (Cronbach's $\alpha = .854$). However, the subscale scores of positive interaction showed low internal consistency (Cronbach's $\alpha = .504$). More efforts are needed to revise the observation instrument for better internal consistency reliability. This limitation was further discussed in Chapter Five.

Data Collection

The current study required two sources of data: the observational data from the district evaluation and the training list from the behavior department. Teacher training status indicated in the training list was used as criteria to define different intervention groups. The following

description in this section explains the data collection procedure of the district evaluation. A large portion of these data has been used for the analyses of current research.

As aforementioned, an observation tool and clearly defined protocol were developed by the researcher and other district personnel prior to the data collection. The researcher and a colleague used the observation instrument in field practices and collected data for reliability checks. The second rater received one-hour training provided by the researcher. The content of the training was to review and discuss the RP observation instrument and protocol. Both the researcher and the second rater are licensed trainers for restorative practices through IIRP.

The two raters conducted eight classroom observations in a local elementary school together. The first two observations were field practices. At the end of each observation, raters compared notes and clarified parameters for each observation item. The observational data for the next six elementary classrooms were used for reliability checks. Observers did not discuss the data for these observations. In January 2020, the same raters observed ten additional classrooms in a secondary school: five middle school classrooms, and five high school classrooms. Observational data from a total of 16 classrooms were analyzed to assess the interrater reliability of the instrument (Appendix F).

The researcher was the only observer for the data collection in elementary schools for the district evaluation. Classroom observations were conducted in December 2019 and February 2020. The third round of observations initially scheduled in April was canceled due to the Covid-19 pandemic (Appendix B). School principals received an email notification a week before the visits. The observation instrument was shared with the principals in the initial email in December 2019. They were encouraged to share the document with teachers. However, it was unknown if teachers were aware of the observation content. At the first arrival of the researcher, schools

provided class schedules, building maps, and keys to access to classrooms. These school documents were kept for both visits.

Each observation took ten minutes and conducted in a general education classroom with a regular room teacher. All observations took place in the morning between 9 am and 12 pm of a typical school day. Special areas, Exceptional Child Education (ECE), rooms with substitute teachers, and classrooms unavailable due to testing or accommodations were not observed.

On an observation day, the researcher coordinated a sequence of room visits based on the information on class schedules. The goal was to maximize the number of observations by avoiding lunchtime, testing, teacher planning period, and substitute rooms.

Required information at the top of the instrument was filled in right before entering a class (Appendix C). The observer did not have any conversation with teachers or students unless there was a request. Tally marks were recorded in corresponding items as the observer coded each interaction between the teacher and students. Eleven observational items for each observation were coded and documented. The observation protocol was followed (Appendix E). Data were recorded on hard copies of the observation tool.

After all observations, recorded data were entered in web surveys in Survey Monkey created by the district data department. The hard copies of the data were kept in a double-locked cabinet in the behavior department with the only access from the researcher. Table 3.3 indicates the numbers of observed classes in each round of observations. At the time of the observations, the observer had no information regarding teacher training status.

For the current study, observational data and participant training status were coded and entered in Excel files. Each case number represented a classroom observation. Observational data were matched prior to data entering. The identifiable information of participants was not

recorded. The data sets were imported and analyzed using SPSS 26, a statistics software package.

Data Analysis

Raw data included the counts of the occurrences for observation items. These counts were converted into percentages for further analyses (Table 3.6 & Table 3.7). The observation contexts were vastly different from classroom to classroom and from content to content. The use of percentage, instead of count, allowed some control over the variances to make the data comparable between classrooms. Details about definitions and observation procedures were provided in the instrument protocol (Appendix E).

The first research question evaluated whether there were different outcomes in teacher practices and student behavior between the intervention and comparison groups. As indicated in the sampling section, teachers were identified for intervention and comparison groups according to their training status. The second and third questions explored differences between the four condition levels and differences within the group between times. The last question explored the correlations between teacher practices and student behavior.

Table 3.8 lists 14 outcome variables and three independent variables for the study. Teacher negative physical affect was not included due to low inter-rater reliability (Appendix F).

Table 3. 6

Converting Total Counts into Percentages (Teacher Observation)

Teacher Observation		**Count**		**% of Total**	
Positive Affect Practice	**(T_PAP)**	**#_T_PAP**	=#_T_RL+#_T_NSP+#_T_OTR	**%_T_PAP**	**=** #_PAP **/ #_Total_TP**
Restorative Language	(T_RL)	#_T_RL	Total number of recorded occurrences of the practice	%_T_TL	= #_T_RL / **#_Total_TP**
Non-Specific Praise	T_NSP	#_T_NSP	Total number of recorded occurrences of the practice	%_T_NSP	= #_T_NSP/ **#_Total_TP**
Opportunity-to-Respond (content-related)	T_OTR	#_T_OTR	Total number of recorded occurrences of the practice	%_T_OTR	= #_T_OTR/ **#_Total_TP**
Negative Affect Practice	**T_NAP**	**#_T_NAP**	= #_T_NLA + #_T_NLA	**%_T_NAP**	= #_T_NAP / **#_Total_TP**
Negative Language Affect	T_NLA	#_T_NLA	Total number of recorded occurrences of the practice	%_T_NLA	= #_T_NLA/ **#_Total_TP**
Negative Physical Affect	T_NPA	#_T_NPA	Total number of recorded occurrences of the practice	%_T_NAP	= #_T_NPA/ **#_Total_TP**
	Total	**#_Total_TP**	= #_T_PAP+#_T_NAP	**100%**	= %_T_PAP + %_T_NAP

Table 3. 7

Converting Total Counts into Percentages (Student Observation)

Student Observation		**Count**		**% of Total**	
Positive Affect Behavior	**S_PAB**	**#_S_PAB**	=#_S_II+#_S_RTQ+#_S_PRBR	**%_S_PAB**	**=** #_S_PAB **/ #_Total_SB**
Initiate an Interaction (with teacher)	S_II	#_S_II	Total number of recorded occurrences of the behavior	%_S_II	= #_S_II / **#_Total_SB**
Respond to Content-Related Questions	S_RTQ	#_S_RTQ	Total number of recorded occurrences of the behavior	%_S_RTQ	= #_S_RTQ/ **#_Total_SB**
Positively Respond to Behavioral Request	S_PRBR	#_S_PRBR	Total number of recorded occurrences of the behavior	%_S_PRBR	= #_S_PRBR/ **#_Total_SB**
Negative Affect Behavior	**S_NAB**	**#_S_NAB**	= #_S_DVL+#_S_DM+#_S_NRBR	**%_S_NAB**	= #_S_NAB / **#_Total_SB**
Distracting Voice Level	S_DVL	#_S_DVL	Total number of recorded occurrences of the behavior	%_S_DVL	= #_S_DVL / **#_Total_SB**
Distracting Movement	S_DM	#_S_DM	Total number of recorded occurrences of the behavior	%_S_DM	= #_S_DM / **#_Total_SB**
Negatively Respond to Behavioral Request	S_NRBR	#_T_NPA	Total number of recorded occurrences of the behavior	%_S_NRBR	= #_S_NRBR/ **#_Total_SB**
	Total	**#_Total_SB**	= #_S_PAB+#_S_NAB	**100%**	= %_S_PAB + %_S_NAB

Table 3.8

Dependent and Independent Variables

Variable [a]	Definition [d]	Type
%_T_PAP	Percentage of teacher positive affect practices	DV[b], Continuous
%_T_RL	Percentage of teacher use of restorative language	DV, Continuous
%_T_NSP	Percentage of teacher use of non-specific praise	DV, Continuous
%_T_OTR	Percentage of instructional opportunities to respond	DV, Continuous
%_T_NAP	Percentage of teacher negative affect practices	DV, Continuous
%_T_NLA	Percentage of teacher negative language affect	DV, Continuous
%_S_PAB	Percentage of student positive affect behavior	DV, Continuous
%_S_II	Percentage of student initiating interaction with teacher	DV, Continuous
%_S_RTQ	Percentage of student responding instructional questions	DV, Continuous
%_S_PRBR	Percentage of student positive responses to teacher's behavioral requests	DV, Continuous
%_S_NAB	Percentage of student negative affect behavior	DV, Continuous
%_S_DVL	Percentage of student displaying interruptive voice levels	DV, Continuous
%_S_DM	Percentage of student displaying interruptive movements	DV, Continuous
%_S_NRBR	Percentage of student negative responses to teacher's behavioral requests	DV, Continuous
Intervention	Intervention: Classrooms with teachers who received initial RP[e] training; Comparison: Classrooms with teachers who have never received RP training.	IV[c], Categorical
Condition level	0: Classrooms with teachers never received RP training; 1: Classrooms with teachers received RP training 1 year ago; 2: Classrooms with teachers received RP training 2 years ago; 3: Classrooms with teachers received RP training 3 years ago;	IV, Categorical
Time	Time 1: First observation in Dec. 2019; Time 2: Second observation in Feb. 2020	IV, Categorical

Note. **a.** Negative physical affect (%_T_NPA) is not assessed in the study due to a low reliability; **b.** DV=Dependent Variable; **c.** IV=Independent Variable; **d.** Detailed definitions of each variable can be found in Appendix E; **e.** RP = Restorative practices

Research Question 1: Are there differences in teacher practices and student behavior between classrooms where teachers have been trained in restorative practices and classrooms where teachers have not been trained in restorative practices?

Independent t-tests were selected to investigate whether there were mean differences between the intervention and comparison groups in each item of teacher practices and student behavior. The intervention group was comprised of randomly selected observations with teachers who have been trained for restorative practices. The comparison group included observations with teachers who have never officially received training.

Research Question 2: Are there differences in teacher practices and student behavior among classrooms with teachers in different intervention condition levels (trained 3 years ago/ 2 years ago/ 1 year ago/ not trained)?

Research Question 3: Do teacher practices and student behavior change overtime differently between the four condition levels?

Questions 2 and 3 investigated the differences between different condition groups and within the groups across two observations. Participants were further grouped into four condition levels: teachers trained three years ago (coded as 3), teachers trained two years ago (coded as 2), teachers trained one year ago (coded as 1), and teachers never trained (coded as 0).

Dependent variables were evaluated using a mixed-design ANOVA with one four-level between-subject factor, condition level (fixed factor), and one two-level within-subject factor, time (random factor).

Research Question 4: What are the relationships between teacher practices and student behavior?

Correlation analyses were conducted to examine the correlations of the variables between teacher practices and student behavior.

CHAPTER FOUR: RESULTS

Study Approval

The study was approved by the Institutional Review Board of Bellarmine University, IRB # 855. Also, the Institution Review Board of the local school district provided permission to use the district evaluation data for current research, IRB # 422.

Sampling

Figure 4.1 illustrates the sampling process to determine the subsamples for independent-samples test and repeated measures ANOVA. The district conducted two rounds of observations in December 2019 and February 2020. A total of 230 observations involved 4714 students and 140 teachers from 9 elementary restorative schools. Ninety teachers or classrooms were observed twice.

A priori power analysis for an independent-samples t-test recommended a sample size of 156, 78 for each intervention and comparison group, for a statistical power of .80 with an effect size of .40. Among all observations, 71 were qualified for the comparison group. One hundred and fifty-nine observations met the criteria of the intervention group. The decision was made to randomly split these observations into two random intervention groups, which allowed the researcher to replicate the test.

The replication was not part of the proposed method. However, this extension strengthened the validity of the findings. Due to assumption violation and small sample sizes, a nonparametric test alternative to the independent t-test was performed twice using a subsample size of 150 and a subsample size of 151 to examine the mean differences in teacher practices and

student behavior between each random intervention group and the comparison group (Figure 4.1).

Furthermore, the power analysis suggested a minimum of 19 units for each condition level to achieve a statistical power of 95% for an effect size of .10. Random selection was not performed due to the limited observations for each condition group. Therefore, the available sample size was sufficient. The final sample for the analysis comprised 90 classrooms with repeated observations: 19 for condition level 3, 20 for condition level 2, 27 for condition level 1, and 24 for condition level 0 (Figure 4.1).

Moreover, correlation analysis was conducted on the entire sample (n=230) to explore the relationships between teacher practices and student behavior. No additional sampling procedure was performed.

Finally, Table 4.1 provides descriptive statistics for the entire sample (n = 230) by count and percentage. Raw data were collected by recording the behavior occurrences. Those counts were later converted into percentages indicated in Table 3.6 and Table 3.7 to control the variations between classrooms, contents, and contexts. All analyses of the study were conducted on the data in the form of percentages instead of counts. Nevertheless, Table 4.1 indicates that this data transformation somewhat shifted the shapes of the distributions.

Figure 4.1

Sampling Process to Determine Final Subsamples for Analyses

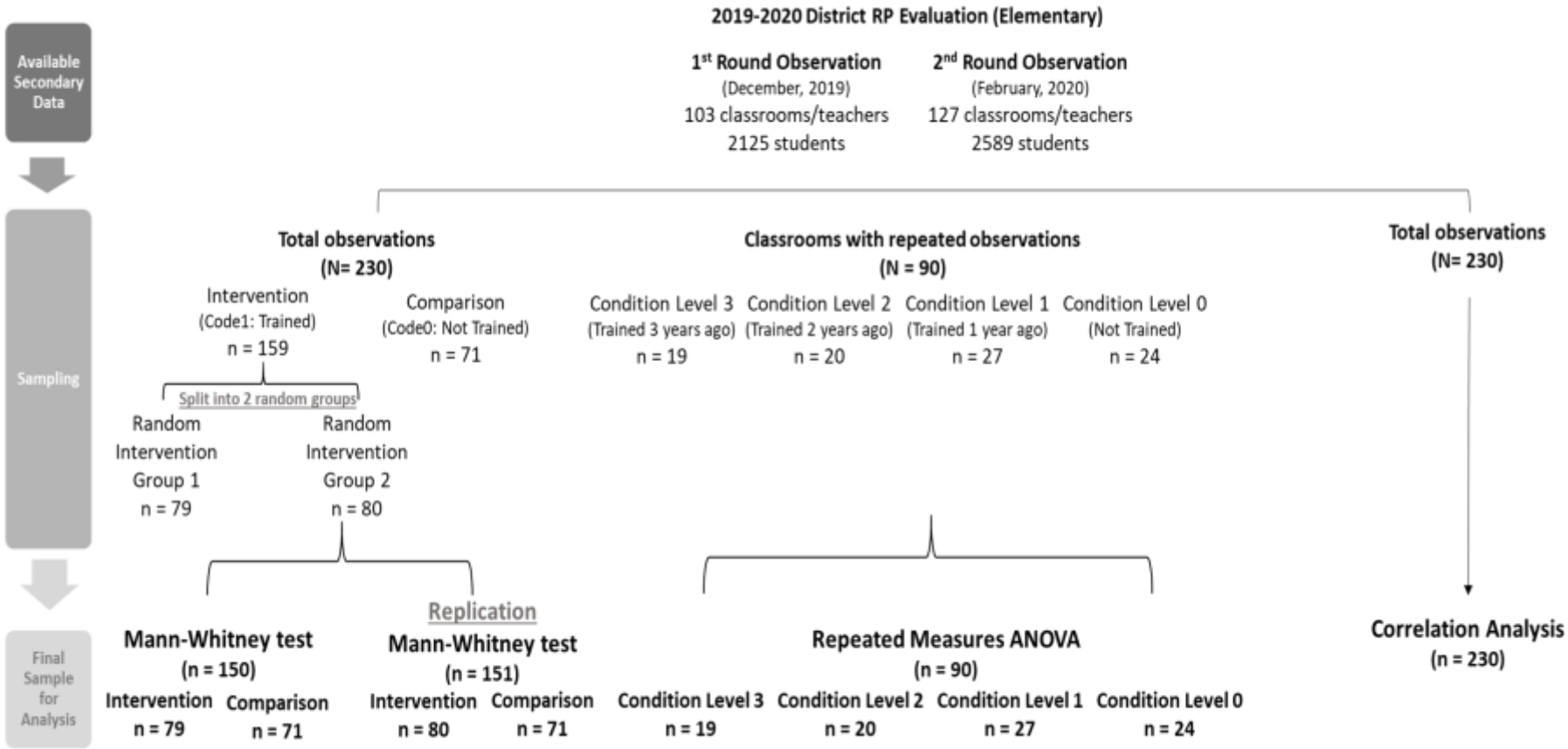


Table 4. 1

Descriptive Statistics for Total Sample (n=230) by Count and Percentage

n = 230	By Count					By Percentage (%)				
	M	*SD*	Range	Skewness	Kurtosis	*M*	*SD*	Range	Skewness	Kurtosis
T_PAP	17.59	12.23	82.00	1.68	4.25	82.10	21.46	90.48	-1.41	1.45
T_RL	3.64	3.30	18.00	1.36	2.33	19.13	16.18	85.71	1.13	1.41
T_NSP	4.92	4.33	25.00	1.49	3.18	22.87	14.71	66.67	.36	-.42
T_OTR	9.04	8.04	42.00	1.40	1.88	40.10	22.50	94.74	-.03	-.74
T_NAP	3.39	4.60	29.00	2.22	6.44	17.91	21.46	90.48	1.41	1.45
T_NLA	2.15	3.42	23.00	2.73	10.17	10.97	15.63	78.95	1.82	3.19
S_PAB	20.00	9.90	45.00	.51	-.48	81.90	19.09	87.50	-1.31	1.27
S_II	2.91	2.52	17.00	1.63	4.49	14.05	13.67	100	2.21	8.34
S_RTQ	9.54	8.19	38.00	1.05	.612	35.98	23.44	89.47	.25	-.87
S_PRBR	7.56	5.16	27.00	1.03	.962	31.88	18.60	91.67	.75	.33
S_NAB	4.22	4.87	23.00	1.71	2.95	18.10	19.09	87.50	1.31	1.27
S_DVL	2.51	2.85	15.00	1.63	3.43	10.90	12.08	75.00	1.571	3.71
S_DM	.80	1.34	6.00	1.94	3.41	3.34	5.52	30.77	1.91	3.77
S_NRBR	.92	1.75	10.00	2.69	7.66	3.87	6.64	35.00	2.05	4.00
Total_TP	20.98	12.20	82.00	1.57	4.03					
Total_SB	24.23	9.92	52	.41	-.24					

Research Question 1

Question 1: Are there differences in teacher practices and student behavior between classrooms where teachers have been trained in restorative practices and classrooms where teachers have not been trained in restorative practices?

Independent-Samples Test (n = 150)

The proposed analysis for this question was independent-samples t-test to compare means between intervention and comparison groups for each outcome measure. However, due to the violation of assumptions, a nonparametric test was used instead. The first random intervention group contained 79 classroom observations. Table 4.2 reveals the numbers of observations for the comparison group and random intervention group 1 from each school.

Table 4.2

Numbers of Observations for Comparison Group and Random Intervention Group 1

School[a]	Comparison	Intervention	Total
1	12	5	17
2	5	7	12
3	17	9	26
4	6	9	15
5	3	12	15
6	15	6	21
7	4	9	13
8	2	17	19
9	7	5	12
Total	71	79	150

Note. a. Schools are not listed in a specific order. School numbers are not assigned school codes.

The majority of the variables demonstrated homogeneity of variances, as assessed by Levene's test ($p>.05$), except for S_RTQ ($p=.001$), S_PRBR ($p=.039$), and S_NRBR ($p=.047$). However, the residuals of most dependent variables did not meet the normality assumption, as assessed by the Shapiro-Wilk test ($P < .05$) and histogram (Table 4.3). Attempts to eliminate the assumption violation via data transformation were not successful. The decision was made to apply the Mann-Whitney U test, a nonparametric test, to all variables due to the assumption violation and small sample size.

Teachers in the intervention group demonstrated higher percentage use of positive practices in most measures compared to teachers in the comparison group: overall teacher positive affect practices (RIG1: M = 84.30%, SD = 19.89%; Comparison: M = 81.80%, SD = 20.70%), restorative language (RIG1: M =22.88%, SD = 17.92%; Comparison: 17.48%, SD = 15.65%; p = .049), and non-specific praise (RIG1: M = 23.23%, SD = 15.11%; Comparison: M = 22.44%, SD = 14.20%). However, teachers in the comparison group provided about 4% more instructional opportunities for students to respond (M = 41.88%, SD = 21.37%) compared to teachers in the intervention group (M = 38.18%, SD = 24.40%). In terms of negative practices, teachers in the intervention group demonstrated less negative affect practices (RIG1: M = 15.70%, SD: 19.89%; Comparison: M = 18.21%, SD: 20.70%) than the comparison group.

Table 4.3

Descriptive Statistics for Measures, Levene's Test for Equality of Variances and Normality Test for Residuals (Comparison Group / Random Intervention Group 1)

Measure	*M* (*SD*)		*MD*[a]	Levene's Test *F* (*p*)	Shapiro-Wilk (*p*)	
	Comp[c]	RIG1[d]			Comp	RIG1
%_T_PAP[b]	81.80	84.30	2.5	.81	.83	.79
	(20.70)	(19.89)		(.368)	(.000)	(.000)
%_T_RL[b]	17.48	22.88	5.4	1.37	.88	.93
	(15.65)	(17.92)		(.243)	(.000)	(.000)
%_T_NSP	22.44	23.23	.79	.67	.96	.97
	(14.20)	(15.11)		(.414)	(.022)	(.081)
%_T_OTR	41.88	38.18	-3.7	1.17	.98	.96
	(21.37)	(24.40)		(.282)	(.257)	(.022)
%_T_NAP[b]	18.21	15.70	-2.51	.81	.83	.79
	(20.70)	(19.89)		(.368)	(.000)	(.000)
%_T_NLA[b]	10.84	10.18	-.66	.27	.77	.67
	(14.42)	(16.64)		(.606)	(.000)	(.000)
%_S_PAB[b]	81.65	85.78	4.13	.95	.89	.85
	(17.48)	(15.42)		(.332)	(.000)	(.000)
%_S_II[b]	13.98	16.49	2.51	.73	.87	.78
	(12.95)	(16.89)		(.395)	(.000)	(.000)
%_S_RTQ	35.94	36.73	.79	11.96	.98	.94
	(19.86)	(27.12)		(.001)[e]	(.528)	(.001)
%_S_PRBR	31.73	32.57	.84	4.32	.97	.95
	(17.22)	(21.10)		(.039)[e]	(.101)	(.003)
%_S_NAB[b]	18.35	14.22	-4.13	.95	.89	.85
	(17.48)	(15.42)		(.332)	(.000)	(.000)
%_S_DVL[b]	10.50	9.10	-1.4	.27	.85	.83
	(10.40)	(10.53)		(.602)	(.000)	(.000)
%_S_DM[b]	3.17	2.35	-.82	1.20	.69	.59
	(4.91)	(4.49)		(.275)	(.000)	(.000)
%_S_NRBR[b]	4.68	2.77	-1.91	4.00	.73	.61
	(6.76)	(5.23)		(.047)[e]	(.000)	(.000)

Note. **a.** Mean difference between intervention group and comparison group. MD = $M_{intervention} - M_{comparison}$; **b.** $p < .05$ for Shapiro-Wilk test for both comparison and intervention groups. The assumption of normality is not met; **c.** Comp = Comparison Group; **d.** RIG = Random Intervention Group; **e.** Equal variances are not assumed.

Students in the intervention group showed more positive behavior than students in the comparison group: overall student positive affect behavior (RIG1: M = 85.78%, SD = 15.42%; Comparison: M = 81.65%, SD = 17.48%), initiating interaction (RIG1: M = 16.49%, SD = 16.89%; Comparison: M = 13.98, SD = 12.95%), responding to instructional questions (RIG1: M = 36.73%, SD = 27.12%; Comparison: M = 35.94%, SD = 19.86%), and positively responding to behavioral requests (RIG1: M = 32.57%, SD = 21.10%; Comparison: M = 31.73%, SD = 17.22%).

Consistently, students in the intervention group demonstrated less negative behavior compared to students in the comparison group on each negative behavior measure: overall student negative affect behavior (RIG1: M = 14.22%, SD = 15.42%; Comparison: M = 18.35%, SD = 17.48%), distracting voice level (RIG1: M = 9.10%, SD = 10.53%; Comparison: M = 10.50%, SD = 10.40%), distracting movement (RIG1: M = 2.35%, SD = 4.49%; Comparison: M = 3.17%, SD = 4.91%), and negatively responding to behavioral requests (RIG1: M = 2.77%, SD = 5.23; Comparison: M = 4.68, SD = 6.76%; p = .026). However, most of the mean differences of these measures between the random intervention group 1 and the comparison group were not statistically significant.

The Mann-Whitney test revealed statistically significant differences between the comparison and intervention groups in two measures (Table 4.4). First, teachers in the intervention group (Mdn = 22.22, Mean Rank 82.10) demonstrated more use of restorative language (%_T_RL) than teachers in the comparison group (Mdn = 13.79, Mean Rank 68.15). A Mann-Whitney test indicated that this difference was statistically significant (U (79, 71) = 2283.00, z = - 1.964, p = .049).

Moreover, students in the intervention group (Mdn = .00, Mean Rank 68.89) displayed less negative responses to teacher's behavioral requests compared to students in the comparison group (Mdn = .00, Mean Rank 82.85). A Mann-Whitney test suggested that the difference was statistically significant (U (79, 71) = 2282.50, z = -2.229, p = .026).

Table 4.4

Mann-Whitney U Test (Comparison Group, n=71 / Random Intervention Group 1, n=79)

Measure	Mann-Whitney *U*	*Z*	*p*
%_T_PAP	2616.00	-.728	.467
%_T_RL	2283.00*	-1.964	**.049**
%_T_NSP	2707.50	-.365	.715
%_T_OTR	2578.50	-.851	.395
%_T_NAP	2616.00	-.728	.467
%_T_NLA	2572.00	-.923	.356
%_S_PAB	2400.50	-1.534	.125
%_S_II	2586.00	-.823	.410
%_S_RTQ	2804.50	.000	1.000
%_S_PRBR	2769.00	-.134	.894
%_S_NAB	2400.50	-1.534	.125
%_S_DVL	2511.50	-1.128	.259
%_S_DM	2499.50	-1.387	.166
%_S_NRBR	2282.50*	-2.229	**.026**

Note. *p < .05

Replicating Independent-Samples Test (n = 151)

The same analysis procedure was applied using the second random intervention group. This random intervention group contained 80 observations (Table 4.5). The replication test intended to check the reliability of the results. The researcher was interested in knowing whether the results were consistent using a different random sample.

Table 4.5

Numbers of Observations for Comparison Group and Random Intervention Group 2

School[a]	Comparison	Intervention	Total
1	12	7	19
2	5	10	15
3	17	6	23
4	6	8	14
5	3	12	15
6	15	7	22
7	4	10	14
8	2	10	12
9	7	10	17
Total	71	80	151

Note. a. Schools are not listed in a specific order. School numbers are not assigned school codes.

Similarly, the majority of the variables demonstrated homogeneity of variances assessed by Levene's test ($p > .05$), except for S_PAB ($p = .016$), S_II ($p = .033$), and S_NAB ($p = .016$). The residuals of most dependent variables did not meet the normality assumption, as assessed by the Shapiro-Wilk test ($P < .05$) and histogram (Table 4.6). Attempts to eliminate the assumption

violation via data transformation were not successful. The decision was made to apply the Mann-Whitney U test to all variables due to the assumption violation and small sample size.

Teachers in the second random intervention group demonstrated marginally less positive practices than teachers in the comparison group in three out of four positive practices measures: overall teacher positive affect practices (RIG2: M = 80.18%, SD = 23.57%; Comparison: M = 81.80%, SD = 81.80%, SD = 20.70%), restorative language (RIG2: M = 16.88%, SD = 14.25%; Comparison: M = 17.48%, SD = 15.65%), and opportunities to respond (RIG2: M = 40.42%, SD = 21.63%; Comparison: M = 41.88%, SD = 21.37%). Teachers in the intervention group (M = 22.88%, SD = 14.92%) used slightly more non-specific praises than teachers in the comparison group (M = 22.44%, SD = 14.20%).

Furthermore, teachers in the intervention group (M = 19.82%, SD = 23.57%) showed about 2% more negative practices than teachers in the comparison group (M = 18.21%, SD = 20.70%). However, the Mann-Whitney test did not indicate any statistical significance in these differences (Table 4.7).

Students in the intervention group (M = 78.30%, SD = 22.85%) appeared 3% less overall positive behavior than students in the comparison group (M = 81.65%, SD = 17.48%), including initiating interaction (RIG2: M = 11.71%, SD = 10.02%; Comparison: M = 13.98%, SD = 12.95%), responding to instructional questions (RIG2: M = 35.26%, SD = 22.73%; Comparison: M = 35.94%, SD = 19.86%), and positively responding to behavioral requests (RIG2: M = 31.33%, SD = 17.32; Comparison: M = 31.73%, SD = 17.22%).

Table 4.6

Descriptive Statistics for Measures, Levene's Test for Equality of Variances and Normality Test for Residuals (Comparison Group and Random Intervention Group 2)

Measure	*M* (*SD*)		*MD*[a]	Levene's Test *F (p)*	Shapiro-Wilk (*p*)	
	Comp[c]	RIG2[d]			Comp	RIG2
%_T_PAP[b]	81.80	80.18	-1.62	.00	.83	.80
	(20.70)	(23.57)		(.966)	(.000)	(.000)
%_T_RL[b]	17.48	16.88	-.60	.13	.876	.912
	(15.65)	(14.25)		(.723)	(.000)	(.000)
%_T_NSP	22.44	22.88	.44	.54	.96	.97
	(14.20)	(14.92)		(.465)	(.022)	(.038)
%_T_OTR	41.88	40.42	-.46	.17	.98	.96
	(21.37)	(21.63)		(.681)	(.257)	(.010)
%_T_NAP[b]	18.21	19.82	1.61	.00	.83	.80
	(20.70)	(23.57)		(.966)	(.000)	(.000)
%_T_NLA[b]	10.84	11.87	1.03	.29	.77	.76
	(14.42)	(15.79)		(.588)	(.000)	(.000)
%_S_PAB[b]	81.65	78.30	-3.35	5.93	.89	.85
	(17.48)	(22.85)		(.016)[e]	(.000)	(.000)
%_S_II[2]	13.98	11.71	-2.27	4.63	.87	.89
	(12.95)	(10.02)		(.033)[e]	(.000)	(.000)
%_S_RTQ	35.94	35.26	-.68	3.02	.98	.96
	(19.86)	(22.73)		(.084)	(.528)	(.014)
%_S_PRBR	31.73	31.33	-.40	.17	.97	.94
	(17.22)	(17.32)		(.682)	(.101)	(.001)
%_S_NAB[b]	18.35	21.70	3.35	5.93	.89	.85
	(17.48)	(22.85)		(.016)[e]	(.000)	(.000)
%_S_DVL[b]	10.50	13.02	2.52	3.97	.85	.83
	(10.40)	(14.48)		(.048)	(.000)	(.000)
%_S_DM[b]	3.17	4.46	1.29	3.21	.69	.72
	(4.91)	(6.69)		(.075)	(.000)	(.000)
%_S_NRBR[b]	4.68	4.22	-.46	.27	.73	.63
	(6.76)	(7.65)		(.604)	(.000)	(.000)

Note. **a.** Mean difference between intervention group and comparison group. MD = $M_{intervention} - M_{comparison}$; **b.** $p < .05$ for Shapiro-Wilk test for both comparison and intervention groups. The assumption of normality is not met; **c.** Comp = Comparison Group; **d.** RIG = Random Intervention Group; **e.** Equal variances are not assumed.

Table 4.7

Mann-Whitney U Test (Comparison Group, n=71 /Random Intervention Group 2, n=80)

Measure	Mann-Whitney *U*	*Z*	*p*
%_T_PAP	2748.00	-.348	.728
%_T_RL	2834.50	-.021	.984
%_T_NSP	2801.00	-.145	.884
%_T_OTR	2789.00	-.190	.849
%_T_NAP	2748.00	-.348	.728
%_T_NLA	2758.00	-.316	.752
%_S_PAB	2759.50	-.302	.763
%_S_II	2675.00	-.616	.538
%_S_RTQ	2745.00	-.354	.723
%_S_PRBR	2743.50	-.360	.719
%_S_NAB	2759.50	-.302	.763
%_S_DVL	2700.00	-.530	.596
%_S_DM	2581.50	-1.072	.284
%_S_NRBR	2546.50	-1.217	.224

For negative behavior, students in the intervention group showed more distracting voice (RIG2: M = 13.02%, SD = 14.48%; Comparison: M = 10.50%, SD = 10.40%) and distracting movement (RIG2: M = 4.46%, SD = 6.69%; Comparison: M = 3.17%, SD = 4.91%). The only negative mean difference in the negative behavior category is negatively responding to behavioral requests (RIG2: M = 4.22%, SD = 7.65%; Comparison: M = 4.68%, SD = 6.76%). The Mann-Whitney test did not find any statistical significance in these differences (Table 4.7).

Summary

Mann-Whitney test revealed statistically significant differences in teacher practices of restorative language (U (79, 71) = 2283.00, z = - 1.964, p = .049) and student negative responses to behavioral requests (U (79, 71) = 2282.50, z = -2.229, p = .026) between the first random intervention group and the comparison group.

The same analysis was performed using the second random intervention group and the same comparison group. However, the results could not be replicated. There were no significant differences between the second random intervention group and the comparison group.

Instead of outperforming the comparison group for the first time ($MD_{T\_RL\_1}$ = 5.4%, p = .049), teachers in the second intervention group used slightly less restorative language ($MD_{T\_RL\_2}$ = - .60%, p > .05) than teachers in the comparison group. Students in the second random intervention group consistently demonstrated less percentage of negative responses to behavioral requests ($MD_{S\_NRBR\_1}$ = -.91 %, p = .026; $MD_{S\_NRBR\_2}$ = -.46%, p> .05). However, the differences for the second analysis were not statistically significant.

The results for both analyses were inconsistent. In terms of mean differences (Table 4.8), only three measures maintained the same trend: negative student responses to behavioral requests

(S_NRBR), teacher non-specific praise (T_NSP), and teacher opportunities to respond (T_OTR). Teachers in random intervention groups consistently outperform teachers in the comparison group on non-specific praise ($MD_{T\_NSP\_1}$ = .79%, $MD_{T\_NSP\_2}$ = .44). However, teachers in the comparison classrooms consistently provided more instructional opportunities for students to respond ($MD_{T\_OTR\_1}$ = -3.7, $MD_{T\_OTR\_2}$ = -.46). Unfortunately, the Mann-Whitney test did not reveal any statistical significance in these mean differences.

Overall, there were no significant differences in teacher practices and student behavior between classrooms where teachers have been trained for restorative practices and classrooms where teachers have not been trained for restorative practices.

Although teachers in the first random intervention group demonstrated significantly more use of restorative language (%_T_RL) than teachers in the comparison group. This result was not able to be replicated using another random intervention group. Likewise, the statistically significant difference in student negative responses to behavioral requests (%_S_NRBR) could not be replicated in the second analysis.

On average, teachers in the intervention group practiced more non-specific behavior praise (%_T_NSP) than teachers in the comparison group. Also, students in the intervention group demonstrated less negative responses to behavioral requests (%_S_NRBR). However, teachers in the comparison group provided more opportunities for students to respond to instructional questions (%_T_OTR).

Table 4.8

Comparing Mean Differences between Two Random Intervention Groups and Comparison Group

Measure	*MD*_1 (n=150) (M_{TIG1}-M_{Comp})	*MD*_2 (n=151) (M_{TIG2}-M_{Comp})	Consistent *MD*[a]
%_T_PAP	2.5	-1.62	
%_T_RL	5.4*	-.60	
%_T_NSP	.79	.44	Y
%_T_OTR	-3.7	-.46	Y
%_T_NAP	-2.51	1.61	
%_T_NLA	-.66	1.03	
%_S_PAB	4.13	-3.35	
%_S_II	2.51	-2.27	
%_S_RTQ	.79	-.68	
%_S_PRBR	.84	-.40	
%_S_NAB	-4.13	3.35	
%_S_DVL	-1.4	2.52	
%_S_DM	-.82	1.29	
%_S_NRBR	-1.91*	-.46	Y

Note. * $p < .05$; **a.** Y = MD_1 and MD_2 show the same direction (positive or negative) of the correlation between the intervention and comparison groups.

Research Question 2 and Research Question 3

Question 2: Are there differences in teacher practices and student behavior among classrooms with teachers in different intervention condition levels (trained 3 years ago/ 2 years ago/ 1 year ago/ not trained)?

Question 3: Do teacher practices and student behavior change overtime differently between the four condition levels?

Repeated Measures Analysis of Variance

The sample for the repeated-measures ANOVA comprised 90 classrooms with repeated observations: 24 for conditional level 0, 27 for conditional level 1, 20 for condition level 2, and 19 for condition level 3 (Table 4.9). A priori power analysis recommended a sample size of 76, 19 for each condition level, for a statistical power of .95 for an effect size of .10 (Table 3.5). Hence, the available sample size is adequate for assessing the effects between and within the condition groups.

As restorative practices experience leveled by four different condition levels, the analysis of its effects on teacher practices and student behavior was conducted using 4 (group: condition level 0, condition level 1, condition level 2, condition level 3) by 2 (time: time 1, time 2) repeated measures ANOVA.

Table 4.9

N*umbers of Classrooms with Repeated Observations for Each Condition Level Group*

School[a]	Condition Level 0	Condition Level 1	Condition Level 2	Condition Level 3
1	1	9		
2	0	12		
3	3	6		
4	2		6	
5	1		9	
6	5		5	
7	4			6
8	2			7
9	6			6
Total	**n = 24**	**n = 27**	**n = 20**	**n = 19**

Note. a. Schools are not listed in a specific order. School numbers are not assigned school codes.

Shapiro-Wilk's test for residuals (Table 4.10) indicated that the residuals of many dependent variables by condition level were not normally distributed ($p < .05$). The decision was made to run the analysis with the violation. Box's M test (Table 4.11) suggested that the assumption of homogeneity of covariance matrices were met by nearly all variables, except for S_II and S_NRBR ($p < .001$). Sphericity assumption was not assessed since there were only two levels for the within factor, time.

Table 4.10

Descriptive Statistics for Variables and Test of Normality for Residuals

		M (*SD*)		*MD*[b]	Shapiro-Wilk (*p*)	
	Code[a]	Time 1	Time 2	T2-T1	Time 1	Time 2
%_T_PAP	0	78.08 (21.74)	84.00 (21.51)	5.92	.89 (.010)[c]	.76 (.000) [c]
	1	77.73 (23.69)	83.04 (21.35)	5.31	.86 (.002) [c]	.79 (.000) [c]
	2	81.10 (20.97)	81.10 (29.72)	0	.79 (.001) [c]	.69 (.000) [c]
	3	75.90 (23.03)	92.52 (12.08)	16.62	.88 (.020) [c]	.66 (.000) [c]
%_T_RL	0	16.12 (15.52)	18.49 (15.87)	2.37	.74 (.000) [c]	.91 (.033) [c]
	1	19.41 (16.03)	22.95 (13.26)	3.54	.93(.057)	.96 (.289)
	2	20.46 (21.53)	20.59 (16.42)	.13	.79 (.001) [c]	.91 (.067)
	3	20.73 (18.39)	18.84 (15.31)	-1.89	.90 (.054)	.86 (.008) [c]
%_T_NSP	0	21.88 (15.78)	23.26 (12.40)	1.38	.95 (.247)	.93 (.088)
	1	28.52 (17.81)	21.02 (12.65)	-7.5	.98 (.745)	.90 (.014) [c]
	2	19.27 (11.86)	21.28 (12.56)	2.01	.95 (.345)	.96 (.456)
	3	14.49 (15.31)	31.07 (14.74)	16.58	.86 (.008) [c]	.98 (.979)
%_T_OTR	0	40.08 (20.00)	42.24 (21.63)	2.16	.96 (.459)	.96 (.522)
	1	29.80 (22.06)	39.08 (21.62)	9.28	.92 (.036) [c]	.98 (.762)
	2	41.37 (25.24)	39.23 (24.75)	-2.14	.94 (.224)	.92 (.107)
	3	40.69 (18.28)	42.61 (22.04)	1.92	.95 (.332)	.93 (.207)
%_T_NAP	0	21.92 (21.74)	16.00 (21.51)	-5.92	.89 (.010) [c]	.76 (.000) [c]
	1	22.27 (23.69)	16.96 (21.35)	-5.31	.86 (.002) [c]	.79 (.000) [c]
	2	18.90 (20.97)	18.90 (29.72)	0	.79 (.001) [c]	.69 (.000) [c]
	3	24.10 (23.03)	7.48 (12.08)	-16.62	.88 (.020) [c]	.66 (.000) [c]
%_T_NLA	0	11.33 (15.03)	9.95 (15.40)	-1.38	.77 (.000) [c]	.68 (.000) [c]
	1	14.49 (19.68)	11.88 (15.65)	-2.61	.75 (.000) [c]	.78 (.000) [c]
	2	9.87 (13.69)	14.97 (23.57)	5.1	.75 (.000) [c]	.70 (.000) [c]
	3	11.58 (13.97)	5.46 (11.58)	-6.12	.82 (.002) [c]	.52 (.000) [c]
%_S_PAB	0	74.11 (19.73)	87.62 (12.22)	13.51	.93 (.080)	.89 (.011) [c]
	1	68.34 (24.51)	84.79 (17.04)	16.45	.92 (.044) [c]	.83 (.001) [c]
	2	77.36 (22.02)	89.72 (14.22)	12.36	.87 (.014) [c]	.69 (.000) [c]
	3	76.25 (23.49)	92.05 (9.03)	15.8	.87 (.013) [c]	.83 (.004) [c]

%_S_II	0	17.43 (11.76)	11.98 (11.55)	-5.45	.95 (.206)	.83 (.001) [c]
	1	16.23 (12.24)	14.50 (16.64)	-1.73	.92 (.037) [c]	.77(.000) [c]
	2	11.24 (13.30)	19.97 (22.38)	8.73	.76 (.000) [c]	.73 (.000) [c]
	3	13.95 (9.93)	8.42 (5.17)	-5.53	.86 (.011) [c]	.95 (.454)
%_S_RTQ	0	29.74 (15.13)	41.33 (22.61)	11.59	.94 (.175)	.97 (.730)
	1	24.30 (20.41)	32.69 (22.89)	8.39	.90 (.011) [c]	.94 (.102)
	2	35.36 (25.43)	38.81 (27.41)	3.45	.96 (.455)	.92 (.84)
	3	33.04 (20.11)	50.10 (29.10)	17.06	.97 (.676)	.92 (.124)
%_S_PRBR	0	26.95 (19.43)	34.31 (15.60)	7.36	.86 (.003) [c]	.97 (.544)
	1	27.81 (20.34)	37.60 (19.46)	9.79	.85 (.001) [c]	.98 (.724)
	2	30.76 (18.65)	30.94 (19.67)	.18	.88 (.017) [c]	.94 (.243)
	3	29.25 (16.85)	33.53 (22.69)	4.28	.93 (.168)	.92 (.122)
%_S_NAB	0	25.89 (19.73)	12.38 (12.22)	-13.51	.93 (.080)	.89 (.011) [c]
	1	31.66 (24.51)	15.21 (17.04)	-16.45	.92 (.044) [c]	.83 (.001) [c]
	2	22.64 (22.02)	10.28 (14.22)	-12.36	.87 (.014) [c]	.69 (.000) [c]
	3	23.75 (23.49)	7.96 (9.03)	-15.79	.87 (.013) [c]	.83 (.004) [c]
%_S_DVL	0	15.94 (12.65)	7.04 (6.75)	-8.9	.87 (.006) [c]	.88 (.007) [c]
	1	19.49 (14.02)	8.67 (9.31)	10.82	.94 (.101)	.84 (.001) [c]
	2	15.97 (18.27)	6.87 (7.09)	-9.1	.80 (.001) [c]	.87(.013) [c]
	3	12.43 (14.56)	4.73 (7.33)	-7.7	.82 (.003) [c]	.71 (.000) [c]
%_S_DM	0	3.45 (5.15)	2.52 (3.81)	-.93	.70 (.000) [c]	.71 (.000) [c]
	1	6.61 (8.03)	2.22 (3.86)	-4.39	.80 (.000) [c]	.65 (.000) [c]
	2	2.63 (4.66)	2.70 (6.96)	.07	.64 (.000) [c]	.44 (.000) [c]
	3	3.89 (5.87)	2.32 (3.43)	-1.57	.70 (.000) [c]	.71 (.000) [c]
%_S_NRBR	0	6.50 (7.29)	2.82 (5.08)	-3.68	.82 (.001) [c]	.62 (.000) [c]
	1	5.56 (9.56)	4.33 (6.71)	-1.23	.66 (.000) [c]	.70 (.000) [c]
	2	4.05 (5.37)	.71 (2.19)	-3.34	.78 (.000) [c]	.36 (.000) [c]
	3	7.44 (9.12)	.91 (1.66)	-6.53	.81 (.002) [c]	.62 (.000) [c]

Note. **a.** Code: Condition Level 0 (0): Not Trained; Condition Level 1 (1): Trained one year ago; Condition Level 2 (2): Trained 2 years ago; Condition Level 3 (3): Trained 3 years ago; **b**. MD = $Mean_{T2}$ - $Mean_{T1}$; **c.** The assumption of normality is not met assessed by Shapiro-Wilk test ($p < .05$).

Table 4.11

Repeated Measures ANOVA and Tests of Equality of Covariance Matrices, Multivariate

Measure	Code[a]	Mean Square	*df*[b]	*F*	*p*	Partial η^2	Box[c] *p*
%_T_PAP	CL	121.019	3, 86	.189	.903	.007	
	T	2137.703	1, 86	6.072	**.016***	.066	.056
	CL*T	472.322	3, 86	1.342	.266	.045	
%_T_RL	CL	139.935	3, 86	.401	.753	.014	
	T	47.411	1, 86	.243	.624	.003	.457
	CL*T	63.897	3, 86	.327	.806	.011	
%_T_NSP	CL	155.357	3, 86	.641	.591	.022	
	T	429.250	1, 86	2.522	.116	.028	.417
	CL*T	1081.126	3, 86	6.352	**.001***	.181	
%_T_OTR	CL	560.200	3, 86	1.085	.360	.036	
	T	347.303	1, 86	.769	.383	.009	.878
	CL*T	270.347	3, 86	.599	.618	.020	
%_T_NAP	CL	121.026	3, 86	.189	.903	.007	
	T	2137.646	1, 86	6.071	**.016***	.066	.056
	CL*T	472.327	3, 86	1.342	.266	.045	
%_T_NLA	CL	185.445	3, 86	.527	.665	.018	
	T	69.262	1, 86	.353	.554	.004	.037
	CL*T	218.157	3, 86	1.112	.349	.037	
%_S_PAB	CL	564.660	3, 86	1.292	.282	.043	
	T	9314.866	1, 86	35.718	**.000***	.293	.130
	CL*T	41.426	3, 86	.159	.924	.006	
%_S_II[2]	CL	168.305	3, 86	.763	.518	.026	
	T	43.734	1, 86	.278	.599	.003	.000[c]
	CL*T	461.318	3, 86	2.935	**.038***	.093	
%_S_RTQ	CL	1370.196	3, 86	2.203	.093	.071	
	T	4518.447	1, 86	10.478	**.002***	.109	.512
	CL*T	322.661	3, 86	.748	.526	.025	
%_S_PRBR	CL	43.946	3, 86	.106	.956	.004	
	T	1289.359	1, 86	4.018	**.048***	.045	.284
	CL*T	194.117	3, 86	.605	.614	.021	
%_S_NAB	CL	564.695	3, 86	1.292	.282	.043	
	T	9314.697	1, 86	35.718	**.000***	.293	.131
	CL*T	41.422	3, 86	.159	.924	.006	
%_S_DVL	CL	226.981	3, 86	1.267	.291	.042	
	T	3678.227	1, 86	36.111	**.000***	.296	.552

	CL*T	19.373	3, 86	.190	.903	.007	
%_S_DM	CL	29.503	3, 86	.913	.438	.031	
	T	128.326	1, 86	4.605	**.035***	.051	.004
	CL*T	45.005	3, 86	1.615	.192	.053	
%_S_NRBR	CL	57.484	3, 86	1.076	.364	.036	
	T	601.924	1, 86	17.533	**.000***	.169	.000[c]
	CL*T	52.351	3, 86	1.525	.214	.051	

Note. **a.** CL = Condition Level, T = Time, CL*T = ConditionLevel*Time; **b.** df: degree of freedom, error degree of freedom; **c.** Box's test of equality of covariance matrices, $p < .001$ indicates the assumption of homogeneity of covariance matrices is not met.

The analysis did not detect any significant main effects of condition level on intended measures. The mean differences in teacher practices and student behavior between the four condition level groups were not statistically significant (Table 4.11).

Statistically significant main effects of time were found in overall teacher positive and negative affect practices ($F (1, 86) = 6.072$, $p = .016$, $\eta^2 = .066$). On average, teachers' overall positive practices increased 7% over time (Time1: M = 78.20, SE = 2.39%; Time2: M = 85.17%, SE = 2.35%; $p = .016$).

Furthermore, time had significant effects on all measures of student behavior, except for student initiation of interaction. Over the time, students substantially increased positive affect behavior ($F (1, 86) = 35.72$, $p < .001$, $\eta^2 = .293$), including responding to instructional questions ($F (1, 86) = 10.48$, $p = .002$, $\eta^2 = .109$), and positive responses to behavioral requests ($F (1, 86) = 4.02$, $p = .048$, $\eta^2 = .045$). Student negative affect behavior ($F (1, 86) = 35.72$, $p < .001$, $\eta^2 = .293$) also reduced significantly in the second observation, including distracting voice level ($F (1, 86) = 36.11$, $p < .001$, $\eta^2 = .296$), distracting movement ($F (1, 86) = 4.61$, $p = .035$, $\eta^2 = .051$), and negative responses to behavior request ($F (1, 86) = 17.53$, $p < .001$, $\eta^2 = .169$).

In addition, there were statistically significant interaction effects of time and condition level on teacher non-specific praise (%_T_NSP: F (3, 86) = 6.35, p =.001, η^2 = .181) and student initiation of interaction (%_S_II: F (3, 86) = 2.94, p =.038, η^2 = .093). There were no main effects of either condition level or time on the two dependent variables. It suggested that the effects of condition level on teacher use of non-specific praise and student initiation of interactions depended on time.

A crossover interaction of condition level and time indicated that the means of the variables crossed over each other in various situations. Teachers with 3-year experience of restorative practices demonstrated significant growth of 17% (Time1: M = 14.49%, SD = 15.31%; Time2 M = 31.07%, SD = 14.74%; $p < .05$) in non-specific praise between two observations. Teachers with 2-year experience (Time1: M = 19.27%, SD = 11.86%; Time2: M = 21.28%, SD = 12.56%) and teachers without any formal experience (Time1: M = 21.88%, SD = 15.78%; Time2: M = 23.26%, SD = 12.40%) of restorative practices also showed small increases in using non-specific praise over the time. On the contrary, teachers with 1-year restorative experience used 8% less non-specific practice (Time1: M = 28.52%, SD = 17.81%; Time2: M = 21.02%; SD = 12.65%) in the second observation (Figure 4.2).

A follow-up one-way ANOVA and post hoc test suggested that the mean difference (MD) of T_NSP over time differed significantly between four condition level groups (Table 4.12). The change in condition level group 3 was significantly higher with a 15% increase over the change in group 0 ($p < .05$), and 24% higher than the change of condition group 1 ($p < .001$).

Figure 4.2

Means of Teacher Non-Specific Praise (%_T_NSP) by Condition Level and Time

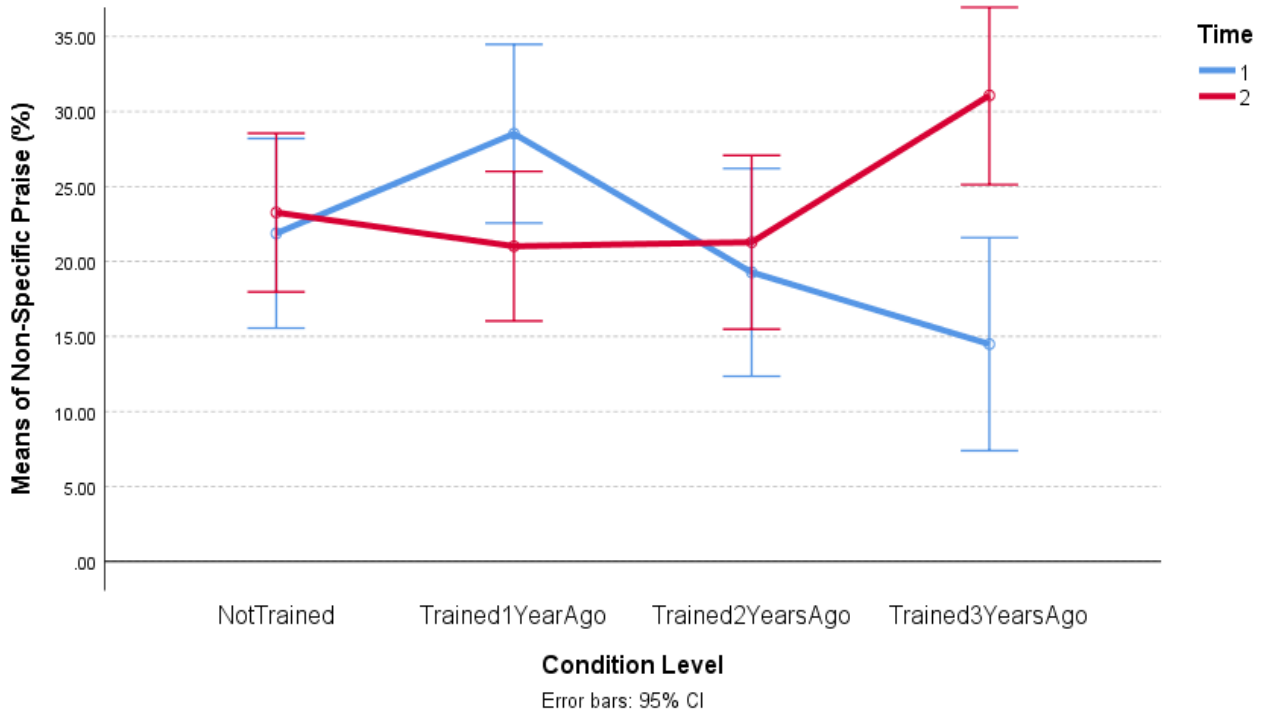


Table 4.12

Tukey HSD Test of One-Way ANOVA[a] Assessing Differences of Time Mean Difference (%_T_NSP) between 4 Condition Level Groups.

	Condition Level 0		Condition Level 1		Condition Level 2	
	MD	*p*	*MD*	*p*	*MD*	*p*
1	-8.89	.321				
2	.62	.999	5.44	.306		
3	15.20*	.043	24.08**	.000	14.57	.073

Note. *p < .05; **p<.001; **a.** DV: Mean difference of %_T_NSP between two times, IV: condition level.

A crossover interaction of condition level and time was also found in student initiation of interaction. None of the groups demonstrated significant changes in student behavior between two observation times (Figure 4.3). However, a follow-up one-way ANOVA and post hoc test (Table 4.13) revealed that the mean difference of %_S_II between times differed significantly between the four condition groups. The change in group 2 was 14% higher than the change in group 0 ($p<.05$).

Figure 4.3

Means of Student Initiation of Interaction (%_S_II) by Condition Level and Time

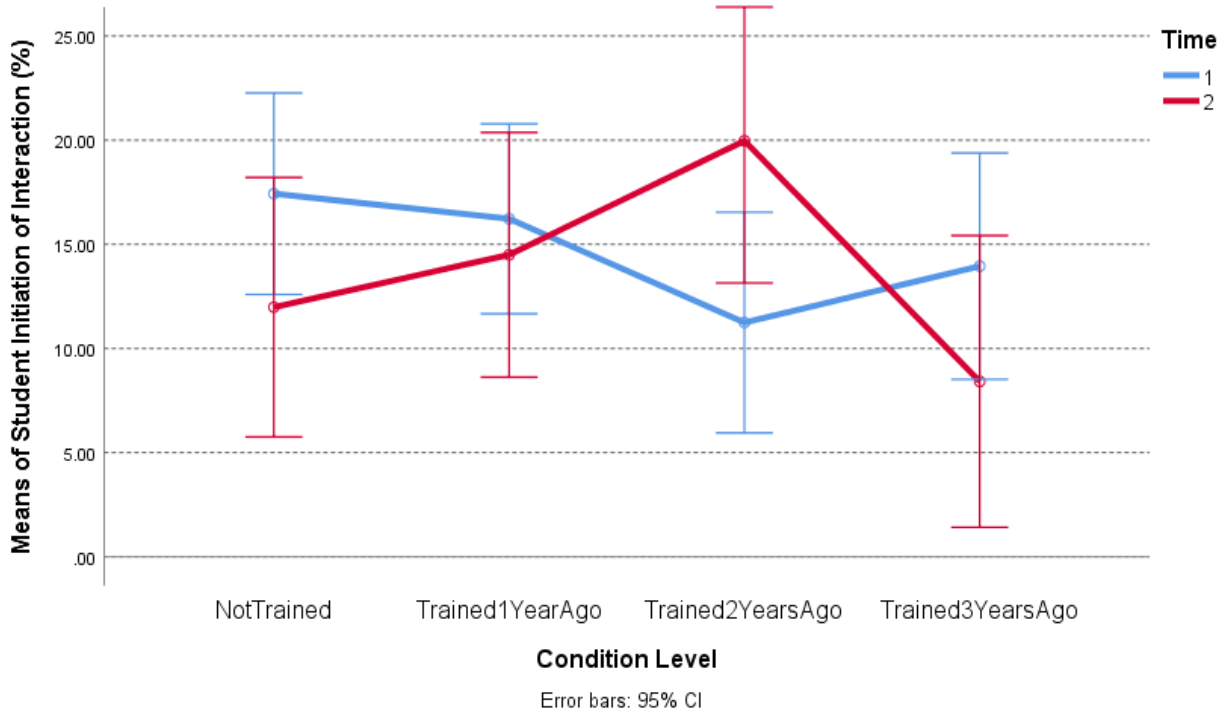


Table 4.13

Tukey HSD Test of One-Way ANOVA[a] Assessing Differences of Time Mean Difference (%_S_II) between 4 Condition Level Groups.

	Condition Level 0		Condition Level 1		Condition Level 2	
	MD	*p*	*MD*	*p*	*MD*	*p*
1	3.72	.878				
2	14.17*	.048	10.46	.197		
3	-.09	1.000	-3.81	.890	-14.26	.065

Note. **p < .001; **a.** DV: Mean difference of %_S_II between two times, IV: condition level.

Summary

Overall, there were no statistically significant differences in teacher practices and student behavior between condition levels. The analysis did not find any significant main effects of condition-level based on the current sample (n=90). Inferred by the current data, different dosages of the intervention, restorative practices experience, do not make significant differences in teacher practices and student behavior.

However, there were statistically significant differences in teacher practices and student behavior within the subjects between two observations. The main effects of time were found in most student behavior measures, overall teacher positive and negative practices.

Besides, there were significant differences in teacher non-specific praise (%_T_NSP) and student initiation of interaction (%_S_II) between condition levels over time. Repeated measures ANOVA revealed crossover interaction effects in these two measures. Teachers with 3-year experience of restorative practices (condition level 3) made significant progress in using non-specific praise than teachers with 1-year experience (condition level 1) and teachers without

formal restorative experience (condition level 0) over time. Also, students in condition level 2 demonstrated significant growth in initiating interaction with their teachers between two observations compared to students in condition level 0.

Research Question 4

Research Question 4: What are the relationships between teacher practices and student behavior?

Correlation Analysis

Pearson correlations examined the relationships between variables of teacher practices and student behavior. Table 4.14 provides descriptive statistics for each variable prior to the correlation analysis. The mean for overall teacher positive affect practices (%_T_PAP) was 82.10% (SD = 21.50%), and the mean for overall student positive affect behavior (%_S_PAB) was 81.90% (SD = 19.09%).

The analysis in Table 4.15 included the entire sample (n=230). Dancey and Reidy's (2017) criteria were used to interpret the Pearson r values. Overall, the relationship between teacher positive practices (%_T_PAP) and student positive behavior (%_S_PAB) was positive, moderate in strength and statistically significant ($r (228) = .571$, $p < .001$).

Positive correlation between teacher use of restorative language (%_T_RL) and student positive behavior (%_S_PAB) was significant, but very small ($r (228) = .183$, $p = .005$). Teacher use of opportunities to respond (%_T_OTR) also indicated small positive correlation with student positive behavior ($r (228) = .342$, $p < .001$). On the contrary, teacher use of non-specific praise did not appear a significant relationship with student behavior ($r (228) = .110$, $p = .097$).

Additionally, all measures of teacher negative practices (%_T_NAP, %_T_NLA, %_T_NAP) and student negative behavior (%_S_NAB, %_S_DVL, %_S_DM, %_S_NRBR) were significantly and positively correlated with moderate strength (Table 4.15).

Table 4.14

Descriptive Statistics for Variables prior to Pearson Correlation Analysis

Measure	*M*	*SD*	Minimum	Maximum	95% CI for Mean	
%_T_PAP	82.10	21.50	9.52	100.00	79.31	84.88
%_T_RL	19.13	16.18	.00	85.71	17.03	21.23
%_T_NSP	22.87	14.71	.00	66.67	20.96	24.78
%_T_OTR	40.10	22.50	.00	94.74	37.18	43.03
%_T_NAP	17.91	21.46	.00	90.48	15.12	20.69
%_T_NLA	10.97	15.63	.00	78.95	8.94	13.00
%_T_NPA	6.94	9.83	.00	44.44	5.66	8.21
%_S_PAB	81.90	19.09	12.50	100.00	79.42	84.38
%_S_II	14.05	13.67	.00	100.00	12.28	15.83
%_S_RTQ	35.98	23.44	.00	89.47	32.93	39.02
%_S_PRBR	31.88	18.60	.00	91.67	29.46	34.29
%_S_NAB	18.10	19.09	.00	87.50	15.62	20.58
%_S_DVL	10.90	12.08	.00	75.00	9.33	12.47
%_S_DM	3.34	5.52	.00	30.77	2.62	4.05
%_S_NRBR	3.87	6.64	.00	35.00	3.00	4.73

Table 4.15

Variable Correlations

	T_PAP	T_RL	T_NSP	T_OTR	**T_NAP**	T_NLA	S_PAB	S_II	S_RTQ	SPRBR	S_NAB	S_DVL	S_DM
T_RL	.361**	-											
T_NSP	.257**	-.041	-										
T_OTR	.526**	-.348**	-.379**	-									
T_NAP	-1.00**	-.361**	-.257**	-.526**	-								
T_NLA	-.907**	-.268**	-.216**	-.531**	.907**	-							
S_PAB	.571**	.183**	.110	.342**	-.571**	-.472**	-						
S_II	.131*	.381**	.055	-.185**	-.131*	-.113	.197**	-					
S_RTQ	.429**	-.217**	-.164*	.672**	-.429**	-.384**	.497**	-.323**	-				
S_PRBR	-.051	.181**	.279**	-.360**	.051	.082	.255**	-.126	-.512**	-			
S_NAB	-.571**	-.183**	-.110	-.342**	.571**	.472**	-1.00**	-.197**	-.497**	-.255**	-		
S_DVL	-.416**	-.095	-.106	-.259**	.416**	.328**	-.889**	-.105	-.459**	-.257**	.889**	-	
S_DM	-.450**	-.197**	-.006	-.283**	.450**	.385**	-.648**	-.120	-.346**	-.140*	.648**	.375**	-
S_NRBR	-.512**	-.188**	-.118	-.275**	.512**	.440**	-.717**	-.273**	-.307**	-.148*	.717**	.424**	.347**

Note. * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

Summary

There was a significantly positive and moderate relationship between the overall teacher positive practices and student positive behavior ($r(228) = .571$, $p < .001$), the same correlation between negative teacher practices and negative student behavior.

Among all the teacher positive practices, opportunities to respond (%_T_OTR) showed most positive correlation with student positive behavior ($r(228) = .342$, $p < .001$). Restorative language had a significantly positive, but very small, association with overall student positive behavior ($r(228) = .183$, $p = .005$). Teacher use of non-specific praise (%_T_NSP) had trivial relationship with student behavior ($r(228) = .110$, $p = .097$).

All variables of teacher negative affect practices and student negative affect behavior were significantly and positively correlated with moderate strength. The analysis indicated that negative teacher practices, either linguistic or physical, were profoundly related to student distractive and destructive behavior, either verbally or physically.

CHAPTER FIVE: DISCUSSION

Introduction

Despite the growing evidence on the ineffectiveness of exclusionary discipline practices, zero-tolerance approaches are still prevalent in American schools (Fabelo et al., 2011; Flannery, 2015; USDOE, 2014; Bradshaw et al., 2010). Research shows that the vast majority of out-of-school suspensions were caused by minor, nonviolent disruptions (Boccanfuso & Kuhfeld, 2011). The long term effects of inequitable discipline approaches result in negative youth development, achievement gap, escalated behaviors, and pipeline to the juvenile justice system (Flannery, 2015).

Classroom referrals exceedingly depend on the subjective judgment of teachers (Bradshaw et al., 2010). Educators must explore alternative approaches to enhance their practices and broaden their perceptions of appropriate behavior. Teachers should re-think ways to engage students in classrooms to learn and narrow the achievement gap.

Restorative practices develop relationships and communities through teaching preventative strategies and promoting positive behavior, which provides students with equitable access to participation in quality learning opportunities. The bonds and connections developed within school and classroom communities help students build empathy, self-efficacy, and confidence (Rodman, 2007). These practices transform adult-student interactions and build a solid foundation for positive classroom and school climate. Ultimately, inclusive environments and positive relationships increase students' possibilities to reach higher academic achievement and personal performance.

In the past two decades, restorative practices have grown in popularity among national and international schools. However, the adoption of the program has outpaced the research related to these practices. There is a growing demand for well-designed empirical and quasi-empirical studies in the field to meet the ESSA requirements of higher-level evidence (USDOE, 2016). Furthermore, the research area regarding the effects of restorative practices on teacher practices in elementary school settings is currently understudied. Also, quantitative observation is rarely used in exploring this topic. This study extends the knowledge of the effects of restorative practices at the practice level using observational data in elementary classrooms.

The purpose of this study is to investigate the impact of restorative practices on teacher practices and student behavior in elementary classrooms through direct observations. This quasi-experimental designed study analyzed the secondary data from a stratified school-randomized evaluation to examine if teachers with more restorative practices experience would show more use of positive practices and make more progress in positive practices over time. The study assessed the progressive differences based on the different dosages or levels of the intervention, defined as the overall experience of restorative practices.

Findings

Effects of Restorative Practices Experience

Research question 1. It was hypothesized that teachers who received formal training of restorative practices would manifest more positive teacher-student interactions and less negative interactions than teachers who have never been formally introduced to restorative practices. However, the overall analysis of the current data did not entirely support this hypothesis.

In conclusion, the inconsistent results suggested that there were no significant differences in teacher practices and student behavior between the intervention and comparison groups. The first analysis revealed that there were significantly more teacher use of restorative language (%_T_RL: Mann-Whitney U (79, 71) = 2283.00, z = - 1.964, p = .049) and less negative student responses to behavioral requests (%_S_NRBR: Mann-Whitney U (79, 71) = 2282.50, z = -2.229, p = .026) in intervention classrooms than comparison classrooms. On average, the first random intervention group demonstrated 3% more positive teacher practices and 4% more positive student behavior than the comparison group. However, these results could not be replicated with another randomly selected intervention group. The replicate tests suggested that the second random intervention group demonstrated about 2% less positive teacher practices and 3% less positive student behavior compared to the comparison group. No significant differences were detected in the second analysis.

Research question 2. In the second question, we continued investigating the differences between groups with different levels of restorative practices experience. Theoretically, there should be progressive differences in teacher-student interaction among groups with different dosages of the intervention. We hypothesized that, as the condition level increases, positive teacher practices and student behavior would increase, and negative practices and behavior would decrease.

Results did not suggest any significant effects of restorative practices experience on teacher practices and student behavior. The analysis did not detect any significant main effects of condition-level in any variables. The results further confirmed the findings in the first question.

Contributions. Most literature in the field has focused on the effects of the whole-school approach in secondary schools. Acosta and her colleagues (2019) also did not find significant

impact of restorative practices on middle school student behavior. Two comprehensive evaluations from the U.K. (Bitel, 2005; Kane et al., 2007) used a mixed method of qualitative and quantitative approaches to assess student outcomes and implementation processes in primary and secondary schools. Survey data did not show statistically significant effects of restorative practices on student attitudes and school improvement. Kane et al. (2007) suggested that primary schools demonstrated a more straightforward change process compared to large-sized secondary schools. A quasi-experimental evaluation conducted in an alternative program setting indicated a significant impact of implementing restorative practices on the overall school environment based on self-reported and school administrative data (McCold, 2008).

To date, no published studies have specifically investigated teacher practices and student behavior at the classroom level using direct observations. Despite the insignificant results, the current findings contribute to foundational knowledge of restorative practices on teacher practices and student behavior at the classroom level and using objective observations. The research also fills a gap in research on elementary schools.

Explanations. There are several possible explanations for the results. First, teacher commitment to implement restorative practices in classrooms is critical to the outcomes but not assessed. Trained teachers may not have been actively implementing restorative practices over the years or practicing at recommended levels. School readiness, including administrative support and staff buy-in, was assessed and considered as one of the primary criteria for the selection process. The sampling process was able to control the teacher turnover effect. Therefore, the study assumed the adequate commitment of practitioners. However, in reality, it is not surprising that teachers and administrators would have been overwhelmed with multiple

required initiatives in the district, which could negatively influence people's commitment to implementation.

Second, we do not know whether teachers in the intervention groups received monthly coaching, either individually or in groups. Although the condition levels were distinguished based on initial training years, the study did not address the training effects beyond the initial intervention. Joyce and Showers (2002) stated that demonstration in training had little effect on changing teacher practices in classrooms. According to their meta-analysis, coaching in classrooms can profoundly increase teachers' knowledge, skill, and use. This implication also means that the intervention would not have substantial impacts on teacher practices if they do not receive adequate individual coaching.

Effects of Time and Restorative Practices Experience

Research question 3. *S*ignificant interaction effects were not found on the overall teacher practices and student behavior. However, the analysis revealed statistically significant interaction effects of time and condition level on two variables: teacher non-specific praise ($p = .001$, $\eta 2 = .181$) and student initiation of interaction ($p = .038$, $\eta 2 = .093$). For the third research question, there are significant differences in teacher use of non-specific praise and student initiation of interaction among four condition level groups between two observations.

Teacher use of non-specific praise changed significantly over time with different dosages of the intervention. Specifically, the progress made by teachers with 3-year experience of restorative practices (MD = 16.58%) were profoundly higher than teachers with 2-year experience (MD = 2.01%, $p > .05$), teachers with 1-year experience (MD = -7.5%, $p < .001$), and teachers with little experience (MD = 1.38%, $p = .043$). The findings suggested that teachers

with more experience with restorative practices showed more intention in using basic positive practices.

The interpretation for student initiation of interaction is less straightforward. The changes made by students in condition level 2 classrooms (MD = 8.73%) were substantially higher than students in condition level 0 (MD = - 5.45%, $p = .048$), condition level 1 (MD = -1.73%, $p > .05$), and condition level 3 (MD = -5.53%, $p > .05$). However, it is also notable that this particular positive student behavior was moderately reduced over time in three of the four condition levels.

Contributions. Little research has investigated the interaction effect of restorative practices and time. McCold and Wachtle (2002) found significant effects of restorative practices on improving student attitude and behavior (e.g., referrals, delinquency, graduation rates), which were positively associated with the time students had been involved in restorative practices. Current findings initiate and invite conversations about the overtime impact of restorative practices on teacher-student interaction in classrooms.

Explanations. One potential explanation of the increasing use of non-specific praise with increasing experience could be the rising awareness of trained practitioners and the adaptability of this particular practice. Teachers with formal training of restorative practices may be well aware that they should use more positive practices. However, they may lack essential skills to utilize more sophisticated positive practices like restorative language (affective statements and specific praise) or questioning techniques (instructional questions for common content and circles). Non-specific praise is a low-hanging fruit compared to other positive practices. It requires less knowledge and techniques to master. This finding suggests that the length of

restorative practices experience has a meaningful impact on teacher awareness or attitude of positive practices, but not much on skill levels.

The explanation of student initiation of interaction may be more related to the measurement rather than the intervention. This category measures the percentage of teacher-student interaction initiated by students, an indicator of a student-centered learning environment. The significant surge in condition level 2 could be a result of an outlier. Exploration of individual case data showed that one classroom had a substantial increase in the category between two observations. It could be that the learning activity in the second observation generated very few interactions that all happened to be initiated by students.

An explanation for the lack of significant effects of intervention and time in most variables could be the short period time between two observations. This time frame may not be sufficient to observe substantial changes in practices.

Relationships between Teacher Practices and Student Behavior

Research question 4. Results showed that the overall positive teacher practices and the overall positive student behavior were moderately and positively associated ($r(228) = .571$, $p < .001$), the same relationship between negative teacher practices and negative student behavior.

Among all the observed positive teacher practices, opportunities to respond (OTR) appeared to be the strongest correlation with student positive behavior ($r (228) = .342$, $p < .001$). Correlation between restorative language and student positive behavior is smaller and also positive ($r (228) = .183$, $p < .001$). Teacher use of non-specific praise did not significantly relate to student behavior.

Contributions. Opportunities to respond (OTR) as an instructional strategy have been well studied, in particular, in special education settings. A large body of literature suggested positive effects of OTR on student academic performance and positive behavior, such as on-task behavior and learning rates (Carnine, 1976; Skinner et al., 1994; West & Sloane, 1986; Sutherland et al., 2001). Gunter et al. (1993) pointed out that increasing correct responses would provide more opportunities for teacher praises and positive interactions between teachers and students. The current findings confirm the significant positive relationship between OTR and student positive behavior and extend the previous findings to general education classrooms.

Stichter et al. (2009) studied the relationships between different components of OTR and classroom management procedures in elementary classrooms. They suggested that teacher instructional talk was significantly and negatively associated with student negative verbalizations. This study also found significant and positive relationships between teacher verbal negatives and student verbal outbursts and overall behavior disruptions. Consistently, the findings of the current study also found significant and negative correlation between OTR and student distractive voice level ($r(228) = -.259$, $p < .001$), and significantly positive correlation between negative teacher language and student distractive voice ($r(228) = .328$, $p < .001$).

Furthermore, Cameron and Pierce (1994) reviewed nearly 100 empirical studies about teacher praise published from 1971 to 1991. They concluded that praise must be behavior specific in order to affect the behavior effectively. Their findings echoed Brophy's (1981) assertion. The current results provide additional support to the previous findings. Our analysis also suggests significant relationships of student behavior with restorative language and insignificant correlation with non-specific praise.

The current findings on relationships between teacher practices and student behavior extend the previous work in two aspects. First, the consistent findings add additional evidence to the growing literature of effective instructional strategies and classroom management practices. Moreover, the study initiates an exploration of restorative language in relationship to student behavioral outcomes.

Restorative language, in the current study, primarily includes affective statements and behavior-specific praise. A growing literature has focused on examining restorative practices as a whole-school program and its responsive component, such as conferencing. To date, no published study has examined affective statements individually. Its effects or relationships with other variables remain unknown. Although the current study was not able to provide specific insight directly related to affective statements, it is imperative to generate a conversation about its effectiveness.

Explanations. Affective statements are named as one of the fundamental practices of restorative practices. Instead of hosting an individual category, the study combined this practice and behavior-specific praise into a single item, restorative language. The decision was grounded in three reasons. First, affective statements were rarely observed in the district evaluations in the past two years. Unlike affective statements that are more emotion-based, behavior-specific praise is more observable and natural for practitioners to master. Besides, both affective statements and behavior-specific praise aim at generating positive affect of the participants through specific language. Also, the effectiveness of behavior-specific praise is well-established in the literature.

There are a couple of aspects to explain the stronger correlation of student behavior with OTR than with restorative language. First, OTR aims to engage students in instructional participation, which effectively reduces disruptions. As aforementioned, OTR is a well-studied

evidence-based practice. On another side, restorative language focuses on the acknowledgment of individual feelings and behavior. In terms of measurement, there are many more occurrences of OTR observed in classrooms than restorative language. The standard deviation of restorative language was relatively large to the mean. The distribution of OTR is much more normal in comparison. OTR was observed in a variety of contents and forms, including restorative circles. Teachers appeared to be more comfortable to ask content-related questions rather than express their feelings or even praise students. This outcome may relate to the lack of coaching and intentional practices.

Limitations

It is critical to acknowledge the limitations while interpreting the study results. These limitations will guide researchers and practitioners to utilize the findings appropriately and to design future research effectively.

Limitations of Implementation Fidelity

The first limitation of the study is the absence of fidelity data. Studies have emphasized the imperative role of implementation fidelity in examining restorative practices (Acosta et al., 2019; Gregory et al., 2016). For this study, the fidelity data at the classroom level were not collected by the district, such as coaching sessions, average time on using restorative practices, and perceived administrative support by teachers. The impact of these additional interventions was not addressed, for instance, coaching, consultation, or on-going training. Hence, we cannot determine if, or how, fidelity factors relate to the findings.

Moreover, the study did not have relevant data to quantify and separate the impact of other district initiatives or similar programs that may be working to either compete or

complement restorative practices. All sample schools are currently implementing Positive Behavior Interventions and Supports (PBIS) as required by the district. Many of these schools have also adopted various social-emotional programs, other school improvement models, or classroom management systems, such as Leader in Me and CHAMPS. Limited knowledge about the impact of these competing programs on teacher practices and student behavior brings cautions to interpreting the current findings.

Limitations of Measures

Another limitation area in the current study is the measures included in the instrument. First, as aforementioned, restorative language is a combined category rather than separated scores for affective statements and behavior-specific praise. Unlike affective statements, behavior-specific is not a signature term for restorative practices. This measure limited the possibility to explicitly assess the impact of affective statements to build a foundation for future research.

Second, OTR is also not a specified element for restorative practices. Questioning is a common measure of interaction. OTR is a technique broadly promoted for teaching since there is strong empirical evidence supporting the effectiveness. The study did not use well-known components as measures, such as affective statements, circle, small impromptu conversations, or conferences (IIRP, 2010). It is because these components are rarely seen in short observations. However, the essence of those practices was observed and recorded in break-down elements. For observed circles and small impromptu conversations, teacher practices and student behavior were specified and recorded for the particular items, such as OTR, specific and non-specific praise, affective statements, and positive or negative physics and language. Although these measures limited the interpretation and comparison of the findings with previous and future research, they

also shone a light on how to integrate restorative practices with other evidence-based approaches to maximize student outcomes.

Limitations of Study Design

The method design of the study also presents some limitations. First, the control condition or the comparison group in the study has been exposed to the intervention. The idea was that if the hypothesized changes in practices and behavior show a similar trend with the progressive dosages of the intervention, it would help us draw a causal conclusion about the effects of restorative practices experiences on teacher practices and student behavior. Although it was clear that teachers in the comparison group had never had official training on restorative practices, the contamination effect was unknown since they also worked at one of the nine restorative schools. Also, the proposed random selection was not performed due to limited samples at each level. As a result, other contextual factors could not be effectively controlled. This limitation brings caution to interpret the current findings. It could explain the insignificant results between the intervention and comparison groups.

Second, the initial training year alone is not sufficient to define the dosage of the intervention. Due to the lack of fidelity data at the practice level, training time, as the only well-defined and objective measure, was used to classify condition levels. Despite the convenience of grouping, it does not reflect the actual amount of restorative experiences.

Third, the time between observations and numbers of repeated observations limited the accuracy of the results. There were two months between the two observations. It might not be long enough to detect changes in practices. Also, the district initially planned for three observations. Due to the Covid-19 pandemic, only two observation rounds were conducted. This

sudden change affected the sample size for the independent test and the robustness of the findings.

Implications

The current findings present important implications for researchers and educators with the acknowledgment of the limitations. First, the findings suggest that restorative practices experience significantly increases teachers' awareness of promoting a positive classroom environment over time. As the dosage of the intervention increases, teachers demonstrate increasing use of non-specific praise. However, the second implication of the findings also reveals that awareness or non-specific praise alone is not enough to change student behavior. Moreover, the third implication from the correlation analysis indicates that opportunities to respond and more specific restorative language, but not non-specific praise, are significantly associated with positive student behavior.

Implications for Researchers

Researchers and evaluators have investigated the impact of restorative practices on participants' attitudes and behavior change based on survey data (Bitel, 2005; Gregory et al., 2016; Kane et al., 2007). The current implication based on objective observational data confirms the behavior changing process that awareness and attitude change before behavior or practices change.

Many practices or interventions with strong evidence are fundamental with clearly-defined definitions and specific implementation procedures. Therefore, empirical evidence can be re-tested and accumulated through replications. The practices included in restorative practices are much more general in comparison. It makes this intervention or interventions challenging to

examine. It is difficult to determine which techniques (e.g., questioning, praising) or intervention forms (e.g., circles, conversations) contribute to the effects.

The adaptation of restorative practices appears broad variations from country to country, district to district, school to school, classroom to classroom, and practice to practice (Acosta et al., 2019; Bitel, 2005; Kane et al. 2007). Without unified implementation procedures, the assessment of its effects can be extremely challenging and misleading. The mixture of various practicing forms and techniques could be an alternative explanation for the lack of empirical evidence for restorative practices in the existing literature.

Implications for Educators

For educators, these implications recommend educators to focus on several efforts to improve their practices and student engagement. First, teachers are encouraged to integrate opportunities to respond into various forms of restorative practices, such as circles and small impromptu conversations. These opportunities produce positive affects among participants. They are fundamental building blocks for positive teacher-student interaction. More importantly, instructional opportunities for students to respond can come in diverse forms and shapes. Effective use of proven effective strategy can increase the chance for student success.

Second, non-specific praise is the transition point for specific restorative language. The findings also indicate that this transition is difficult and requires intentions. Two or three years of experience without intentional practices will not automatically grant this shift. In other words, purposeful coaching and progress monitoring are critical if educators aim to improve this area.

Extending from the first two points, coaching at the classroom level must be emphasized. Although fidelity data were not available, it is crystal clear that teacher practices will not

improve solely based on training sessions through lectures. Training at the whole-school or even large-group level will not produce much effect on behavior change. The bottom line is that practitioners must advance their practices through direct coaching and ongoing feedback from experts (Joyce & Showers, 2002).

Besides the positive practices, the findings also implicate that teachers' negative language and physical expressions are significantly and positively associated with student negative behavior. While focusing on positive practices, educators also need to recognize that negative practices take a substantial toll on negative teacher-student interaction. The challenge is that negative practices are often unconscious for practitioners. To improve this area, coaches are recommended to utilize the current instrument for data collection and guide conversations for continuous improvement. The observation tool can be useful for establishing intentionality and monitoring progress.

Finally, the study implicates the importance of developing implementation fidelity measures and data collection procedures at the school and practitioner levels. It is common to have school-level fidelity data. However, any intervention will not take effect until it reaches to the individual level. To help the matter, districts and schools should develop a logic model to explicitly identify the active components for the program and illustrate the causal path from implementation to positive student outcomes. The absence of a shared understanding among stakeholders may result in inappropriate adaption and ineffective use of resources.

Future Research

Future studies should provide additional data and analyses to deepen the understanding of the current findings. First, there should be more observations with more extended time in

between. For instance, future studies can collect observational data several times during a school year or even multiple years. Additional observational data will bring more robust findings.

Furthermore, implementation fidelity data should be collected and analyzed for future research. The mediating factors related to fidelity should be examined if the main effects of the intervention are detected. This process may involve an intentional design of data collection to track the fidelity data on coaching, administrative support, staff readiness, and commitment through surveys or interviews.

Study design should also be improved for future studies with additional resources. First, although the observation instrument is practical for practitioners, further development is required for better reliability and validity for research use. Future research with adequate funding should consider adapting observation intervals to produce better internal consistency, more normally distributed data, and more accurate results.

Moreover, future studies should consider a randomized control trial (RCT) design to investigate the current topic. It would be ideal for a longitudinal study with the baseline and ongoing observational data from randomly selected control and treatment schools and classrooms to illustrate the trends of teacher practices and student behavior along with the implementation. Two published studies using RCT have focused on the whole-school program (Acosta et al., 2019; Green et al., 2019). It will be insightful to understand the evolution of classroom practices and behavior with in-depth observations.

Additional recommendations for future research include the investigation of racial disproportionality and disparity in teacher practices. Gregory et al. (2016) have investigated restorative practices to address the racial discipline gap from the perspective of teacher-student

relationships using survey data. Future studies are encouraged to use observational data to explore the root of the equitable learning environment, teacher practices.

Final Remarks

Researchers and educators have reached a consensus on the priority of investing money, time, and energy into proven effective interventions and practices. I became an educator so that I can be part of the force for change. I am becoming a researcher so that I can force the change effectively. It is critical to know if an intervention is effective in improving student outcomes. It is also imperative to know if an intervention can be replicated among practitioners effectively and broadly. Increasing levels of resources and funding are being allocated to programs to improve overall climate and address urging discipline concerns. We are reaching a reflection point to ask ourselves whether we are moving towards our goals, either to change behavior, to close gaps, or to improve the climate. It is undeniable that spending top dollar on top-quality programs for all students is a matter of social equity.

References

Abramson, L. (2015), Restorative approaches in the next decade: The how may be as important as the what, *Restorative Practices in Action Journal: For School and Justice Practitioners*, New York State Permanent Judicial Commission on Justice for Children and the Center on Race, Crime and Justice of the John Jay College of Criminal Justice, CUNY.

Acosta, J., Chinman, M., Ebener, P., Malone, P. S., Phillips, A., & Wilks, A. (2019). Evaluation of a whole-school change intervention: Findings from a two-year cluster-randomized trial of the restorative practices intervention. *Journal of Youth and Adolescence, 48*, 876-890. doi: 10.1007/s10964-019-01013-2.

American Psychological Association. (n.d.). APA dictionary of psychology. Retrieved from https://dictionary.apa.org/

American Psychological Association Zero Tolerance Task Force. (2006). Are zero tolerance policies effective in the schools? An evidentiary review and recommendations. American Psychological Association.

Bitel, M. (2005). *National evaluation of the restorative justice in schools programme.* London: Youth Justice Board for England and Wales.

Blood, P. (2005). *The Australian context - restorative practices as a platform for cultural change in schools.* Paper presented at the XIV World Congress of Criminology, Philadelphia, PA.

Boccanfuso, C. and Kuhfeld M. (2011). *Multiple Responses, Promising Results: Evidence-Based, Nonpunitive Alternatives to Zero Tolerance* (Publication #2011-09). Washington, DC: Child Trends, citing Skiba, R. (2000). *Zero Tolerance, Zero Evidence: An Analysis*

of School Disciplinary Practice. Bloomington, IN: Education Policy Center Indiana University.

Bradshaw, C. P., Mitchell, M. M., O'Brennan, L. M., & Leaf, P. J. (2010). Multilevel exploration of factors contributing to the overrepresentation of black students in office disciplinary referrals. *Journal of Educational Psychology, 102*(2), 508–520. doi:10.1037/a0018450

Braithwaite, J. (1989). *Crime, shame and reintegration.* Cambridge: Cambridge University Press.

Brophy, J. E. (1981). Teacher praise: A functional analysis. *Review of Educational Research, 51*(1), 5–32.

Broden, M., Bruce, C., Mitchell, M. A., Carter, V., & Hall, R. V. (1970). Effects of teacher attention on attending behavior of two boys at adjacent desks. *Journal of Applied Behavior Analysis, 3*, 199-203.

Cameron, J., & Pierce, W. D. (1994). Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research, 64*(3), 363–423.

Camine, D. W. (1976). Effects of two teacher-presentation rates on off-task behavior, answering correctly, and participation. *Journal of Applied Behavior Analysis, 9*, 199-206.

Chafouleas, S. M., Riley-Tillman, T. C., & Christ, T. J. (2009). Direct Behavior Rating (DBR): An emerging method for assessing social behavior within a tiered intervention system. *Assessment for Effective Intervention, 34*, 195–200.

Colvin, G. S., Sugai, C. L. G., & Patching, W. (1993). Precorrection: An instructional approach for managing predictable problem behaviors. *Intervention in School and Clinic, 28*, 143–150.

Costello, B., Wacthtel, J., & Wachtel, T. (2009). The restorative practices handbook: Building a culture of community in schools. International Institute of Restorative Practices.

Dancey, C. P., & Reidy, J. (2017). *Statistics without maths for psychology* (7th ed.). New York: Person.

Darwin, C. R. (1859). *On the Origin of Species by Means of Natural Selection, or, the Preservation of Favored Races in the Struggle for Life.* London: John Murray.

Darwin, C. R. (1871). *The descent of man, and selection in relation to sex.* London: John Murray.

Darwin, C.R. (1872). *The expression of the emotions in man and animals.* London: John Murray.

Darwin, C. R. (1890). *The expression of the emotions in man and animals*. 2nd edition. Edited by Francis Darwin. London: John Murray.

Dharwadker, V. (2015). Emotion in motion: The Natyashastra, Darwin, and affect theory. *Modern Language Association of America, 130* (5), 1381-1404.

Dignan, J. (2002). *Reparation orders, in B. Williams (ed.) Reparation and victim-focused social work.* London: Jessica Kingsley Publishers.

Doumen, S., Verschueren, K., Buyse, E., Germeijs, V., Luyckx, K., & Soenens, B. (2008). Reciprocal relations between teacher-child conflict and aggressive behavior in kindergarten: A three-wave longitudinal study. *Journal of Clinical Child & Adolescent Psychology, 37,* 588–599.

Eyberg, S. M., Nelson, M. M., & Boggs, S. R. (2008). Evidence-based psychosocial treatment for children and adolescents with disruptive behavior. *Journal of Clinical Child and Adolescent Psychology, 37,* 215-237.

Eyberg, S. M, Nelson, M. M., Duke, M, & Boggs, S. R. (2005). *Manual for the Dyadic Parent-Child Interaction Coding System* (3rd edition).

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: Council of State Governments Justice Center; Public Policy Research Institute at Texas A&M University. Retrieved from: http://csgjusticecenter.org/youth/breaking-schools-rules-report.

Ferguson, E., & Houghton, S. (1992). The effects of contingent teacher praise, as specified by Canter's assertive discipline programme, on children's on-task behavior. *Educational Studies, 18,* 83-93.

Flannery, M. E. (2015). The school-to-prison pipeline: Time to shut it down. *NEA Today Magazine, 33*(1), 42–45.

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, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).

Forman, S. G., Shapiro, E. S., Codding, R. S., Gonzales, J. E., Reddy, L. A., Rosenfield, S. A., & Stoiber, K. C. (2013). Implementation science and school psychology. School Psychology Quarterly, 28, 77–100.

Green, A. E., Willging, C. E., Zamarin, K., Dehaiman, L. M., & Ruiloba, P. (2019). Cultivating healing by implementing restorative practices for youth: Protocol for a cluster randomized trial. *International Journal of Educational Research, 93,* 168-176. doi: 10.1016/j.ijer.2018.11.005.

Gershenson, R. A., Lyon A. R., & Budd, K. S. (2010). Promoting positive interactions in the classroom: Adapting Parent-Child Interaction Therapy as a universal prevention program. *Education & Treatment of Children, 33*(2), 261-287.

Gregg, M., & Seigworth, G. J. (2010). *The affect theory reader*. Durham, NC: Duke University Press.

Gregory, A., Clawson, K., Davis, A., & Gerewitz, J. (2016). The promise of restorative practices to transform teacher-student relationships and achieve equity in school discipline. *Journal of Educational and Psychological Consultation, 26*(4), 325-353. doi: 10.1080/10474412.2014.929950.

Gunter, P. L., & Coutinho, M. J. (1997). Negative reinforcement in classrooms: What we're beginning to learn. *Teacher Education and Special Education, 20,* 249–264.

Gunter, P. L., Denny, R. K., Jack, S. L., Shores, R. E., & Nelson, C. M. (1993). Aversive stimuli in academic interactions between students with serious emotional disturbance and their teachers. *Behavioral Disorders, 18,* 265–274.

Gunter, P. L., Denny, R. K., Shores, R. E., Reed, T. M., Jack, S. L., & Nelson, M. (1994). Teacher escape, avoidance, and counter control behaviors: Potential responses to disruptive and aggressive behaviors of students with severe behavior disorders. *Journal of Child and Family Studies, 3,* 211–223.

Hall, R. V., Lund, D., & Jackson, D. (1968). Effects of teacher attention on study behavior. *Journal of Applied Behavior Analysis, 1,* 1-12.

Hall, R. V., Panyan, M., Rabon, D., & Broden, M. (1968). Instructing beginning teachers in reinforcement procedures which improve classroom control. *Journal of Applied Behavior Analysis, 1,* 315-322.

Haydon, T., Mancil, G. R., & Van Loan, C. (2009). Using opportunities to respond in a general education classroom: A case study. *Education & Treatment of Children, 32*, 267–278.

Holdaway, S., Davidson, N., Dignan, J., Hammersley, R., Hine, J., & Marsh, P. (2001). *New strategies to address youth offending: The national evaluation of the pilot youth offending teams.* RDS Occasional Paper, (69). London: Home Office.

Hophins, B., Masters, G. (2003). *Just schools: A whole school approach to restorative justice*. Jessica Kingsley Publishers.

International Institute for Restorative Practices. (2010). Restorative Practices: Whole-school implementation. Edwards Dolquist Implementation Tactics. Retrieved from http://www.iirp.edu/pdf/SSS_Implementation_Overview.pdf.

International Institute for Restorative Practices. (n.d.). What is restorative practices? Retrieved from https://www.iirp.edu/restorative-practices/what-is-restorative-practices.

Jackson, S. (1998). *Family justice? An evaluation of the Hampshire youth justice family group conferencing project.* Southampton: University of Wellington.

James, W. (1890). *The principles of Psychology*. Cambridge: Harvard Up.

James, W., and Lange, C. G. (1922). *The Emotions.* Baltimore: Johns Hopkins University Press. doi: 10.1007/BF00186675.

Jenkins, L. N., Floress, M. T., & Reinke, W. M. (2015). Rates and types of teacher praise: A review and future directions. *Psychology in the Schools, 52*, 463–476.

Joyce, B., & Showers, B. (2002). *Student achievement through staff development (3rd ed.).* Alexandria, VA: Association for Supervision and Curriculum Development.

Kane, J., Lloyd, G., McCluskey, G., Riddell, S., Stead, J., & Weedon, E. (2007). *An evaluation of restorative practices in 3 Scottish authorities,* Research Report. Edinburgh: SEED.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155-163.

Ladd, G. W., & Burgess, K. B. (1999). Charting the relationship trajectories of aggressive, withdrawn, and aggressive/withdrawn children during early grade school. *Child Development, 70*, 919–929.

LeBuffe, P. A., & Naglieri, J. A. (1999a). *Devereux early childhood assessment: User's guide.* Lewisvill, NC: Kaplan Press.

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.

Lyon, A R., Gershenson, R A., Farahmand, F. K., Thaxter, P. J., Behling, S., & Budd, K. S. (2009). Effectiveness of teacher-child interaction training (TCIT) in a preschool setting. *Behavior Modification, 33*, 855-884.

Ludwig, R. J., & Welch, M. G. (2019). Darwin's other dilemmas and the theoretical roots of emotional connection. *Frontiers in Psychology*. Retrieved from https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00683/full.

Marsh, P., & Crow, G. (1998). *Family group conferences in child welfare.* Oxford: Blackwell Science.

Marshall, T., & Merry, S. (1990). *Crime and accountability.* London: Home Office.

Maxwell, G., & Morris, A. (1993). *Family victims and culture: Youth justice in New Zealand.* Wellington: Social Policy Agency, and Institute of Criminology, Victoria University of Wellington.

McAllister, L. W., Stachowiak, J. G., Baer, D. M., & Conderman, L. (1969). The application of operant conditioning techniques in a secondary school classroom. *Journal of Applied Behavior Analysis, 2*, 277–285.

McCluskey, G., Lloyd, G., Stead, J., Kane, J., Riddell, S., & Weedon, E. (2008). 'I was dead restorative today': From restorative justice to restorative approaches in school. *Cambridge Journal of Education, 38*(2), 199-216. doi:10.1080/03057640802063262.

McCold, P. (2008). Evaluation of a restorative milieu: Restorative practices in context. In H. Ventura Miller (Ed.) *Restorative justice: From theory to practice* (pp. 99–137). Bingley, UK: Emerald Group Publishing Limited.

McCold, P., & Wachtel, T. (2002). Restorative justice theory validation. In: E. Weitekamp, H.-J. Kerner (Eds.) *Restorative justice: Theoretical foundations* (pp. 110–142). Devon, UK: Willan Publishing.

Miers, D. (2001). *An international review of restorative justice.* London: Home Office.

Nathanson, D. (1992). *Shame and pride: Affect, sex and the birth of the self.* New York: Norton & Company.

Newburn, T., Earle, R., Goldie, S., Campbell, A., Masters, G., Crawford, A., Sharpe, K., Hale, C., Saunders, R., & Uglow, S. (2001). *The introduction of referral orders into the youth justice system. Second interim report*. RDS Occasional Paper, (73). London: Home Office.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Oakland Unified School District (2014, September). *Restorative justice in Oakland schools implementation and impacts: An effective strategy to reduce racially discipline,*

suspensions and improve academic outcomes. Prepared for the Office of Civil Rights, U.S. Department of Education.

Reinke, W. M., Herman, K. C., Newcomer, L. (2016). The brief student-teacher classroom interaction observation: Using dynamic indicators of behaviors in the classroom to predict outcomes and inform practice. *Assessment for Effective Intervention, 42*(1), 32-42. doi: 10.1177/1534508416641605.

Rodman, B. (2007). Reclaiming youth: What restorative practices add to the thinking. *Reclaiming Children and Youth: The Journal of Strength-Based Interventions, 16*(2), 48–51.

Rosenberg, M. S. (1986). Maximizing the effectiveness of structured classroom management programs: Implementing rule-review procedures with disruptive and distractible students. *Behavioral Disorders, 11*, 239–248.

Schott Foundation, Advancement Project, American Federation of Teachers, & National Education Association. (2014). Restorative practices: Fostering healthy relationships & promoting positive discipline in schools. The Atlantic Philanthropies.

Shores, R. E., Gunter, P. L., & Jack, S. L. (1993). Classroom management strategies: Are they setting events for coercion? *Behavioral Disorders, 18*, 92–102.

Shriver, M. D., & Allen, K. D. (2008). *Working with parents of noncompliant children: A guide to evidence-based parent training for practitioners and students.* Washington, DC: American Psychological Association.

Skinner, C. H., Ford, J. M., & Yunker, B. D. (1991). A comparison of instructional response requirements on the multiplication performance of behaviorally disordered students. *Behavioral Disorders, 17*, 56-65.

Skinner, C. H.,& Shapiro, E. S. (1989). A comparison of taped-words and drill interventions on reading fluency in adolescents with behavior disorders. *Education and Treatment of Children, 12,* 123-133.

Skinner, C. H., Smith, E. S., & McLean, J. E. (1994). The effects of interatrial interval duration on sight-word learning rates in children with behavioral disorders. *Behavioral Disorders, 19*, 98-107.

Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology, 85,* 571–581.

Sprick, R. S. (2009). *Champs: A proactive and positive approach to classroom management for grades K-9.* Eugene, OR: Pacific Northwest Publishing.

Stichter, J. P., Lewis, T. J., Whittaker, T. A., Richter, M., Johnson N. W., & Trussell, R. P. (2009). Assessing teacher use of opportunities to respond and effective classroom management strategies: Comparison among high- and low-risk elementary schools. *Journal of Positive Behavior Interventions, 11*(2), 68-81.

Stormont, M., & Reinke, W. M. (2009). The importance of precorrection and behavior specific praise strategies. *Beyond Behavior, 18,* 26–32.

Sugai, G., & Horner, R. H. (2005). *School-wide positive behavior support: School leadership teams training reference manuals*. Retrieved from http://ea.niusileadscape. org/docs/FINAL_PRODUCTS/LearningCarousel/ SchoolWidePBS08.pdf

Sutherland, K. S., & Wehby, J. H. (2001). Exploring the relationship between increased opportunities to respond to academic requests and the academic and behavioral outcomes of students with EBD: A review. *Remedial and Special Education, 22,* 113–121.

Sutherland, K. S., Wehby, J. H., & Copeland, S. R. (2000). Effect of varying rates of behavior-specific praise on the on-task behavior of students with EBD. *Journal of Emotional and Behavioral Disorders, 8*, 2–8.

Sutherland, K. S., Alder, N., & Gunter, P. L. (2003). The effect of varying rates of opportunities to respond to academic requests on the classroom behavior of students with EBD. *Journal of Emotional and Behavioral Disorders, 11*, 239–248.

Tiano, J. D., & McNeil, C. B. (2006). Training head start teachers in behavior management using parent-child interaction therapy: A preliminary investigation. *Journal of Early and Intensive Behavioral Intervention, 3*(2), 220-233.

Tomkins, S. (1962). *Affect, imagery, consciousness (1).* New York: Springer.

Tomkins, S. (1963). *Affect, imagery, consciousness (2).* New York: Springer.

U.S. Department of Education. (2014). *Guiding Principles: A Resource Guide for Improving School Climate and Discipline*, Washington, D.C.

U.S. Department of Education. (2016). *Non-regulatory guidance: Using evidence to strengthen education investments.*

Umbreit, M., & Roberts, A. (1996). *Mediation of criminal conflict in England: An assessment of services in Coventry and Leeds.* Saint Paul, MN: Center for Restorative Justice and Peacemaking, University of Minnesota.

West, R. P., & Sloane, H. N. (1986). Teacher presentation rate and point delivery rate: Effect on classroom disruption, performance accuracy, and response rate. *Behavior Modification, 10*, 267-286.

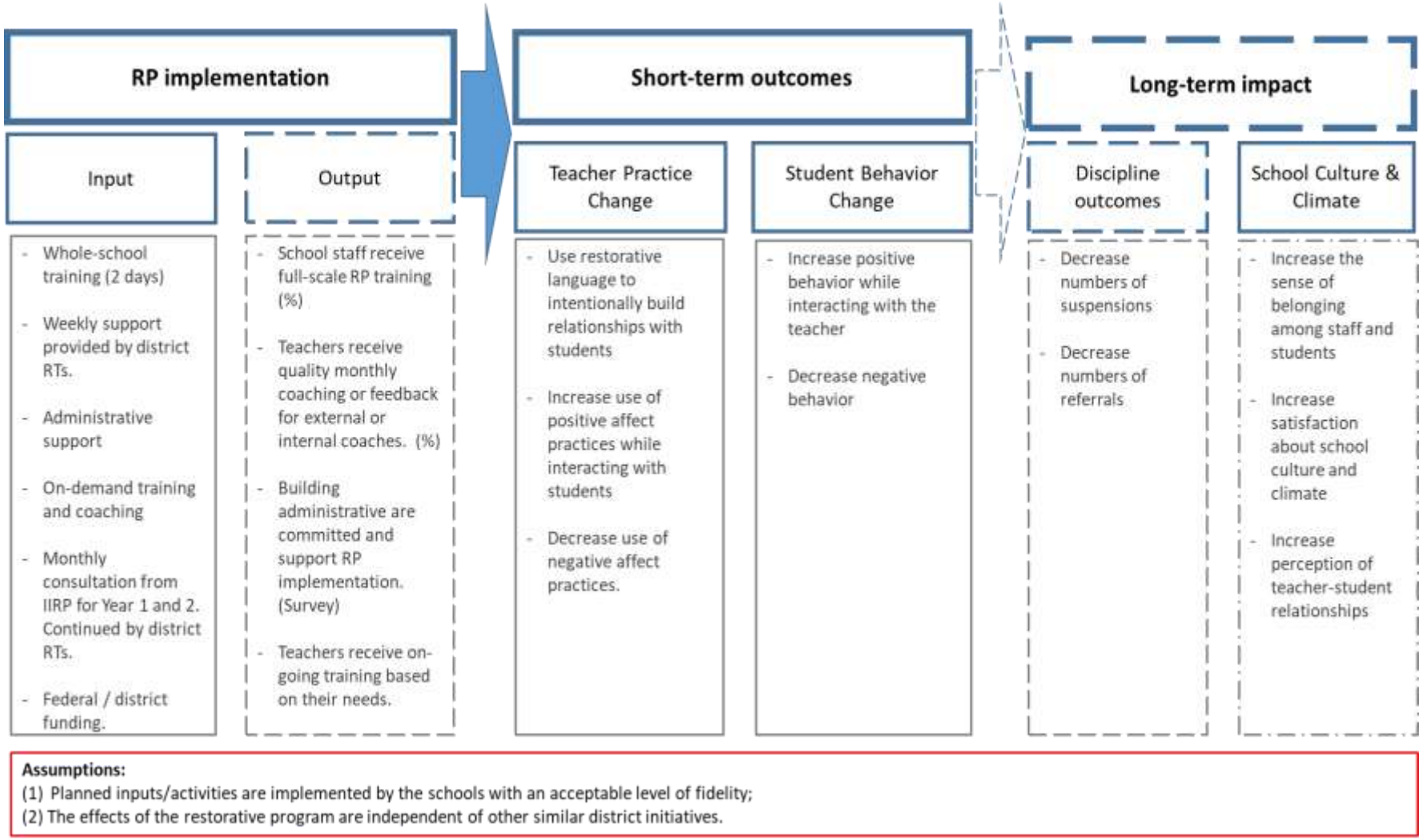
Wong, H. K., & Wong, R. T. (1998). *The first days of school: How to be an effective teacher: Mountain View.* CA: Harry Wong Publications.

Yoder, P., & Symons, F. (2010). *Observational measurement of behavior*. New York, NY: Springer.

Zehr, H. (1990). *Changing lenses: A new focus for criminal justice.* Scotsdale, PA: Herald Press.

Appendix

Appendix A. A Logic Model of the Study



Appendix B. Data Collection Timetable

Time	**District Evaluation** (requested and approved by the district without a need of IRB)	**Current Study** (must obtain IRB approvals from Bellarmine University and the school district)
November 2019	Observation tool and protocol have been developed and approved. School-level **randomized selection**: Cohort 1 (summer of 2017): 3 ES Cohort 2 (summer of 2018): 3 ES Cohort 3 (summer of 2019): 3 ES	
December 2019	Field test of the instrument (report IRR) Conduct **1st round observations**: All classrooms at all schools (9).	
February 2020	Conduct **2nd round observations**: All classrooms at all schools (9).	
April 2020*	Conduct **3rd round observations**: All classrooms at all schools (9).	
July 2020		Obtain IRB approvals Qualifier's list (2 criteria): the teachers must 1. Have all 3 observations 2. Be trained in initial training in the same school. Classroom-level **randomized selection** with IRB: Six (6) teachers from each school Sample of the study: 6 classrooms x 3 observations x 9 schools

Note. * Observations scheduled in April 2020 were not proceeded due to state-wide school closing as a result of the COVID-19 pandemic. The determination of sample size and the timeline for IRB application are re-evaluated and adjusted.

Appendix C. RP Classroom Observation Tool (report-use)

Restorative Practices (RP) Classroom Observation Tool

Date _____ **School** _____ **Grade** _____ **#of students** _____
Time _____ **Teacher** _____ **Observer** _____

Teacher Observation			% of Total
Positive Affect Practice	Subtotal		
Restorative Language			
Non-Specific Praise			
Opportunity-to-Respond (OTR) (content-related)			
Negative Affect Practice	Subtotal		
Negative Language Affect			
Negative Physical Affect			
	Total		100%

Student Observation			% of Total
Positive Affect Behavior	Subtotal		
Initiate an Interaction (with teacher)			
Respond to Content-Related Questions (with teacher)			
Positively Respond to Behavioral Request			
Negative Affect Behavior	Subtotal		
Distracting Voice Level			
Distracting Movement			
Negatively Respond to Behavioral Request			
	Total		100%

Student Engagement	Total	%
# Students on-task during ENTIRE observation		

Response to Problem Behavior

(ONLY applicable when problem behavior occurred during the observation.)

Note: Continue recording on page 1. Complete this page at the end of the observation.

Teacher Observation	Yes	No
Did the teacher use Restorative Questions?		
Did the teacher use an appropriate voice level?		
Did the teacher talk to the student in private?		
Did the teacher's physical affect contribute to de-escalation?		

Student Observation	Yes	No
Was the student's behavior de-escalated as a result of the teacher's effort?		
Was the student removed from the classroom?		

Circle Observation

(ONLY applicable when a circle occurred during the observation.)

Circle Observation	Yes	No
Was an academic circle observed?		
Was a community-building circle observed?		
Was a responsive circle observed?		

Appendix D. RP Classroom Observation Tool (Observer-use)

Restorative Practices (RP) Classroom Observation Tool

Date _____ **School** _____ **Grade** _____ **#of students** _____
Time _____ **Teacher** _____ **Observer** _____

Teacher Observation		
Positive Affect Practice		**Total**
Restorative Language		
Non-Specific Praise		
Opportunity-to-Respond (OTR) (content-related)		
Negative Affect Practice		
Negative Language Affect		
Negative Physical Affect		

Student Observation		
Positive Affect Behavior		**Total**
Initiate an Interaction (with teacher)		
Respond to Content-Related Questions (with teacher)		
Positively Respond to Behavioral Request		
Negative Affect Behavior		
Distracting Voice Level		
Distracting Movement		
Negatively Respond to Behavioral Request		

Student Engagement	
# Students on-task during ENTIRE observation	

Response to Problem Behavior

(<u>ONLY</u> applicable when problem behavior occurred during the observation.)

<u>Note:</u> Continue recording on page 1. Complete this page at the end of the observation.

Teacher Observation	Yes	No
Did the teacher use Restorative Questions?		
Did the teacher use an appropriate voice level?		
Did the teacher talk to the student in private?		
Did the teacher's physical affect contribute to de-escalation?		

Student Observation	Yes	No
Was the student's behavior de-escalated as a result of the teacher's effort?		
Was the student removed from the classroom?		

Circle Observation

(<u>ONLY</u> applicable when a circle occurred during the observation.)

Circle Observation	Yes	No
Was an academic circle observed?		
Was a community-building circle observed?		
Was a responsive circle observed?		

Appendix E. Observation Protocol (observer-use)

Restorative Practice (RP) Observation Protocol

Basic Procedure

- Write down the *basic information* at the top of the form. Record the *time* at the beginning of the observation.
- *Observe* the teacher's and students' behavior during classroom activities.
- *Record the* occurrences of the *behavior* (for multiple students or an individual) through tally marks on *page 1* (RP Classroom Observation Tool). Count the total number of students who are on-task during the entire observation. Record the number in the student engagement column.
- *Move* to a different location within the classroom *every 2 minutes*.
- If *problem behavior* occurs during the observation, *continue recording* on page 1.
- *Leave* the classroom after *10 minutes*. Record the time at the end of the observation.
- If a problem behavior incident was observed, *turn to* page 2 (Response to Problem Behavior), and *check* each question item (Yes or No) based on the observation of the incident.
- If multiple incidents occurred during the observation, only record *the first one*.
- If a circle was observed, *turn to* page 2 (Circle Observation), and *check* each question item (Yes or No) based on the observation of the circle.
- *Count* the tally marks. Record the total number in the Total box.
- Move to the next classroom. Start a new observation form.

Operational Definition of Terms

Teacher Observation	
Positive Affect Practice	
Restorative Language	The teacher provides a positive expression of student(s) behavior, including an affective statement, behavior-specific praise, small prompt conversation, personal talk for relationship building, and positive physical affects.
Non-Specific Praise	The teacher provides a positive verbal feedback related to behavior or academics. (Ex. Good job! Way to go!)
Opportunity-to-Respond (OTR) (content-related)	A verbal inquiry (question) or comment is provided to the student(s) with an opportunity to answer. It needs to be content-related. Ex. "Turn to page 87" is not counted as an OTR; "Turn to your partner and tell her/him the main idea of the story." is counted as an OTR. If a circle is observed, record a tally mark for each time the teacher asks a circle (or follow-up) question, even if not content-related.
Negative Affect Practice	
Negative Language Affect	The teacher uses a negative verbal expression of student(s) behavior, including sassy, sarcastic, rude, or impudent comments.
Negative Physical Affect	The teacher uses a negative non-verbal expression of student(s) behavior consciously or unconsciously. Ex. The teacher may shake her/his finger or put hands on her/his hips unintentionally.

Student Observation	
Positive Affect Behavior	
Initiate an Interaction (with teacher)	Student initiates a positive interaction (question, comment, etc.) with the teacher.
Respond to Content-Related Questions (with teacher)	Student provides a verbal or non-verbal response to a content-related question that answers or attempts to answer the question from the teacher. An appropriate choral response receives one tally mark for each question from the teacher. If a circle is observed, record a tally mark for each time that a student answers the circle (or a follow-up) question appropriately.
Positively Respond to Behavioral Request	Student performs or attempts to perform a behavioral request following the teacher's direction.
Negative Affect Behavior	
Distracting Voice Level	Student uses inappropriate voice level, which interrupts the classroom activity, including loud talking, screaming, or shouting.
Distracting Movement	Student uses physical movement, which interrupts the classroom activity.
Negatively Respond to Behavioral Request	Student does not attempt to answer the request or responds with negative comments or actions.

Student Engagement	
#Students on-task during ENTIRE observation	Students work on the current assignment and ignore distractions.

For Problem Behavior (complete the form at the end of the observation)

Teacher Observation	
Did the teacher use Restorative Questions?	If the teacher asked restorative questions (in some forms of variation) to the student, check YES. Otherwise, check NO.
Did the teacher use an appropriate voice level?	If the teacher talks to the student with a calm and respectful voice, check YES. Otherwise, check NO.
Did the teacher talk to the student in private?	If the teacher spoke with the student privately did not embarrass (shame) the student in front of the peer (intentionally or unintentionally), check YES. Otherwise, NO.
Did the teacher's physical affect contribute to de-escalation?	If the teacher presented in a respectful physical manner that contributed to the de-escalation of the situation, check YES. Otherwise, check NO.

Student Observation	
Was the student's behavior de-escalated as a result of the teacher's effort?	If the student started to calm down as a result of the teacher's effort, check YES. Otherwise, check NO.
Was the student removed from the classroom activity or the classroom?	If the student was removed from the classroom activity or the classroom (SRT call to the office), check YES. Otherwise NO.

For Circle Observation (complete the form at the end of the observation)

Circle Observation	
Was an academic circle observed?	If an academic circle was observed, check YES, Otherwise NO. Ex. A circle was used for students to share or deepen the understanding of academic content.
Was a community-building circle observed?	If a community-building circle was observed, check YES. Otherwise NO. Ex. A circle was used for students to share feelings, ideas, and experiences to build mutual understanding and positive relationships within the classroom community.
Was a responsive circle observed?	If a responsive circle was observed, check YES. Otherwise NO. Ex. After an incident, a circle was used for students to share feelings, repair relationships, and solve problems.

Appendix F. Inter-Rater Reliability of the Instrument

Restorative Practices (RP) Classroom Observation Tool
Inter-Rater Reliability

Observed classrooms: 6 (ES)

Observation time: 9:00am-11:00am

Raters: 2 consistent raters

Teacher Observation	Intraclass Correlation	95% Confidence Interval	
		Lower Bound	Upper Bound
Positive Affect Practice	**.546**	**-.355**	**.921**
	.706	***-1.101***	***.959***
Restorative Language	.645	-.213	.941
	.784	*-.542*	*.970*
Non-Specific Praise	.415	-.494	.891
	.586	*-1.955*	*.942*
Opportunity-to-Respond (OTR) (content-related)	.887	.399	.983
	.940	*.570*	*.992*
Negative Affect Practice	**.546**	**-.355**	**.921**
	.706	***-1.101***	***.959***
Negative Language Affect	.792	.095	.968
	.884	*.173*	*.984*
Negative Physical Affect	-.006	-.757	.752
	-.011	*-6.227*	*.858*

Student Observation	Intraclass Correlation	95% Confidence Interval	
		Lower Bound	Upper Bound
Positive Affect Behavior	**.950**	**.691**	**.993**
	.974	***.817***	***.996***
Initiate an Interaction (with teacher)	.966	.781	.995
	.983	*.877*	*.998*
Respond to Content-Related Questions	.995	.962	.999
	.997	*.981*	*1.000*
Positively Respond to Behavioral Request	.992	.943	.999
	.996	*.971*	*.999*
Negative Affect Behavior	**.950**	**.691**	**.993**
	.974	***.817***	***.996***
Distracting Voice Level	.951	.695	.993
	.975	*.820*	*.996*
Distracting Movement	.917	.526	.988
	.957	*.689*	*.994*
Negatively Respond to Behavioral Request	.583	-.306	.929
	.737	*-.880*	*.963*

Note: Intraclass correlation coefficient (ICC) is based on a two-way random effects model and consistency type. Statistics in black are single measures. Statistics in gray are average measures.

Restorative Practices (RP) Classroom Observation Tool

Inter-Rater Reliability

Observed classrooms: 6 (ES), 5 (MS), 5 (HS)
Observation time: 9:00am-11:00am
Raters: 2 consistent raters

Teacher Observation	Intraclass Correlation	95% Confidence Interval	
		Lower Bound	Upper Bound
Positive Affect Practice	**.755**	**.430**	**.907**
	.861	***.601***	***.951***
Restorative Language	.702	.332	.885
	.825	*.498*	*.939*
Non-Specific Praise	.703	.334	.885
	.826	*.501*	*.939*
Opportunity-to-Respond (OTR) (content-related)	.924	.797	.973
	.960	*.887*	*.986*
Negative Affect Practice	**.755**	**.430**	**.907**
	.861	***.601***	***.951***
Negative Language Affect	.780	.478	.917
	.877	*.647*	*.957*
Negative Physical Affect	.289	-.224	.677
	.448	*-.579*	*.807*

Student Observation	Intraclass Correlation	95% Confidence Interval	
		Lower Bound	Upper Bound
Positive Affect Behavior	**.934**	**.822**	**.976**
	.966	***.902***	***.988***
Initiate an Interaction (with teacher)	.956	.879	.984
	.978	*.936*	*.992*
Respond to Content-Related Questions	.967	.908	.988
	.983	*.952*	*.994*
Positively Respond to Behavioral Request	.948	.858	.981
	.973	*.923*	*.991*
Negative Affect Behavior	**.934**	**.822**	**.976**
	.966	***.902***	***.988***
Distracting Voice Level	.860	.645	.949
	.925	*.784*	*.974*
Distracting Movement	.751	.422	.905
	.858	*.593*	*.950*
Negatively Respond to Behavioral Request	.877	.685	.955
	.935	*.813*	*.977*

Note: Interclass correlation coefficient (ICC) is based on a two-way random effects model and consistency type. Statistic in black are single measures. Statistic in gray are average measures.