

---

# **Restorative Justice Programming and Student Behavioral and Disciplinary Outcomes**

**Lauren Rich, Nicholas Mader, and Aida Pacheco-Applegate**

**Restorative Justice Programming and Student Behavioral and Disciplinary Outcomes**

**Lauren Rich, Nicholas Mader, and Aida Pacheco**

**Recommended Citation**

**Rich, L., Mader, N., & Pacheco-Applegate, A. (2017, November). *Restorative justice programming and student behavioral and disciplinary outcomes*. Chicago, IL: Chapin Hall at the University of Chicago.**

**© 2017 by Chapin Hall at the University of Chicago**

**1313 East 60th Street**

**Chicago, IL 60637**

**ISSN:1097-3125**

---

**Acknowledgements:**

**This study was funded by the Pritzker Pucker Family Foundation.**

---

**Disclaimer:**

**This material is based on data provided by Chicago Public Schools (CPS) and a local community-based organization (CBO). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of CPS or the CBO.**

---

## Table of Contents

|  |    |
|--|----|
| Executive Summary .....  | 1  |
| Restorative Justice Programming and Student Behavioral and Disciplinary Outcomes ..... | 3  |
| Definition of Restorative Justice .....  | 3  |
| Implementation of Restorative Justice in Chicago Public Schools .....                  | 4  |
| Previous Research on Effect of Restorative Justice on Student Outcomes .....           | 6  |
| Model of Student Participation in the ‘Peace Room’ and Research Questions.....         | 8  |
| Study Context .....  | 11 |
| 2014–15 Statistics .....   | 12 |
| 2015–16 Statistics .....   | 14 |
| Description of Methods and Study Sample .....  | 15 |
| Propensity Score Matching—Overview.....  | 15 |
| Analysis of Peace Room Interactions/Generation of Program Participants .....           | 17 |
| Propensity Score Matching: Description of Process .....                                | 19 |
| Findings .....   | 21 |
| Characteristics of the Study Samples .....   | 21 |
| Results of PSM Analysis.....   | 23 |
| Summary.....   | 27 |
| Limitations and Recommendations for Future Research.....                               | 28 |
| References.....  | 30 |

---

---

**List of Tables and Figures**

|   |    |
|---|----|
| Figure 1. Percentage of Students with or without Disciplinary Infractions .....   | 12 |
| Figure 2. Disciplinary Infractions and Suspensions, by Student Demographics .....   | 13 |
| Figure 3. Percentage of Students who Interacted with PR Staff .....   | 14 |
| Table 1. Characteristics of Students Attending Focal Schools .....  | 22 |
| Figure 4. Propensity Estimation: Odds Ratios for Predictors of Likelihood of Participating in Peace Room Programming..... | 24 |
| Table 2. Estimated Differences between Peace Room Participants and Matched Comparisons .....                              | 25 |

---

# Executive Summary

Over the past two decades, policymakers, educators, and others, both in the US and abroad, have increasingly proposed restorative justice (RJ) practices as an alternative to “zero tolerance” policies in schools. Starting in the mid-1990s, these policies were enacted in almost every state in the US. While the purpose of these policies was to ensure safe learning environments, evidence has emerged that an overreliance on the exclusionary disciplinary practices associated with them actually increased the likelihood, for some youth, of entering the juvenile justice system. By contrast, RJ proponents maintain that RJ practices—such as the use of mediation for resolving conflicts—can reduce the use of suspensions and expulsions, thus boosting the chances that at-risk youth stay in high school and graduate.

This report examines the effect of a primary component of the RJ programming of one community-based organization (CBO), on students’ behavioral and disciplinary outcomes in four high schools in the city of Chicago from 2014 to 2016. Specifically, we examine the effect of students’ use of RJ programs offered in the CBO’s “Peace Room” (PR) on rates of disciplinary infractions, in-school suspensions (ISS) and out-of-school suspensions (OSS). To do this, we used data on student utilization of the PR compiled by the CBO staff and linked it with data on student disciplinary infractions and suspensions obtained from Chicago Public Schools (CPS). We then employed the method of propensity score matching to compare outcomes of students who had one or more infractions and utilized the PR to similar students who attended the same four high schools prior to the implementation of the PR.

We found that the percentage of PR participants who had one or more misconducts and received out-of-school suspensions (OSS) was 30 percentage points lower than the percentage for students in the comparison group. Specifically, 43 percent of the former received OSS versus 73 percent of the latter. We also found that students who had one or more incidents of misconduct and utilized the PR experienced about 4 fewer days of OSS than students in the comparison group. However, it should be noted that, because CPS adopted policies to reduce schools’ reliance on OSS during the time period covered by this study, it is difficult to separate out effects due to the policy changes from those due solely to utilizing the PR. Therefore, these results most likely overestimate the effect of utilizing RJ programs offered by the PR.

We also found that the percentage of PR participants who received in-school suspensions (ISS) was 14 percentage points lower than the percentage for students in the comparison group; 25 percent of the former received ISS versus 39 percent of the latter. In addition, we found that students who utilized the PR experienced about 1 fewer day of ISS than students in the comparison group. Given that the changes in CPS policies during the study time period focused attention primarily on OSS, and that the overall percentages of students receiving ISS in CPS increased during the study time period, we are more confident that these results primarily reflect the effects of participating in programs offered by the PR.

The research team also looked at whether different demographic groups of students benefited more from participating in PR restorative justice practices than others. We found that, with respect to rates of ISS among students with one or more misconducts, the beneficial effects of participating in PR programming were (1) larger for females than for males; (2) larger for Hispanics than for African Americans; (3) larger for Hispanic females than for Hispanic males; (4) larger for ninth graders than for students in other grades; (5) larger for students receiving free or reduced-price lunch than for those not; and (6) larger for

students with an individual education plan (IEP) than for students without. Finally, while the main results did not show statistically significant differences between PR participants and comparison students in the number of misconducts in the current year, we did find significant benefits for African Americans versus Hispanics, African American males versus all other race/gender groups, students receiving free or reduced-price lunch versus those not, and students with an IEP versus those without.

In sum, this study provides promising evidence of the positive effects of some RJ practices on student disciplinary outcomes. It also suggests that the positive effects may be larger in some cases for students who are at greater risk for poorer outcomes.

---

# Restorative Justice Programming and Student Behavioral and Disciplinary Outcomes

Over the past two decades, policymakers, educators, and others, both in the US and abroad, have increasingly proposed restorative justice (RJ) practices as an alternative to “zero tolerance” policies in schools. These policies were enacted in almost every state in the US, starting in the mid-1990s. While the purpose of these policies was to ensure safe learning environments, evidence has emerged that an overreliance on the exclusionary disciplinary practices associated with them actually increased the likelihood, for some youth, of entering the juvenile justice system (Stinchcomb, Bazemore, & Riestenberg, 2006). By contrast, RJ proponents maintain that RJ practices can reduce the use of suspensions and expulsions, thus boosting the chances that at-risk youth stay in high school and graduate.

This report examines the effect of a primary component of the RJ programming of one community-based organization (CBO) on students’ behavioral and disciplinary outcomes in four high schools in the city of Chicago from 2014 to 2016.<sup>1</sup> Specifically, we examine the effect of students’ use of the CBO’s “Peace Room” (PR) on rates of disciplinary infractions, in-school suspensions (ISS) and out-of-school suspensions (OSS). We begin with background information on how restorative justice has been defined and implemented in schools, including the Chicago Public Schools (CPS). We then briefly review previous studies of the effects of RJ practices on student outcomes. Finally, we describe the RJ programming offered by the CBO, the assumptions we made about the process by which students participated in the CBO’s Peace Room, our methods of analysis, and the results of our analyses.

---

## Definition of Restorative Justice

According to the Youth Justice Board for England and Wales (2004), the fundamental principle of restorative justice is that, when one person has harmed another, the most useful response is to try and repair the harm done, rather than to retaliate against the perpetrator. Additionally, a number of scholars

---

<sup>1</sup> The four high schools are Lincoln Park, South Shore International College Preparatory (SSICP), Sullivan, and Tilden Career Community Academy (Tilden). Also, while we have data on the CBO’s RJ programming for 2013-14 for two of the schools, for reasons explained later in this report we are focusing on the school years 2014-15 and 2015-16.

have proposed more expansive definitions of the term. For example, Van Ness and Strong (1997) suggest the following definition:

Restorative justice is different from conventional justice processes in that it views crime primarily as injury (rather than primarily as law-breaking), and the purpose of justice as healing (rather than punishment alone). It emphasizes accountability of offenders to make amends for their actions, and focuses on providing assistance and services to the victims. Its objective is the successful reintegration of both victim and offender as productive members of safe communities.

Initially, as suggested by the definition above, restorative justice practices were implemented in the criminal justice system, but they have since been introduced into schools worldwide. According to McCluskey et al. (2008), when RJ is implemented in schools there is less of an emphasis on giving victims a voice. Instead, schools more often focus on fostering good relationships following conflict or harm and, more generally, on developing a school culture (including policies and procedures) that reduces the likelihood of conflict. According to Cameron and Thorsborne (2001), schools with a RJ orientation view student misconduct not as a violation against the school as an institution, but as a violation against people and relationships in the school and wider community. Such schools respond to student misconduct in a way that emphasizes the importance of repairing and maintaining those relationships.

Schools in the UK have implemented such RJ practices as conflict resolution circle time, peer mediation, and restorative conferencing (Youth Justice Board for England and Wales, 2004). During regular circle times, a teacher and his or her students meet in a circle to talk and share concerns. Such meetings are meant to give students a safe place to air their concerns and thus prevent the escalation of problems (Armstrong, 2015). In peer mediation, a group of students receives mediation training and then helps other students to reach agreement and reconciliation following disputes (Powell, Muir-McClain, & Halasyamani, 1995). Finally, a restorative conference is a facilitated meeting between two or more students involved in a conflict, as well as teachers, peers, and others who might have a significant relationship to the students. The purpose of such meetings is to allow all who are affected by negative behavior to deal with the consequences of that behavior and jointly decide how to repair the harm done (Youth Justice Board for England and Wales, 2004).<sup>2</sup>

---

## **Implementation of Restorative Justice in Chicago Public Schools**

In July 2006, the Chicago Board of Education released the following policy statement:

---

<sup>2</sup> A facilitated meeting without the presence of others with significant relationships to the students involved in the conflict is called a “restorative mediation”.

. . . the Board recognizes and embraces the philosophy of restorative justice. Restorative justice principles involve those who have a stake in a specific offense in collectively identifying and addressing the harm done, the needs and obligations of all involved in order to heal and correct the situation as fully as possible. The Board encourages principals and administrators to adopt and implement restorative justice philosophies and practices as additional tools to address student misconduct. When restorative justice practices are available and adequate and when all parties voluntarily embrace and participate in restorative justice practices, these practices should be utilized as outlined in the SCC (Student Code of Conduct).

In addition, the Board stated that, during the 2006–07 school year, the Chief Executive Officer would work with stakeholders to develop school supports that would “reflect a comprehensive approach to student discipline” that included components of restorative justice.

By the fall of 2010, some community-based organizations in Chicago concluded that CPS had not made enough progress on the RJ front. So, in the fall of 2010, seven organizations launched the High HOPES (Healing Over the Punishment of Expulsions and Suspensions) campaign (High HOPES Campaign, 2012). As part of this campaign, the organizers held a community summit on using restorative justice practices in the school environment. They also met with CPS officials and held a public session with the Chicago Board of Education. This campaign appears to have at least partially resulted in a subsequent series of CPS reforms:

- In late 2010, CPS initiated a new program aimed at increasing student safety and improving school climate. The program, the Culture of Calm Initiative, took place in 47 high schools. At least some of the funds provided by the program were used by schools to implement peer mediation during the 2010–11 school year (Levenstein, Spote, & Allensworth, 2011).
- In the fall of 2012, CPS revised the Student Code of Conduct (SCC) to reduce the length of suspensions (Stevens, Sartain, Allensworth, & Levenstein, 2015). Specifically, the changes eliminated automatic 10-day suspensions and required principals to seek district approval to suspend students for more than five days. They also included a recommendation for the use of “nonexclusionary” practices, such as peace circles, to help resolve conflicts.
- In June 2014, the SCC was revised again. The changes limited the use of out-of-school suspensions (OSS) to instances where a student’s attendance at school endangered others, or when the student caused chronic or extreme interruption to other students’ participation in school activities and prior interventions have been utilized (Chicago Public Schools, 2014). They also removed OSS as an available consequence for minor misconducts, and lowered the number of suspension days permitted for repeated or more serious misconducts. Additionally, a new district-designed professional

development plan required RJ training for all security staff and that RJ training be made available to high-need schools.

These policy changes must be kept in mind when interpreting the results of this study. In particular, it is possible that the changes, especially those occurring in the summer of 2014, were successful in reducing the likelihood that a student who committed an offense during that year would receive an OSS.

Additionally, as Stevens et al. (2015) show, the percentage of all high school students who received an OSS declined from 24 percent in the 2009–10 school year to 16 percent in 2013–14, although they could not determine the extent to which these changes were due to policy changes. Without additional information, when OSS rates are declining across the school district, it is difficult to determine whether an OSS decline in a particular school is due to programs that school implemented or to factors also affecting other schools. However, it should be noted that, while OSS rates across CPS have declined during the time period covered by this study (i.e. the 2010-11 through 2015-16 school years), rates of ISS have increased somewhat. Specifically, during the time period covered by this study, the percentage of CPS high school students who had one or more misconducts and received at least one OSS declined from 71.2% to 38.9%. However, during this same period, the percentage of CPS high school students who had one or more misconducts and received at least one ISS increased from 40.4% to 46.3%. Thus, we can be more confident that results suggesting a positive impact of RJ programming on ISS in the four schools involved in the study are not simply a reflection of policy changes. Similarly, because the policy changes did not attempt to impact the number of students who had misconduct reports, it is more likely that any decreases in misconduct (especially severe misconduct) found at the four schools in the study could be interpreted as being a result of RJ interventions.

---

## **Previous Research on Effect of Restorative Justice on Student Outcomes**

RJ practices first emerged in Australia in the 1990s and have been used in schools worldwide for over two decades (Stinchcomb et al., 2006). However, few studies have attempted to rigorously assess their impacts on students and schools.<sup>3</sup> A search of both peer-reviewed journal articles and unpublished reports yielded only three studies that attempted such an assessment: Youth Justice Board for England and Wales (2004); Wong, Cheng, Ngan, and Ma (2011); and Gregory, Clawson, Davis, and Gerewitz (2016). Below we briefly summarize these three studies and discuss their limitations.

---

<sup>3</sup> However, some researchers have documented changes in suspensions and/or disciplinary referrals over time in schools implementing RJ (though not in relation to a comparison group). For example, Stinchcomb et al. (2006) report that ISS, OSS and behavior referrals declined, sometimes dramatically, after the introduction of RJ practices in schools in South St. Paul, Minnesota. Additionally, some researchers have documented perceived changes in school climate based on interviews and survey data collected from staff and students (McCluskey et al., 2008; Ortega, Lyubansky, Nettles, & Espelage, 2016).

The Youth Justice Board for England and Wales (2004) study is an evaluation focused on 26 schools in England that participated in a restorative justice program funded by the Youth Justice Board (YJB). The schools implemented a range of restorative practices, including “active listening and restorative enquiry,” circle time and mediation, and peer mediation. As part of the evaluation, students and school staff in the 26 schools, as well as students and staff in similar schools, completed surveys on the prevalence of bullying and feelings of safety prior to and after RJ implementation. Overall, YJB researchers did not find any statistically significant differences over time between schools that implemented RJ practices and schools that did not, although trends suggested a potentially positive impact of RJ.

However, the researchers did find positive results for a subgroup of schools. In schools where RJ had been implemented for the longest period of time, students were significantly more likely to say that their schools were doing a good job of stopping bullying and significantly less likely to say bullying was a serious problem at their schools. Additionally, the percentage of staff who said student behavior had improved over the past year increased significantly, and the percentage who said behavior had worsened decreased significantly. In contrast, no statistically significant changes were seen in the comparison schools, and trends appeared to be moving in the opposite direction of those in the intervention schools.

The main limitation of this study is that the researchers report changes in student outcomes between baseline and follow-up for the program and comparison schools, but do not provide data on student outcomes in the schools at baseline. Thus, it is possible that differences in levels of the outcomes at baseline could have influenced some of the differences in changes observed. For example, if the comparison schools began with higher levels of bullying, the results could be biased towards showing an impact from RJ. Additionally, as pointed out by the authors, both program and comparison schools had a range of other initiatives designed to have an impact on student behavior, making it difficult to isolate the effect of RJ.

Wong et al. (2011) examined the impact of implementation of a “restorative whole school approach” on seventh through ninth grade students attending four schools in Hong Kong. School staff received training on drafting antibullying policies, the use of mediation for resolving conflicts, implementation of a peace education curriculum, and the use of student competitions to “facilitate the building of a harmonious school climate”. Following the training, the evaluation team determined the extent to which the approach had been implemented in each school. They determined that one school had fully implemented the approach, while two partially implemented it, and one did not implement it at all.

Students at the four schools completed baseline and follow-up surveys on their own behavior including: bullying and caring behaviors, frequency of hurting others, lack of empathy, sense of belonging and perceptions of school harmony. The researchers found (1) a small decline in bullying at the fully and

partially implementing schools, and a small increase at the non-implementing school; (2) a small increase in empathy for others at the fully implementing school (and no change at the other schools); and (3) small declines in perceived school harmony and sense of belonging at the non and partially implementing schools and no change at the fully implementing school. Thus, overall, there appeared to be some positive benefits of implementing the restorative whole school approach. However, a limitation of this study is the use of the non- and partially implementing schools as the sole comparisons. The fact that these schools received training but did not fully implement the program suggests that there might have been factors that prevented full implementation and also impacted the outcomes examined. In addition, as noted by the authors, students at the implementing school reported less bullying, more caring behavior, and more positive perceptions of school climate at baseline. This may have influenced some of the results.

Finally, Gregory et al. (2016) examined the relationship between the extent of teachers' implementation of restorative practices and two outcomes: the quality of teacher-student relationships and teachers' likelihood of referring students for misconduct or defiance. The study focused on over 400 students in 29 classrooms in two high schools located in a small city on the east coast of the United States. All of the participating teachers received training in the use of restorative practices at the beginning of the school year. Researchers assessed their implementation of these practices later in the year by surveying students and teachers. Surveys asked students how often their teachers used various restorative practices and treated them in a respectful manner. Teachers completed surveys on their own implementation of RJ practices. Also, researchers obtained data on misconduct and defiance referrals from school records. The results showed that, after accounting for student race/ethnicity and teacher rating of student cooperativeness, teachers who implemented RJ more fully were more likely to have better relationships with students and to issue fewer misconduct and defiance referrals to Latino and African American students. A limitation of this study is the lack of a comparison sample of teachers who did not receive training in RJ. In particular, it is possible that teachers who were low implementers of RJ had other characteristics that also negatively impacted their relationships with students or increased their tendency to refer students for misbehavior.

Overall, these three studies are suggestive of potentially positive effects of RJ on student outcomes, but more research employing carefully selected comparison groups is necessary to confirm these effects.

---

## **Model of Student Participation in the 'Peace Room' and Research Questions**

The CBO that is the focus of this study utilizes a "whole school restorative justice model", which includes:

- Strategic planning with administrators and school leaders with respect to school-wide discipline and behavior systems
- A Peace Room (PR) which serves as a hub for restorative practices and interventions, including peace circles, mediations, restorative chats & re-entry circles (e.g., for students returning from a suspension)
- A disciplinary intervention curriculum which teaches social-emotional & leadership skills
- Professional development on restorative justice practices for teachers and school staff

In this study we focus specifically on student utilization of the Peace Room, since we did not have access to implementation data for the other three components of the CBO's programming.

Our first task in assessing the impact of student utilization of the PR was to develop an underlying model, or representation, for how students come to participate in the PR and how that participation might impact student behavioral and disciplinary outcomes. Developing such a model is important, because it provides clarity with respect to which students should be the main focus of the analysis. It also helps to clarify the most important variables to be used in the analysis and the potential relationships between these variables.

In order to gain an understanding of the ways in which students and staff used the PR, we conducted an examination of the PR log sheets maintained by CBO staff. These sheets contained information for each student who had an interaction with the staff and included the student's ID number, date of the interaction, the role of the person who referred the student to the PR (e.g., teacher or administrator), and the type of RJ practice utilized. Some sheets also contained a Student Code of Conduct (SCC) code for the incident associated with the referral, as well as detailed notes about the incident. We noted that there were primarily two different ways in which students might ultimately end up interacting with PR staff. First, students might come to the PR of their own volition. We noted that students might elect to come to the PR for a wide variety of reasons, including to: process grief over the loss of a family member; express frustrations about relationships with peers, teachers, or family members; get advice on ways to navigate difficult relationships with peers, teachers, or family members; get advice on how to make friends; discuss problems at home (other than relationship problems); and discuss progress made following previous interactions with PR staff. Second, a teacher, dean or other administrator, security guard, or other staff member might refer students to the PR. Teachers appeared to refer students most often for classroom disruptions or negative interactions with the teacher. Deans, administrators, and security staff appeared to refer most often for physical or verbal altercations, threats, and bullying. Finally, other staff appeared to refer for a mixture of conflicts and personal issues.

In considering the important features of the model, we decided to focus on student interactions with PR staff members that occurred as a result of a student being referred by a school staff member. We decided to put aside student self-referrals on the assumption that, given the limitations of the data available to us, it would be much harder to identify the characteristics of students who decided to come to the PR on their own. Identifying those characteristics is a crucial part of the analysis, as those characteristics allow us to match students who use the PR to comparison group students who did not interact with PR staff.

Below are our assumptions and hypotheses about the process by which a student ends up going to the PR. These assumptions and hypotheses constitute our model:

1. An incident occurs in which one or more students engage in a behavior that could be considered to be a violation of the CPS SCC.<sup>4</sup>
2. A staff member who witnesses the incident decides whether to refer the student(s) to the PR. We do not have information on the criteria that staff members might employ in making this decision, but we surmise that the severity of the behavior might be a factor.
3. We assume that the student(s) can elect to go to the PR or not. In cases where students elect to go to the PR, an appropriate consequence would then be assigned; this consequence might or might not include an in- or out-of-school suspension. However, we assume that if a student elects to go to the PR the consequence would be much less likely to include suspension, especially if the PR activity results in a successful resolution of the incident.<sup>5</sup> For this reason, there might be a direct relationship between going to the PR and the likelihood of receiving a suspension.
4. Finally, we hypothesize that the skills that students gain through their interactions with the PR might make it less likely that they engage in future behavior of the sort that led to disciplinary referrals in the first place. For example, students who are able to learn conflict resolution skills might be more likely to employ those in future instances when there is the possibility of a verbal or physical altercation. For this reason, we might expect participation in RJ programming to impact students' future behavior and any subsequent disciplinary outcomes.

Utilizing this model as the framework for our study, we addressed the following research questions:

---

<sup>4</sup> According to the CPS SCC incidents of misconduct are assigned to one of six groups based on the severity of the behavior, ranging from "inappropriate behaviors" at level 1 to "illegal behaviors" at level 6.

<sup>5</sup> We recognize that, given the available information, it is not possible to identify students who are referred to the PR but don't go. Therefore, in the actual analysis, we only consider the likelihood that a student is referred to the PR.

- Do students who attended the four schools, and were referred to and utilized the PR, have lower rates of misconduct and suspensions compared to similar students who attended the four schools before the CBO entered them?<sup>6</sup>
  - ♦ Do students who utilize the PR have lower rates of misconduct and suspensions in the same year in which they participate in PR programming?
  - ♦ Do students who utilize the PR have lower rates of misconduct and suspensions in the year subsequent to the year in which they participate in PR programming?
- Does the effect of participating in PR programming vary across different subgroups of students?

To answer these questions we employed the method of propensity score matching (described in detail below) using administrative data from CPS on misconducts and suspensions, and data on student interactions with PR staff recorded by the CBO.

---

## Study Context

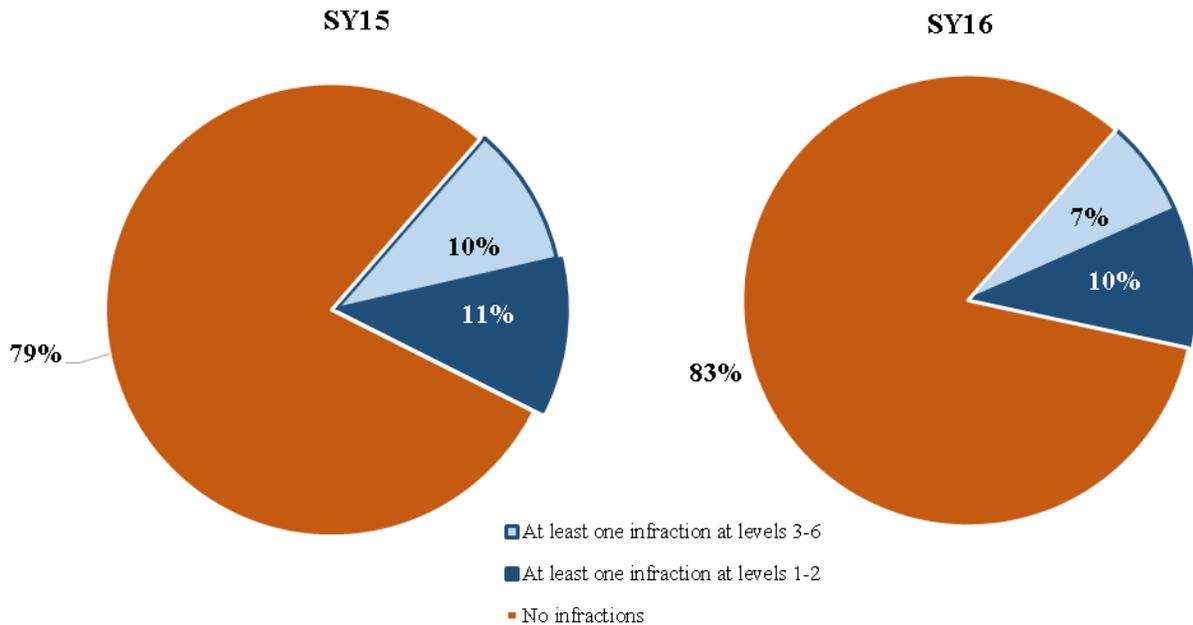
This study focuses on four schools in which the CBO worked during the 2014-15 and 2015-16 school years: Lincoln Park High School (Lincoln Park), South Shore International College Preparatory (SSICP), Sullivan High School (Sullivan), and Tilden Career Community Academy (Tilden). Briefly, Lincoln Park is a racially and ethnically diverse neighborhood school. It has a student population that is 55% low-income, and is located on the north side of Chicago. Sullivan is also a racially and ethnically diverse neighborhood school located on the north side. In contrast to Lincoln Park, however, the vast majority of its students are low-income. SSICP is a selective enrollment school serving a predominately African-American and low-income population on the south side of Chicago. Finally, Tilden is a neighborhood school that also serves a predominately African-American and low-income population on the south side.

We generated some basic descriptive statistics in order to gain additional insight into the context of the four schools. These statistics were generated using administrative data from CPS and data on student interactions with PR staff recorded by the CBO.

---

<sup>6</sup> In the remainder of the report, for simplicity we will use the term “utilized the PR” to mean “referred to and utilized the PR”.

**Figure 1. Percentage of Students with or without Disciplinary Infractions**



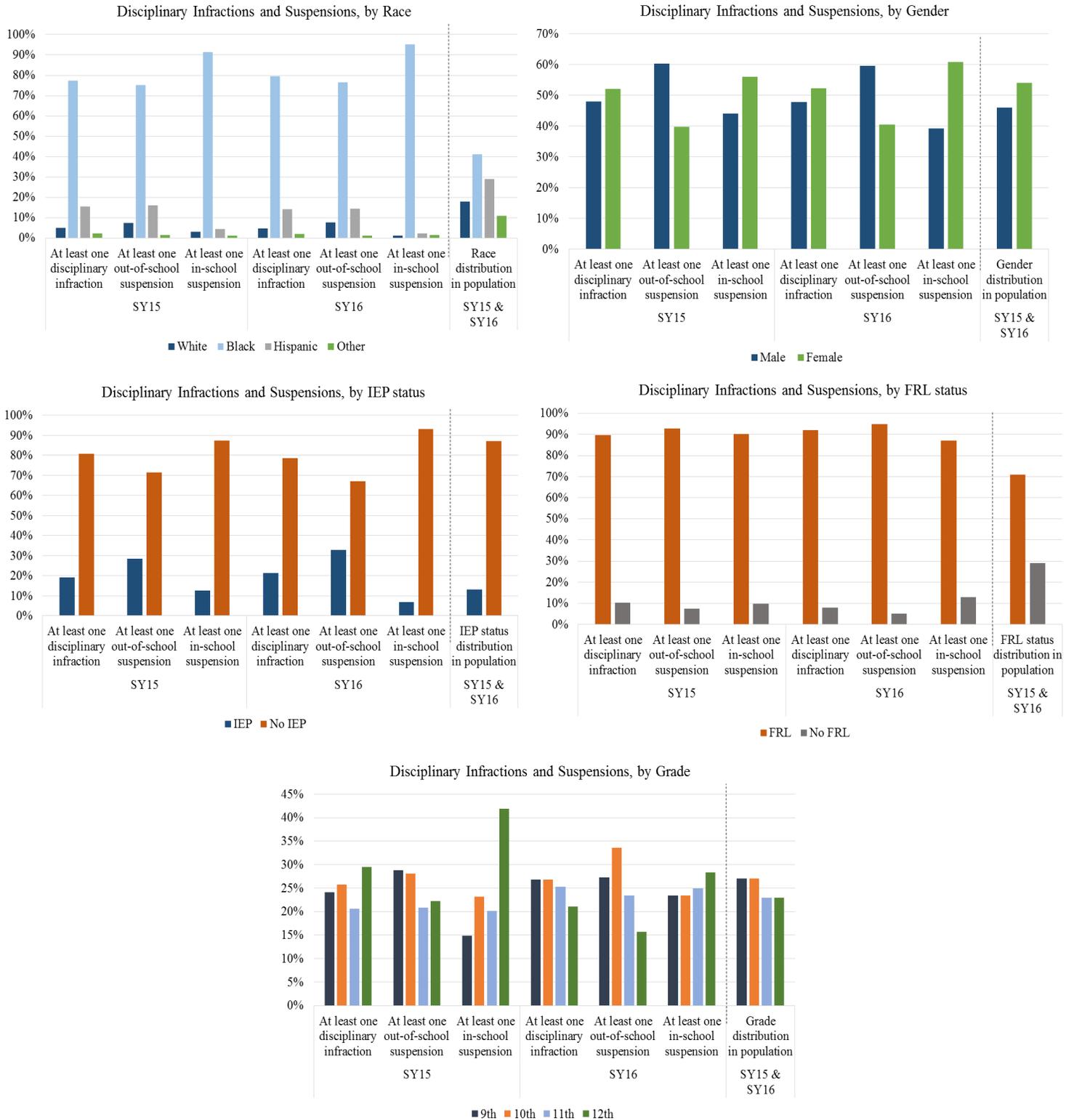
### **2014–15 Statistics**

According to these statistics, there were a total of 4,024 students who attended the four schools in 2014–15 (SY15). As shown in Figure 1, 21 percent of these students had at least one disciplinary infraction (or misconduct) at any level in that school year, and 10 percent of them had at least one infraction at levels 3 through 6.

We also looked at how different demographic groups of students compared in terms of receipt of disciplinary infractions and suspensions. The first panel of Figure 2 shows that, in SY15, African American students were overrepresented among students who had at least one disciplinary infraction. Specifically, while African Americans constituted 41% of the population, they made up 77% of those with at least one infraction. Students who had an IEP and students who received free or reduced-price lunch were also overrepresented among those with at least one infraction. However, there weren't notable differences in receipt of disciplinary infractions between male and female students, or between students at different grade levels.

Figure 2 also shows that African American students, students with an IEP, and students receiving free or reduced-price lunch were much more likely to receive at least one out-of-school suspension. In addition, African American students, those receiving free or reduced-price lunch and 12th graders were much more likely to receive at least one in-school suspension.

**Figure 2. Disciplinary Infractions and Suspensions, by Student Demographics**



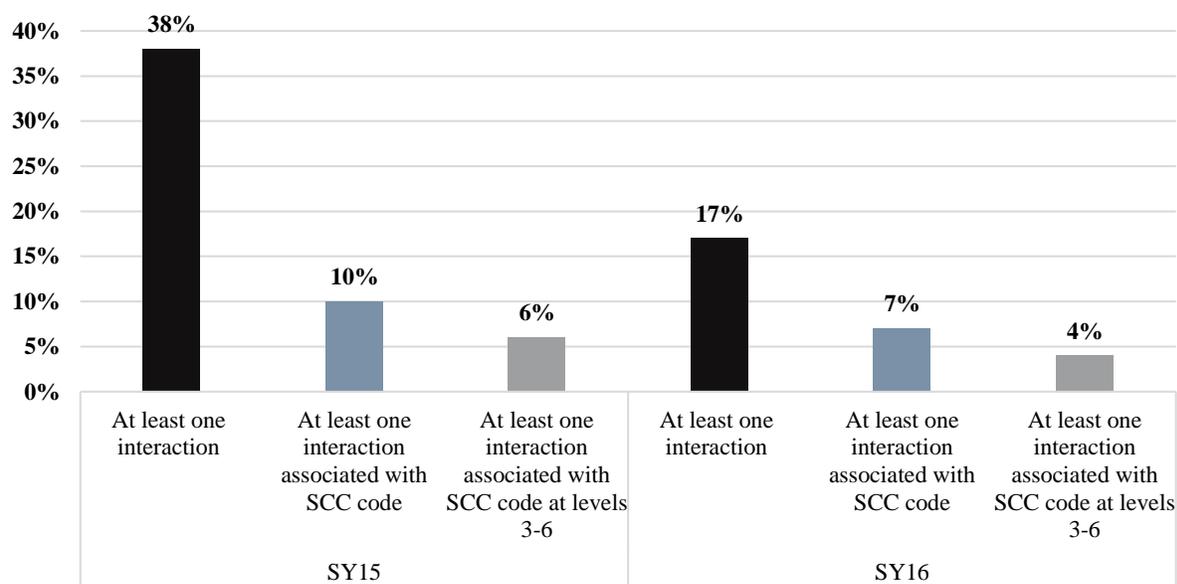
### 2015–16 Statistics

A total of 3,734 students attended the four schools in 2015–16 (SY16). Compared to SY15, a somewhat smaller percentage of these students had one or more disciplinary infractions, both overall (17% vs. 21%) and at levels 3 through 6 (7% vs. 10%). This could be because these students were less likely to engage in behavior that would be associated with an infraction, which could in turn be a result of maturation.

However, it is possible that lower levels of infractions could also be attributed in part to students' past participation in PR programming or to the fact that school staff were less likely to record infractions.

Figure 2 shows that, in general, the percentages of students with at least one disciplinary infraction or at least one suspension (both in- and out-of-school) declined across all demographic groups from SY15 to

**Figure 3. Percentage of Students who Interacted with PR Staff**



SY16. However, it continued to be the case that African American students, students with IEPs, and students receiving free or reduced-price lunch were much more likely to receive misconducts and suspensions.

We also looked at the percentages of students across the two years who showed up in the PR log sheets. As shown in Figure 3, in SY15, 38 percent of students had at least one interaction with PR staff (including self-referrals), 10 percent had at least one interaction that was associated with an SCC code,

and 6 percent had at least one interaction associated with an SCC code of 3 or higher.<sup>7</sup> However, in SY16, only 17 percent of students had at least one interaction with PR staff, although similar percentages of students in both years participated in PR programs due to an incident that either had an SCC code assigned or had an SCC code of 3 or higher assigned.

---

## Description of Methods and Study Sample

### Propensity Score Matching—Overview

The main challenge of any impact evaluation is to figure out what would have happened to individuals in the absence of their participation in a given program or intervention. However, since participants' outcomes are, by definition, only observed after they participate in the program, statistical or other methods have to be used to estimate what their outcomes would likely have been had they not participated in the program.<sup>8</sup> In the context of the current study, our primary task was to determine how many disciplinary infractions or suspensions would have been incurred by students who participated in PR programming *in the event the programming had not actually existed*. However, since the programming did exist, this means that our primary task was to find students similar to the students who participated in PR programming, but who did not have access to the programming.

Ideally, we would have used an experimental evaluation, in which we would have randomly assigned some of the students with disciplinary infractions to participate in PR programming, while the rest would not have had the opportunity to participate. For example, we might have tracked all students for the first several months of the school year and then randomly referred (e.g., by flipping a coin) half of the students who ended up with disciplinary infractions during that time period to the Peace Room. This would ensure that the only difference between the two groups of students—those referred and those not referred—was whether or not they utilized the PR. If assignments to programming were nonrandom, then participants and nonparticipants might differ in other characteristics that affect both participation in the program and their likelihood of incurring discipline infractions or suspensions. If we could have randomly assigned student to the two groups, a simple comparison of the outcomes of the two groups at the end of the school year and in the following school year would provide an accurate estimate of the effect of being referred to and utilizing PR programming.<sup>9</sup>

---

<sup>7</sup> Information on student participation in PR programming in the 2013–14 school year did not contain SCC codes.

<sup>8</sup> It is true that the variable the program is expected to impact (e.g. disciplinary infractions) could potentially be measured before and after individuals participate in the program. However, any positive change in the variable could not be attributed solely to participation in the program, since other unaccounted-for factors could potentially explain the change.

<sup>9</sup> In addition, we would have conducted an implementation study in order to determine the extent to which PR participants were receiving the intervention as planned.

Since a random assignment study was not feasible for this evaluation (given budget and other constraints), we had to select a nonexperimental method in order to evaluate the effect of participating in PR programming. We elected to use the method of propensity score matching (PSM), which involves comparing the outcomes of participants in a program or intervention to nonparticipants who are *observationally similar* to the participants (Heinrich, Maffioli, & Vazquez, 2010). Essentially, the method entails finding a nonparticipant who is “similar” to each participant, and then estimating the impact of the program as the difference in outcomes between the participant and his or her matched comparison. Then, averaging across all participants provides an estimate of the impact of the program on participants.

The method also involves the calculation of a “propensity score” (or probability of participation), which is defined as the probability that an individual in the combined sample of participants and nonparticipants participates in the program, given a set of observed variables. If all information relevant to participation and outcomes is observable to the researcher (i.e., there is data available to assess it), the propensity score will produce valid matches for estimating the impact of the program (thus, defining the variables on which the matches are to be made is of critical importance; we discuss the matching variables we selected below).

After deciding on the method of PSM, we then had to decide how to select the comparison students. There were two main possibilities: we could select the comparison students from other similar schools that did not have PR programming during the time period under study or we could select them from students who attended the four schools before they implemented PR programming. The advantage of the former choice is that the comparison is conducted between students attending school in the same time period, thus avoiding any influences of changes over time that might have impacted the likelihood of students incurring disciplinary infractions or being suspended. This is a potentially important consideration given that, as described above, CPS revised the SCC to reduce the length of suspensions in the fall of 2012 and adopted changes to the SCC in June 2014 to reduce the use of out-of-school suspensions.<sup>10</sup> On the other hand, the disadvantage of the first option is that, in order to draw valid conclusions, the schools selected as comparison schools would ideally be those that would have opted to adopt PR programming if they had been given the opportunity to do so. However, we believe it would be very difficult to identify these types of schools based on the limited amount of information generally available about schools, and that the biases arising from using this comparison sample would be larger than those arising from using the second option. Given this, we decided to draw our comparison sample

---

<sup>10</sup> Related to this, the Consortium on Chicago School Research found that, in 2009-10, 24 percent of high school students received an out-of-school suspension. However, this rate declined in each year after that through 2013-14. On the other hand, rates of in-school suspensions among African American high school students doubled during this time period.

from students who attended the focal schools before RJ programming started. We chose the 2010-11 school year in order to minimize the possibility of overlap with the students attending the schools after RJ programming began.

---

## **Analysis of Peace Room Interactions/Generation of Program Participants**

Having established a model for the underlying process by which students end up in the PR, and selected the method of propensity score matching to determine the effect of participating in PR programming on student outcomes, we next identified which students to consider as PR program participants (from among all those who had had any interaction with PR staff).

In order to do this, we determined that we needed to use the PR log sheets to establish which disciplinary infractions (incurred by which students) had resulted in referrals to the PR. In addition, we also determined that, to be sure that we were catching disciplinary infractions that were noted by both school and PR staff, we would need to try to match the disciplinary infractions identified in the PR log sheets with those identified in the CPS discipline data files. Accordingly, we restricted ourselves only to referrals that were made to the PR in SY15 or SY16, because these were the years for which PR staff recorded SCC codes. We then tried to match up code violations recorded in the CPS discipline file to code violations in the CBO data files, according to the date on which the incident occurred and the SCC codes that were assigned by PR and school staff.

For SY15, we found that there were 851 students who attended one of the four focal schools and who had one or more incidents recorded in the CPS discipline file. These 851 students incurred a total of 1,892 code violations (2.2 incidents per student). We also found that there were 570 students in the SY15 PR log sheet that had one or more disciplinary infractions (or code violations) with an SCC code assigned. In total, there were 1,122 of these code violations (2.0 incidents per student)

We then checked to see how many students showed up in both the CPS file and the PR log sheet; there were 398 such students. This means that 70 percent of the students in the SY15 PR log sheet (who had at least one disciplinary infraction that was assigned an SCC code by PR staff) also had at least one disciplinary infraction that had been noted by school staff. Therefore, 30 percent of the students had a code violation that was identified by the PR staff but was not recorded by the school, not reported to CPS central administration, or both. Additionally, this means that 453 students with code violations noted by school staff (or 53 percent of students with violations) were not referred to the PR or were referred but didn't go.<sup>11</sup>

---

<sup>11</sup> It is also possible that the students were referred but the SCC code was not recorded by PR staff.

It should be noted that the 398 students incurred a total of 1,710 incidents, or 4.3 incidents per student on average. Since the average number of incidents per student in the CPS file and PR log sheet (before combining) was about 2, this implies that the students who showed up in both files were more likely to incur disciplinary infractions.

Next, we compared the disciplinary infractions recorded by CPS and PR staff for students who showed up in both files. Of the 1,710 incidents generated by these students, only 217 (or 12.7 percent) of these incidents matched on date.<sup>12</sup> However, these incidents were generated by 190 unique students, implying that 48 percent of the students had an incident in the PR log sheet that matched, according to date, with an incident in the CPS file. On this basis, we decided to include these 190 students in the “treatment group” (the group of students considered to have participated in PR programming, where participation is defined according to the model previously described).

We also wanted to allow for the fact that some incidents might have inadvertently been recorded by PR staff a day after they occurred. Thus, we also checked to see how many incidents matched if their dates were allowed to differ by one day. Additionally, we imposed the restriction that the first part of the SCC code had to match.<sup>13</sup> We found that there were 25 students with incidents that matched on date within a day and also matched on the SCC code. We decided to include these students in the “treatment group” as well.

Next we repeated this process for SY16. We found that there were 623 students who attended one of the four focal schools and whom had one or more incidents recorded in the CPS discipline file. These 623 students incurred a total of 1,305 code violations (2.1 incidents per student). Thus, the number of students with disciplinary infractions in SY16 is lower than the number in SY15, but the number of infractions per student is similar.

We also found that there were 375 students in the SY16 PR log sheet that had one or more disciplinary infractions (or code violations) with an SCC code assigned. In total, there were 720 of these code violations (1.9 incidents per student). Thus, as with the CPS file, the number of students in the PR log sheet with disciplinary infractions in SY16 is lower than the number in SY15, but the number of infractions per student is similar.

---

<sup>12</sup> When we attempted to match on date and SCC code, we found an even smaller percentage of matches. We subsequently reviewed a random sample of cases where the dates matched but the SCC codes differed and determined that, based on the notes recorded in the PR logs, in the majority of cases the SCC codes in the PR and CPS files appeared to be referring to the same incident (e.g., both codes referred to a fight). Therefore, in order to obtain a large enough sample for analysis, we decided to retain students who had at least one infraction that matched on date, but not necessarily on SCC code.

<sup>13</sup> There are two parts to the CPS SCC codes. The first part identifies the level of severity of the infraction, and the second part identifies a specific behavior that falls within the level of severity. We matched only on the first part of the code.

As with the SY15 files, we then checked to see how many students showed up in both the CPS file and the PR log sheet; there were 250 such students. This means that 67 percent of the students in the SY16 PR log sheet (who had at least one disciplinary infraction that was assigned an SCC code by PR staff) also had at least one disciplinary infraction that had been noted by school staff. This is similar to the percentage we found for SY15 (70%). Additionally, these numbers indicate that 60 percent of the students with disciplinary infractions in the CPS file did not show up in the PR log sheet, which is a little higher than what we found for SY15.

The students who showed up in both the CPS file and the PR log sheet incurred a total of 1,026 incidents, or 4.1 incidents per student on average, which is similar to what we found for SY15. Also, we found that only 138 (or 13.5%) of these incidents matched on date, which is, again, similar to the percentage we found for SY15. These incidents were generated by 91 unique students, indicating that 36 percent of the students had an incident in the PR log sheet that matched on date with an incident in the CPS file (this is lower, by 12 percentage points, than what we found for SY15). These 91 students were also added to the treatment group.

As with SY15, we also checked to see how many incidents matched if SCC codes were the same but their dates were allowed to differ by one day. We found that there were 4 students with incidents that satisfied these criteria. We included these students in the treatment group as well. Overall, this process yielded a treatment group of 300 students across the two years.

### **Propensity Score Matching: Description of Process**

As mentioned above, a critically important part of carrying out PSM is the selection of the variables that will be used to match program participants (the treatment group) to similar nonparticipants. In selecting these variables, we were necessarily limited to using the variables available in the CPS data files. We first hypothesized that school staff might be less willing to refer students with the worst eighth grade behavior to the PR, perhaps on the assumption that these students would be less likely to benefit from RJ activities. Thus, we decided to match on eighth grade misconducts (total number of misconducts, number of level 1 and 2 misconducts, and number of level 3 through 6 misconducts) and suspensions (having at least one in-school suspension, and having at least one out-of-school suspension).<sup>14</sup> We also decided to match on eighth grade test scores and attendance, on the assumption that school staff might be more likely to refer students perceived as more able and/or engaged. For this same reason, we also matched on an indicator of whether a student was old for grade. Additionally, because the descriptive statistics above showed

---

<sup>14</sup> We measured these variables in eighth grade in order to make sure we were capturing characteristics prior to when students would have had any exposure to the Peace Room.

differences in behavioral and disciplinary outcomes by race/ethnicity, gender, grade level, Individualized Education Plan (IEP) status, and free/reduced lunch status we included these variables as well. Finally, we included a variable indicating the school attended by the student, on the assumption that the implementation of RJ programming might vary according to differences in school climate, effectiveness of school leadership and other factors across schools.

Additionally, we decided to look at the effect of utilizing the PR on student outcomes in two different years: the year in which students participated in PR programming and the year subsequent to the year in which they participated. We looked at the latter on the assumption, for example, that skills learned in the PR might persist and affect disciplinary outcomes in future years. We also decided to examine six different outcomes in each year: (1) the total number of misconducts incurred, at any level; (2) the number of level 3 through 6 misconducts; (3) the likelihood of having at least one out-of-school suspension; (4) the total number of days spent in out-of-school suspensions; (5) the likelihood of having at least one in-school suspension; and (6) the total number of days spent in in-school suspensions. In addition, we examined the likelihood of having at least one misconduct at any level and the likelihood of having at least one misconduct at levels 3 through 6 in the subsequent year.<sup>15</sup>

We used a two-step process to obtain estimates of the association between participation in PR programming and the aforementioned outcomes. In the first stage, we estimated the relationship between participating in PR programming and each of the matching variables. These relationships were estimated using the sample of students who attended the focal schools in SY15 and SY16 and who had at least one recorded disciplinary infraction. Some of these students were identified, in the process described in the section above, as having utilized the PR but the remainder were not. The estimates from this first stage were used to compute a propensity score (i.e., the probability of PR participation) for the SY15/SY16 treatment group and the sample of students who attended the focal schools in 2010–11 and who had at least one disciplinary infraction. The second stage estimation then used “nearest neighbor” PSM to match students on the metric of the estimated probability of PR participation.<sup>16</sup> The resulting sample of treatment and comparison students was then used to estimate the effect of PR participation on the selected outcomes.

---

<sup>15</sup> By definition, all of the students in the sample had at least one misconduct at any level in the first year.

<sup>16</sup> In nearest neighbor PSM, the student from the comparison group with the propensity score closest to a given student from the treatment group is chosen as the match for the latter student.

---

## Findings

### Characteristics of the Study Samples

Table 1 shows the characteristics of five different samples of students who attended the schools with which the CBO partnered. The first column contains all students who attended the four focal schools during the 2010–11 (SY11) school year. The second contains all students who attended the focal schools during SY15 and SY16. The third column is a subset of students in the first column, and is restricted to those who had one or more disciplinary infractions in SY11. This is the group of students from which we drew our comparison sample. The fourth column is a subset of the students in the second column, and contains those who had one or more disciplinary infractions in SY15 and SY16. Finally, column five is a subset of column four and contains the students who we identified as being participants in PR programming, as defined by our model (the “treatment” group). Comparing columns one and two reveals that students who attended the focal schools in SY11 and in SYs 15 and 16 were very similar to each other in terms of demographics. In particular, the percentages of students from each race/ethnic group were virtually the same, as were the percentages that were female and possessed an IEP. Also, students attending the schools in SY15 and SY16 were only somewhat more likely to receive free or reduced-price lunch than students in SY11. On the other hand, students attending the schools in SY11 were more likely to be old for grade.

In terms of disciplinary outcomes, students attending the schools in SY15 and SY16 had lower levels of misconducts and suspensions. In particular, they were substantially less likely to have at least one OSS and had significantly lower total OSS days. As mentioned previously, these differences are likely to be, at least in part, a reflection of changes in disciplinary practices in the schools as a result of changes in CPS policies. This, again, is a limitation of using students attending the schools in earlier years as the comparison group.

Comparing columns 2 and 4 shows that students who attended the focal schools in SY11 and had one or more disciplinary infractions were more likely to be African American, less likely to be white or Asian/Pacific Islander, less likely to be female, and less likely to be in 12th grade than the overall population of students in that year. They were also more likely to have an IEP, to receive free or reduced-price lunch, to be old for their grade, and to be in tenth grade. Many of these differences are mirrored when comparing all students who attended the schools in SYs 15 and 16 with students attending in those years with one or more infractions (columns 2 and 4), although in many cases they are even more pronounced. For example, compared to all students who attended the schools in SYs 15 and 16, those who had one or more infractions were much more likely to be African American and to qualify for free or reduced-price lunch. They were also much less likely to be white or Hispanic. Notably, however, there was little difference between the two groups in terms of the percent who were female.

**Table 1. Characteristics of Students Attending Focal Schools**

|   | Full Sample | SY11  | SY15-16 | SY11 with >=1 infr. | SY15-16 with >=1 Infr. | SY15-16 in PR |
|---|-------------|-------|---------|---------------------|------------------------|---------------|
| Participated in PR programming                          | 2%          | 0%    | 4%      | 0%                  | 19%                    | 100%          |
| White   | 18%         | 18%   | 18%     | 13%                 | 5%                     | -             |
| Black   | 40%         | 38%   | 41%     | 55%                 | 78%                    | 74%           |
| Native American   | -           | -     | -       | -                   | -                      | -             |
| Hispanic  | 29%         | 30%   | 29%     | 26%                 | 15%                    | 19%           |
| Multi-racial  | 3%          | 3%    | 3%      | 3%                  | 1%                     | -             |
| Asian/Pacific islander                                  | 9%          | 11%   | 8%      | 3%                  | 1%                     | -             |
| Hawaiian/Pacific Islander                               | -           | -     | -       | -                   | -                      | -             |
| Female  | 54%         | 53%   | 54%     | 40%                 | 52%                    | 56%           |
| Has IEP   | 13%         | 11%   | 13%     | 19%                 | 20%                    | 25%           |
| Receives free/reduced-price lunch                       | 70%         | 67%   | 71%     | 76%                 | 91%                    | 94%           |
| Old for grade   | 16%         | 21%   | 14%     | 28%                 | 19%                    | 22%           |
| Attendance rate, prior SY                               | 90%         | 87%   | 91%     | 82%                 | 87%                    | 88%           |
| Attendance rate, 8th Grade                              | 95%         | 95%   | 95%     | 93%                 | 94%                    | 93%           |
| In 9th grade  | 27%         | 26%   | 27%     | 27%                 | 25%                    | 32%           |
| In 10th grade   | 28%         | 30%   | 27%     | 40%                 | 26%                    | 29%           |
| In 11th grade   | 23%         | 24%   | 23%     | 22%                 | 22%                    | 18%           |
| In 12th grade   | 22%         | 20%   | 23%     | 11%                 | 26%                    | 22%           |
| Attends Sullivan in Current Year                        | 19%         | 23%   | 16%     | 31%                 | 22%                    | 45%           |
| Attends Tilden in Current Year                          | 11%         | 14%   | 9%      | 16%                 | 19%                    | 21%           |
| Attends Lincoln Park in Current Year                    | 60%         | 63%   | 58%     | 53%                 | 16%                    | 8%            |
| Attends SSICP in Current Year                           | 11%         | 0%    | 16%     | 0%                  | 42%                    | 27%           |
| Number of schools attended, prior SY                    | 1.04        | 1.03  | 1.04    | 1.04                | 1.09                   | 1.09          |
| Number of misconducts, prior SY                         | 0.8         | 1.5   | 0.5     | 3.2                 | 1.5                    | 1.6           |
| Number of level 3-6 misconducts, prior SY               | 0.3         | 0.4   | 0.2     | 0.9                 | 0.7                    | 0.8           |
| Has at least one level 3-6 misconduct, prior SY         | 12%         | 17%   | 10%     | 35%                 | 29%                    | 36%           |
| Number of level 1-2 misconducts, prior SY               | 0.5         | 1.1   | 0.3     | 2.3                 | 0.9                    | 0.9           |
| Has at least one level 1-2 misconduct, prior SY         | 15%         | 21%   | 12%     | 40%                 | 37%                    | 36%           |
| Number of schools attended, 8th Grade                   | 1.02        | 1.03  | 1.02    | 1.05                | 1.05                   | 1.05          |
| Number of misconducts, 8th Grade                        | 0.2         | 0.4   | 0.2     | 0.9                 | 0.5                    | 0.7           |
| Number of level 3-6 misconducts, 8th Grade              | 0.2         | 0.3   | 0.1     | 0.6                 | 0.4                    | 0.5           |
| Has at least one level 3-6 misconduct, 8th Grade        | 9%          | 12%   | 9%      | 24%                 | 20%                    | 26%           |
| Number of level 1-2 misconducts, 8th Grade              | 0.1         | 0.1   | 0.1     | 0.3                 | 0.2                    | 0.3           |
| Has at least one level 1-2 misconduct, 8th Grade        | 5%          | 5%    | 5%      | 11%                 | 10%                    | 18%           |
| Has at least one misconduct, current SY                 | 23%         | 31%   | 19%     | 100%                | 100%                   | 100%          |
| Has at least one misconduct, following SY               | 22%         | 27%   | 16%     | 51%                 | 49%                    | 55%           |
| Number of misconducts, current SY                       | 0.6         | 0.9   | 0.4     | 2.9                 | 2.2                    | 3.1           |
| Number of misconducts, following SY                     | 0.5         | 0.7   | 0.3     | 1.6                 | 1.2                    | 1.6           |
| Has at least one level 3-6 misconduct, current SY       | 11%         | 15%   | 9%      | 48%                 | 46%                    | 60%           |
| Has at least one level 3-6 misconduct, following SY     | 11%         | 15%   | 7%      | 32%                 | 22%                    | 31%           |
| Number of level 3-6 misconducts, current SY             | 0.2         | 0.3   | 0.1     | 0.9                 | 0.8                    | 1.2           |
| Number of level 3-6 misconducts, following SY           | 0.2         | 0.3   | 0.1     | 0.7                 | 0.4                    | 0.7           |
| Has at least one out of school suspension, current SY   | 12%         | 22%   | 6%      | 72%                 | 34%                    | 43%           |
| Has at least one out of school suspension, following SY | 11%         | 16%   | 6%      | 34%                 | 19%                    | 24%           |
| Out of school suspension days, current SY               | 0.6         | 1.4   | 0.2     | 4.4                 | 1.1                    | 1.6           |
| Out of school suspension days, following SY             | 0.6         | 1     | 0.2     | 2.3                 | 0.6                    | 0.9           |
| Has at least one in school suspension, current SY       | 9%          | 11%   | 9%      | 34%                 | 46%                    | 25%           |
| Has at least one in school suspension, following SY     | 12%         | 17%   | 7%      | 32%                 | 24%                    | 24%           |
| In school suspension days, current SY                   | 0.2         | 0.3   | 0.2     | 1.1                 | 0.9                    | 0.6           |
| In school suspension days, following SY                 | 0.3         | 0.4   | 0.2     | 0.8                 | 0.5                    | 0.5           |
| N   | 11,244      | 3,575 | 7,669   | 1,115               | 1,461                  | 279           |

Note: We use "-" to indicate cells for which there are fewer than 10 students (except in cases where there are zero by definition).

Comparing columns 5 and 6 shows that students whom we designated as PR participants were very similar in terms of demographics to students who attended the schools in SYs 15 and 16 and who had one or more infractions. However, they were much more likely to attend Sullivan High School, and less likely to attend Lincoln Park or SSICP. Also, they were more likely to have at least one level 3 through 6 misconduct (60% vs 46%), and somewhat more likely to have at least one OSS, though less likely to have at least one ISS (25% vs 46%).

Finally, comparing columns 4 and 6 reveals that students who attended the focal schools in SY11 and had one or more infractions were more likely than PR program participants to be white or Hispanic, and less likely to be African American or female. They were also less likely to receive free or reduced-price lunch or to be in 12th grade, and more likely to be in tenth grade. Also, the distribution of students across schools is different, in part because SSICP did not open until the 2011–12 school year. Additionally, in terms of eighth grade characteristics and outcomes, the 2010–11 students had higher standardized test scores, were less likely to have a level 1 and 2 misconduct, and somewhat less likely to have been suspended.<sup>17</sup> These results suggest that the demographics and eighth grade characteristics are potentially important variables to account for when attempting to find matches between the PR participants and students with one or more disciplinary infractions who attended the schools in 2010–11.

### **Results of PSM Analysis**

Figure 4 shows the results of the first stage of the PSM analysis, in which we examine the extent to which the matching variables we selected are related to the likelihood that a student utilized the PR.

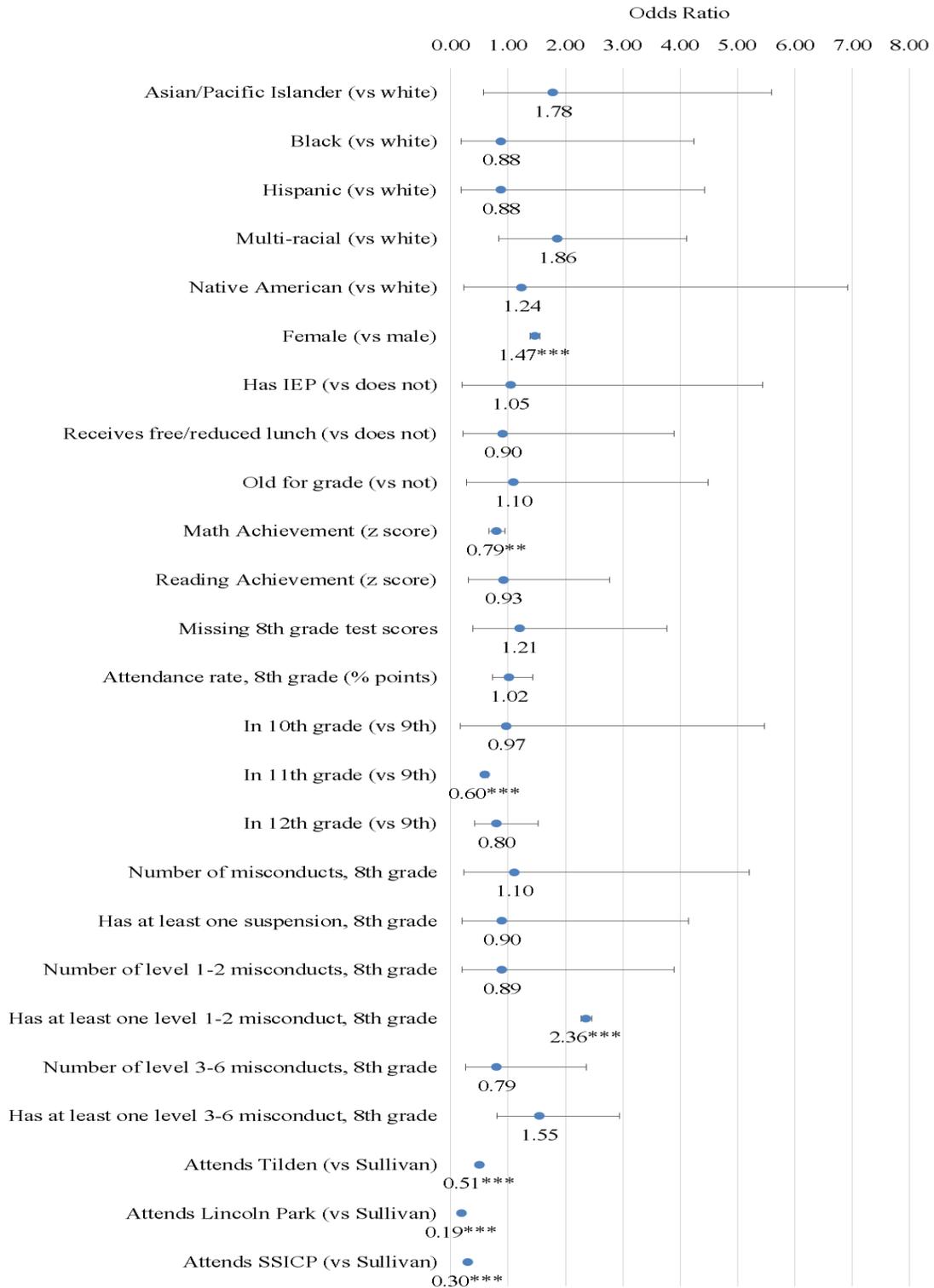
The results are presented as odds ratios. Odds ratios can be thought of as relative measures of effect that allow comparison of the treatment group to the comparison group. In particular, if the outcome is the same in each group, then the odds ratio will be equal to one. If the outcome is more likely in relation to the treatment group, then the odds ratio will be greater than one. Similarly if the outcome is less likely in relation to the treatment group, the odds ratio will be less than one.

The results in Figure 4 show that, controlling for other characteristics, female students and students who had at least one level 1 and 2 misconduct in eighth grade were significantly more likely to utilize the PR. On the other hand, students with higher math scores in eighth grade, and students currently enrolled in 11th grade (as opposed to ninth) were significantly less likely to utilize the PR. Additionally, we found significant differences between the focal schools in the likelihood that a student participated in PR programming. Specifically, relative to students attending Sullivan High School, students in all three of the other schools were significantly less likely to participate in PR programming.

---

<sup>17</sup> The test scores have been standardized to have a mean of zero and a standard deviation of one.

**Figure 4. Propensity Estimation: Odds Ratios for Predictors of Likelihood of Participating in Peace Room Programming**



\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

**Table 2. Estimated Differences between Peace Room Participants and Matched Comparisons**

| Outcome   | Avg for Treated Sample | Avg for Matched Controls | Difference | Std. Error of Difference | p-value |
|---|------------------------|--------------------------|------------|--------------------------|---------|
| Has at least one misconduct, following SY               | 55.0%                  | 44.8%                    | 10.2%      | 0.07                     | 0.17    |
| Number of misconducts, current SY                       | 3.13                   | 3.69                     | -0.56      | 0.38                     | 0.15    |
| Number of misconducts, following SY                     | 1.60                   | 1.56                     | 0.04       | 0.35                     | 0.91    |
| Has at least one level 3-6 misconduct, following SY     | 30.7%                  | 36.3%                    | -5.6%      | 0.07                     | 0.41    |
| Number of level 3-6 misconducts, current SY             | 1.21                   | 1.13                     | 0.08       | 0.17                     | 0.65    |
| Number of level 3-6 misconducts, following SY           | 0.69                   | 0.77                     | -0.07      | 0.18                     | 0.68    |
| Has at least one out of school suspension, current SY   | 43.0%                  | 73.2%                    | -30.2%     | 0.05                     | 0.00    |
| Has at least one out of school suspension, following SY | 23.6%                  | 36.5%                    | -13.0%     | 0.07                     | 0.06    |
| Out of school suspension days, current SY               | 1.60                   | 5.12                     | -3.52      | 0.77                     | 0.00    |
| Out of school suspension days, following SY             | 0.86                   | 2.26                     | -1.41      | 0.50                     | 0.00    |
| Has at least one in school suspension, current SY       | 25.4%                  | 38.9%                    | -13.5%     | 0.05                     | 0.01    |
| Has at least one in school suspension, following SY     | 23.6%                  | 23.6%                    | 0.0%       | 0.07                     | 1.00    |
| In school suspension days, current SY                   | 0.56                   | 1.76                     | -1.21      | 0.30                     | 0.00    |
| In school suspension days, following SY                 | 0.51                   | 0.90                     | -0.39      | 0.27                     | 0.15    |

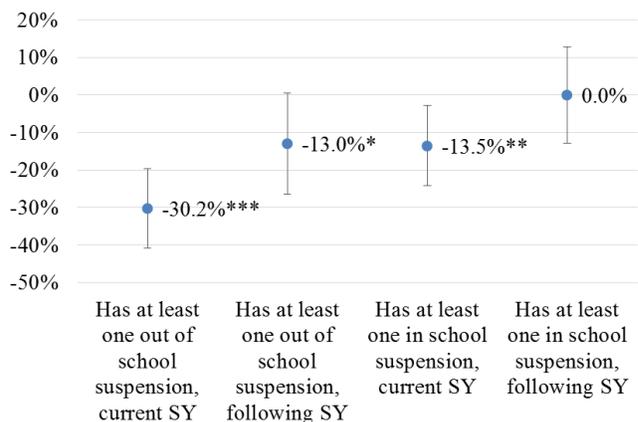
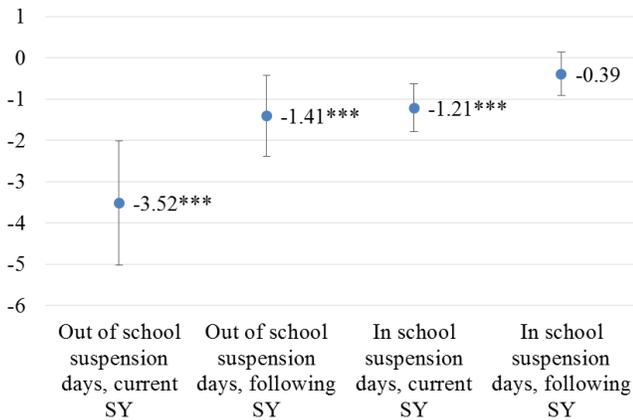
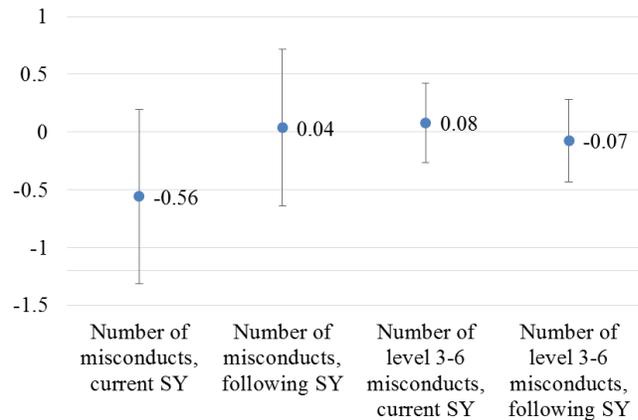
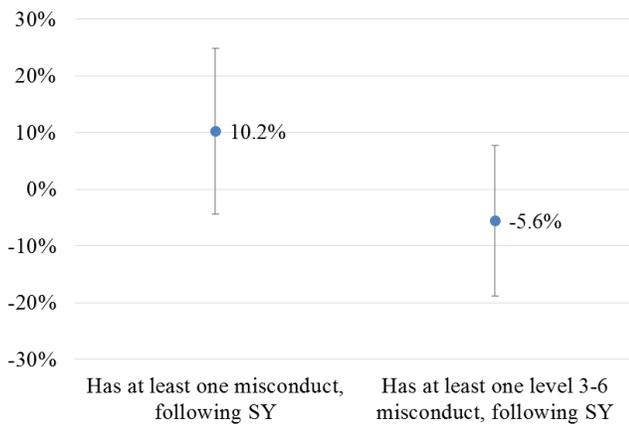


Table 2 shows the results of the second stage of the PSM analysis, in which we examine the estimated effect of participating in PR programming on outcomes in the current and subsequent years. Looking first at the effects of participating in PR programming on misconducts in the year of participation, the results in Table 2 show no statistically significant differences between PR participants and their matched comparisons in either the number of misconducts at any level or the number of misconducts at levels 3 through 6.

In terms of suspensions, Table 2 shows that the average out-of-school suspension rate among students who had one or more misconducts and participated in PR programming was about 30 percentage points lower than that for similar students who attended the focal schools in 2010–11. Also, students who had one or more misconducts and utilized the PR experienced about 4 fewer days of OSS. However, as noted previously, given the changes in CPS policies regarding OSS during the time period of this study, it is difficult to separate out effects due to the policy changes from those due to utilizing the PR. Therefore, it is likely that these results overestimate the effect of participating in PR programming. In addition, as noted above, there were many students with misconducts that were recorded by school staff who either were not referred to the PR or did not show up in the PR logs. If school staff were more likely to refer students who would not have received a suspension (i.e. in the absence of the PR), this would also result in an overestimate of the effect of participating in PR programming.

On the other hand, Table 2 also shows that the average in-school suspension (ISS) rate among students who had one or more misconducts and participated in PR programming was about 14 percentage points less than the rate for similar students who attended the focal schools in 2010–11. Additionally, these students experienced a little over 1 fewer ISS days on average. Given that overall ISS rates among high school students increased during the study period, these estimates are less likely to overestimate the effects of participating in PR programming.

Looking next at the difference between PR participants and their matched comparisons in misconducts in the year following the year of participation, we did not find any statistically significant differences. Thus, our hypothesis that students might learn skills from utilizing the PR that would lessen the likelihood of their incurring misconducts in subsequent years was not supported. Similarly, the estimated differences in the in-school suspension rate and days of ISS were also smaller and statistically insignificant. We did find that differences in the OSS rate and days of OSS continued to be statistically significant. Specifically, the following year OSS rate among students who had one or more misconducts and participated in PR programming was estimated to be 13 percentage points lower than the rate for similar students who attended the focal schools in 2010–11 (though it was only significant at the 10% level). Also, students who had one or more misconducts and participated in PR programming experienced about 1.5 fewer days

of OSS. However, as noted above, given CPS policy changes, these differences are likely to be overestimates of the true effect.

Finally, we looked at the effects of participating in PR programming for different subgroups of students. Overall, we found some evidence that females tended to benefit more than males, Hispanics tended to benefit more than African Americans, ninth graders benefited more than older students, and students with IEPs benefited more than those without IEPs. Specifically, we found that:

- The differences between PR participants and comparison students in the average OSS rate in the current year were larger for African American females, Hispanic males, and Hispanic females versus African American males; ninth graders versus all other grades; and students with an IEP versus those without.
- The differences between PR participants and comparison students in the average OSS rate in the following year were larger for females versus males, Hispanics versus African Americans, and students with an IEP versus those without.
- The differences between PR participants and comparison students in the average ISS rate in the current year was larger for females versus males, Hispanics versus African Americans, Hispanic females versus Hispanic males, ninth graders versus all other grades, students receiving free or reduced-price lunch versus those who were not, and students with an IEP versus those without.

In addition, while the main results did not show statistically significant differences overall between PR participants and comparison students in the number of misconducts in the current year, we did find significant benefits for African Americans versus Hispanics, African American males versus all other race/gender groups, students receiving free or reduced-price lunch versus those not, and students with an IEP versus those without. Similarly, while the main results did not show significant differences in the ISS rate in the following year, we did find significant differences for Hispanics versus African Americans, and students with an IEP versus those without.

---

## Summary

Although RJ practices have been adopted in schools worldwide over the past two decades, there are relatively few studies that have attempted to rigorously assess the impact of these practices on student and school outcomes. To help remedy this gap in knowledge, we employed the method of propensity score matching (PSM) to examine the effects of participation in a key component of RJ programming offered by a local community-based organization on student behavioral and disciplinary outcomes in four high schools in Chicago.

Our analysis is based on an underlying model in which it is assumed that school staff members observe student disciplinary infractions and then make a decision about whether to refer students to the Peace Room. We also hypothesized that students who are referred to the PR are less likely to receive a suspension as a consequence, and that these students may also learn skills from their interactions in the PR that make it less likely that they will engage in future behavior that results in a disciplinary infraction or suspension. In accordance with this model, we identified a sample of students who could be considered to be participants in PR programming. These students (the treatment group) were selected from the sample of students who had one or more disciplinary infractions recorded in the CPS discipline file, as well as incidents with SCC codes recorded in the PR log sheet, and consisted of students who had at least one incident that matched either on date or within a day and SCC code.

We used the method of PSM to compare these students to similar students who attended the focal schools prior to the start of PR programming, and we looked for effects both in the year that students participated in PR programming and in the year after they participated. Our results showed, first, that a significantly lower percentage of students who participated in PR programming experienced at least one out-of-school suspension, and that they had significantly fewer days of OSS. These results were true for the year in which students utilized the PR, as well as in the following year. We also found that a significantly lower percentage of the students who participated in PR programming experienced at least one in-school suspension. In addition, they had about one fewer day of ISS. This was true for the year in which students utilized the PR, but not in the following year.

Finally, we found that the benefits of utilizing the PR, with respect to rates of ISS, were: (1) larger for females than for males, (2) larger for Hispanics than for African Americans, (3) larger for Hispanic females than for Hispanic males, (4) larger for ninth graders than for students in other grades; (5) larger for students receiving free or reduced-price lunch than for those not receiving this benefit; and (6) larger for students with an IEP than for students without an IEP.

---

## **Limitations and Recommendations for Future Research**

We should note several limitations of this study. First, as mentioned previously, PSM can only account for *observable* preparticipation differences between program participants and nonparticipants. Thus, after the matching was completed, there still could have been *unobservable* differences between the participants and nonparticipants that worked in the favor of (or against) the participants. Related to this point, our treatment sample is drawn from students who had infractions recorded in both the CPS administrative data and the PR log sheets. However, over half of the students with infractions in the CPS data did not show up in the PR logs. This might be because they were 1) not referred by school staff to the PR; 2) referred but did not go to the PR; or 3) referred and utilized the PR, but for some reason PR staff

did not or were unable to make note of their presence. To the extent that the first explanation is true, it could be that teachers and administrators tend to refer students to the PR who they think are better suited to resolving differences through dialogue. If their hunches are correct, then the estimates we obtained would overestimate the positive effect of participating in PR programming. On the other hand, to the extent that the third explanation is true, our estimates would underestimate the effect of participation. However, without additional information on the process by which students use the PR, we cannot assess the extent of bias in our results.

Additionally, we know that CPS began changing its approach to student discipline around the time that the CBO began working in the schools and after the comparison group students attended the focal schools. Thus, some of the estimated effects that we found are likely to be due in part to changes in the ways the schools recorded and dealt with disciplinary fractions as a result of the changes in CPS policy, particularly where out-of-school suspensions are concerned. Also, the analysis assumes that the only difference between the schools in SY11 and the schools in SYs 15 and 16 was the introduction of restorative justice practices. However, there could have been other changes in the approach to discipline and behavior management that impacted misconducts and rates of suspension.

Finally, we were not able to examine the quality of implementation of the PR programming, which is a critical aspect of the potential impact of any program. Related to this, we did not include information on the culture and climate of the schools, both of which could undoubtedly influence both student outcomes and the level and quality of implementation of programming.

Overall, however, this study provides promising evidence for the positive effects of some RJ practices on student disciplinary outcomes, and suggests that the positive effects may be larger in some cases for students who are at greater risk for poorer outcomes.

We recommend that future studies focus on the implementation, as well as the impacts, of RJ practices in schools. In addition to examining the relationship between outcomes and the quality of implementation, it would be helpful to have a clearer understanding of the processes by which specific students utilize RJ practices, including how school administrators and other school staff make decisions about which students to refer to staff implementing RJ programs. Additionally, this study only considered the effects, at the individual student level, of participation in one component of the CBO's RJ programming. Future research might also examine the entire range of restorative practices implemented in schools, including their school- as well as student-level impacts.

---

# References

- Armstrong, M. (2015). *Classroom circles: Building positive relationships*. Retrieved from <http://www.rpiassn.org>.
- Cameron, L. & Thorsborne, M. (2001). Restorative justice and school discipline: Mutually exclusive? In H. Strang & J. Braithwaite (Eds). *Restorative justice and civil society* (pp. 180–195). Cambridge, England: Cambridge University Press.
- Chicago Public Schools, Office of College and Career Success. (2014, June). *Chicago Public Schools student code of conduct (SCC) revisions*. Retrieved from [https://www.cpsboe.org/content/documents/student\\_code\\_of\\_conduct\\_revisions\\_06.2014.pdf](https://www.cpsboe.org/content/documents/student_code_of_conduct_revisions_06.2014.pdf).
- 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.
- Heinrich, C., Maffioli, A. & Vazquez, G. (2010) *A primer for applying propensity-score matching: Impact-evaluation guidelines*. Washington, DC: Inter-American Development Bank.
- High HOPES Campaign. (2012). *From policy to standard practice: Restorative justice in Chicago Public Schools*. Retrieved from <http://www.dignityinschools.org/policy-standard-practice-restorative-justice-chicago-public-schools>.
- Levenstein, R., Spote, S., & Allensworth, E. (2011). *Findings from an investigation into the Culture of Calm initiative*. Chicago, IL: University of Chicago Consortium on School Research.
- Mader, N., Sartain, L., & Steinberg, M. P. (2017). When suspensions are shorter: The effects on school climate and student learning. Working paper.

- McCluskey, G., Lloyd, G., Kane, J., Riddell, S., Stead, J., & Weedon, E. (2008). Can restorative practices in schools make a difference? *Educational Review*, 60(4), 405–17.
- Ortega, L., Lyubansky, M., Nettles, S., & Espelage, D. L. (2016). Outcomes of a restorative circles program in a high school setting. *Psychology of Violence*, 6(3), 459–468.
- Powell, K. E., Muir-McClain, L., & Halasyamani, L. (1995). A review of selected school-based conflict resolution and peer mediation projects. *The Journal of School Health*, 65(10), 426-431.
- Stevens, W. D., Sartain, L., Allensworth, E. M., & Levenstein, R. (2015). *Discipline practices in Chicago schools: Trends in the use of suspensions and arrests*. Chicago, IL: University of Chicago Consortium on School Research.
- Stinchcomb, J. B., Bazemore, G., & Riestenberg, N. (2006). Beyond zero tolerance: Restoring justice in secondary schools. *Youth Violence and Juvenile Justice*, 4(2), 123–147.
- Van Ness, D. & Strong, K. H. (1997). *Restoring justice*. Cincinnati, OH: Anderson Publishing Company.
- Wong, D. S. W., Cheng, C. H. K., Ngan, R. M. H., & Ma, S. K. (2011). Program effectiveness of a restorative whole-school approach for tackling school bullying in Hong Kong. *International Journal of Offender Therapy and Comparative Criminology*, 55(6), 846–862.
- Youth Justice Board for England and Wales. (2004). *National Evaluation of the Restorative Justice in Schools Programme*. Retrieved from <https://restorativejustice.org.uk/resources/national-evaluation-restorative-justice-schools-programme>.