

Restorative Interventions and School Discipline Sanctions in a Large Urban School District

Yolanda Anyon

University of Denver

Anne Gregory

Rutgers, The State University of New Jersey

Susan Stone

University of California, Berkeley

Jordan Farrar

Jeffrey Jenson

Jeanette McQueen

University of Denver

Barbara Downing

Eldridge Greer

John Simmons

Denver Public Schools

A large urban district (N = 90,546 students, n = 180 schools) implemented restorative interventions as a response to school discipline incidents. Findings from multilevel modeling of student discipline records (n = 9,921) revealed that youth from groups that tend to be overrepresented in suspensions and expulsions (e.g., Black, Latino, and Native American youth; boys; and students in special education) had similar, if not greater, rates of participation in restorative interventions than their peers. First-semester participants in restorative interventions had lower odds of receiving office discipline referrals (OR .21, $p < .001$) and suspensions (OR .07, $p < .001$) in the second semester. However, the suspension gap between Black and White students persisted. Implications for reform in school discipline practices are noted.

KEYWORDS: office discipline referral, restorative intervention, restorative justice, school discipline, suspension

A collective challenge to conventional wisdom about school discipline has been issued at local, state, and federal levels. No longer is it assumed that suspension should remain the “go to” response to student misconduct

Anyon et al.

and school safety concerns. A growing body of evidence indicates that exclusionary discipline practices, such as out-of-school suspension (OSS) and expulsion, are not effective or equitable approaches to improving student behavior and school safety (American Academy of Pediatrics, 2013). School officials also recognize that aggregated discipline rates obscure disparities between student groups. Many educators are now scrutinizing their

YOLANDA ANYON, PhD, is an assistant professor in the Graduate School of Social Work at the University of Denver, 2148 S. High Street, Denver, CO 80208, USA; e-mail: yanyon@du.edu Her research draws on mixed methods and critical theories to consider the roles of schools and community-based organizations in shaping the life outcomes of low-income youth of color.

ANNE GREGORY, PhD, is an associate professor at the Graduate School of Applied and Professional Psychology at Rutgers University. Her research has focused on the persistent trend that African American adolescents are issued school suspension and expulsion at higher rates than adolescents from other groups. Through program development and evaluation, she aims to improve educational settings for students from diverse racial and ethnic backgrounds.

SUSAN STONE, PhD, is the Catherine Mary and Eileen Clare Hutto Professor of Social Services in Public Education. Her work is focused on social work in education and its impact on the academic progress of vulnerable youth in schools.

JORDAN FARRAR, MSW, is a doctoral candidate at the Graduate School of Social Work at the University of Denver. Her research interests include youth involved in armed conflict, restorative justice and indigenous healing practices, and global social work research.

JEFF JENSON, PhD, is the Philip D. & Eleanor G. Winn Professor for Children and Youth at Risk at the Graduate School of Social Work at the University of Denver. His teaching and research interests address the etiology and prevention of childhood and adolescent problems of bullying, aggression, school dropout, and juvenile delinquency.

JEANETTE McQUEEN, MSW, is an assistant professor and director of the online MSW program at the Graduate School of Social Work at the University of Denver. She has worked in a variety of clinical settings with children, youth, and families, including community-based services, residential treatment, and public schools.

ELDRIDGE GREER, PhD, is the associate chief of student equity and opportunity at Denver Public Schools, where his focus is removing structural barriers that negatively impact educational opportunity. He has been nationally recognized for his work in Denver for discipline reform efforts, with a particular focus on eliminating the racial disparities in discipline.

BARBARA DOWNING, PhD, is a district partner in the Office of Social-Emotional Learning at Denver Public Schools. She is responsible for monitoring the implementation of district discipline policy, improving student information system technology, and training and facilitation of school staff.

JOHN SIMMONS, MA, is the former associate chief of the Division of Student Services at Denver Public Schools. He oversaw division decision-making on resource allocation, policy, professional development opportunities, and assessment and monitoring systems.

data to detect disproportionality along the lines of student race, gender, and special education status (e.g., Public Schools of North Carolina, 2013).

Indeed, there is consensus among many researchers, policy makers, educators, and school-based mental health professionals that exclusionary school discipline practices rarely improve school safety and, in fact, exacerbate racial inequalities in education and incarceration. In 2014 alone, national reports about school discipline were issued from the U.S. Departments of Education and Justice (U.S. Department of Education, 2014), the Council of State Governments Justice Center (Morgan, Salomon, Plotkin, & Cohen, 2014), and the Discipline Disparities Research to Practice Collaborative (Carter, Fine, & Russell, 2014). The reports converge in their recommendations to reduce suspension through alternative practices that have a greater chance of changing student behavior, keeping youth in school, and maintaining a positive school climate.

These recent calls for change reflect evidence that suspension has deleterious effects on student well-being and school safety. Schools with high rates of suspensions, expulsions, and law enforcement referrals are perceived by students, teachers, and parents to be less safe than other schools (Osher, Poirier, Jarjoura, & Brown, 2014; Steinberg, Allensworth, & Johnson, 2014). Moreover, youth who have been suspended or expelled are more likely than students who do not receive disciplinary sanctions to be pushed out of school and into criminal justice systems; this process is often referred to as the “school-to-prison pipeline” (Fabelo et al., 2011; Rausch, Skiba, & Simmons, 2004; Skiba et al., 2014). For instance, a longitudinal study of Florida ninth graders found that each suspension decreased students’ odds of graduating high school by an additional 20% and decreased their odds of enrolling in postsecondary schooling by 12% (Balfanz, Byrnes, & Fox, 2015). Moreover, a Texas statewide study found that students suspended or expelled for a discretionary school violation were about three times more likely than other youth to have contact with the juvenile justice system in the next school year (Fabelo et al., 2011).

Patterns of dropout and juvenile justice involvement are of particular concern given racial disparities in exclusionary school discipline outcomes. Latino, Native American, and Black youth are significantly more likely than students of other backgrounds to be referred to school administrators for discipline problems and to receive OSS, expulsion, or a referral to law enforcement as punishment (Hannon, DeFina, & Bruch, 2013; Payne & Welch, 2010). These students tend to be disciplined more harshly than White students for the same type and number of offenses (Anyon et al., 2014; Bradshaw, Mitchell, O’Brennan, & Leaf, 2010) but are less likely to have access to much needed support services (Reyes, Elias, Parker, & Rosenblatt, 2013). The interlocking nature of the discipline, achievement, and incarceration gaps suggests that, over the long term, whole groups of students who are disproportionately suspended and have lower

Anyon et al.

achievement are less likely to obtain a range of positive life outcomes (Gregory, Skiba, & Noguera, 2010).

Concerns about equity and the detrimental effects of suspension have driven educators to seek alternatives to traditional suspension practices and policies. The U.S. Departments of Education (DOE, 2014) and Justice recommend that students should not only be held accountable for conduct but should also have opportunities to learn from discipline incidents and build social and emotional skills. They note that schools may decide to use restorative interventions (RIs) to enhance and teach a range of individual skills. Similarly, the Council of State Governments Justice Center (Morgan et al., 2014) suggests that after conflict or rule infractions, educators should utilize a “restorative follow-up.” The follow-up, they explain, provides students with opportunities to discuss incidents, accept responsibility for harmful actions, and identify ways to repair harm. Recommendations from these reports reflect a wave of initiatives sweeping the United States (following the lead of many other countries such as New Zealand, Australia, and England) in which schools have implemented RIs as alternatives to suspension (Drewery, 2013; McCluskey et al., 2008; Schiff, 2013). Despite reduced suspension rates in individual schools and descriptive accounts of improvements in districts that implement RIs (e.g., Encarnacao, 2013; Karp & Breslin, 2001; Schiff, 2013; Stinchcomb, Bazemore, & Riestenberg, 2006), multivariate analyses of implementation data that account for between-school variability and the nature of student offenses are rare and have not been published in peer-reviewed journals.

RIs

Arising from a humanist philosophy and with historical roots in a range of diverse cultures (e.g., Native American, Maori) and religions (e.g., Judaism), restorative approaches assume that subjective experiences of harmful acts need to be acknowledged and that it is worthwhile to harness the power of the collective for resolution and repair (Drewery, 2013; Zehr, 2002; Zehr & Toews, 2004). Restorative approaches to school discipline include a variety of practices on the prevention-intervention continuum. Namely, some practices aim to prevent infractions and other practices intervene after infractions have occurred (e.g., Amstutz & Mullet, 2005; Blood & Thorsborne, 2005; McCluskey et al., 2008; Wachtel, Costello, & Wachtel, 2009). At the intervention end of the continuum, restorative approaches have two core features: (a) Those affected by an infraction or crime come together to identify how people were impacted by the incident, and (b) they jointly problem-solve and identify actions that will repair the harm (Coates, Umbreit, & Vos, 2003; McGarrell & Hipple, 2007).

In essence, RIs are problem-solving processes held in a small conference or a larger circle format, which may include people affected by the incident

directly and indirectly. Typically, conferences for serious incidents follow a formal procedure. First, a preconference meeting is held whereby a facilitator meets with a disputant to orient him or her to restorative approaches. At this meeting, a disputant can decline to participate in an RI or a facilitator can determine a conference is not appropriate if the disputant will not accept any responsibility or acknowledge his or her role in the incident and/or is not willing to repair the harm (Wachtel, O'Connell, & Wachtel, 2010). Second, if the conference is to proceed, a range of parties are invited to voluntarily attend, including the disputant, the disputant's supporters, and all those negatively impacted by the incident (McCluskey et al., 2008).

Third, in the conference itself, participants sit in a circle facing one another, and a facilitator uses a structured set of questions to guide the exchange among all the participants. The goal is for everyone (including the victim and the disputant) to voice their perspectives. The set of questions facilitate reflection on the link between actions and subsequent consequences. Typical questions include the following: "What happened?"; "Who has been harmed/affected by what you have done?"; "What part are you responsible for?"; and "How will the harm be repaired?" (Teachers Unite, 2014). Questions also solicit sharing of the emotional experience of the incident to further empathy and understanding (Nathanson, 1997; Wachtel et al., 2010).

Fourth, the participants jointly develop a plan to repair the harm and prevent future incidents. The aim is to hold disputants accountable for breaching trust with the community and at the same time reintegrate those students back into the community (Braithwaite, 1989, 2001; Costello, Wachtel, & Wachtel, 2010). Agreements to repair the harm can take many forms, including the disputant making amends through his or her actions (e.g., community service or repair damaged property). Typically, agreements are written down and agreed upon by all conference participants.

RIs and positive outcomes. Most prior research on restorative practices has examined school-wide reductions in office discipline referrals (ODRs) and OSSs using single group, pre-, and posttest designs (Schiff, 2013). These studies lack comparison groups and seldom use any statistical controls to account for potential confounders. That said, numerous international studies have reported reductions in school-wide ODRs and OSS rates after restorative practices were introduced, including in New Zealand (Buckley & Maxwell, 2007), Scotland (Kane et al., 2007), and China (Wong & Mok, 2011). Studies of restorative practices in the United States have shown similar declines, including in Denver, Colorado (González, 2015), Minneapolis, Minnesota (Riestenberg, 2013), and Philadelphia, Pennsylvania (Lewis, 2009). In addition, in Oakland, the suspension rate of Black students declined at a sharper rate than the suspension rate of White students after the introduction of a range of interventions including restorative approaches

to discipline, Schoolwide Positive Behavioral Intervention Supports, and the Manhood Development Program (Jain, Bassey, Brown, & Kalra, 2014). In the Denver Public Schools (DPS) during years when RIs initially spread to schools throughout the district, González (2015) reported a 4 percentage point decrease in the Black/White suspension gap. These trends are only suggestive of the promise for restorative approaches to reduce exclusionary practices and narrow the racial discipline gap.

Findings from experimental research addressing student outcomes resulting from school-based conferences have yet to be published. However, research on adult and youth restorative conferencing in the criminal justice system suggests that similar programming in school settings may be beneficial. Randomized controlled trials in the United States, Australia, and Great Britain, in which juvenile offenders were assigned to restorative conferences, other diversion programs, or typical court procedures, have found that some restorative conference programs have the ability to reduce rates of reoffense, whereas other programs have no long-term effects on reoffending (Larsen, 2014; Latimer, Dowden, & Muise, 2005). Specifically, Strang, Sherman, Mayo-Wilson, Woods, and Ariel (2013) analyzed results from 10 controlled trials on three continents with a total of 1,879 offenders and 734 interviewed victims. Their results showed that among cases in which both offenders and victims were willing to meet, restorative justice conferences reduced future crime. That said, using propensity score matching with samples of youth offenders in Australia, Smith and Weatherburn (2012) revealed no difference in future offense rates between youth who participated in a conference and those who participated in a business-as-usual condition. Further, although a controlled trial in Indianapolis found that participants randomly assigned to conferencing or to other diversion programming experienced short-term benefits in terms of reduced rates of reoffense after 2 years (McGarrell & Hipple 2007), the benefits were not sustained in a 10-year follow-up (Jeong, McGarrell, & Hipple, 2012). The researchers conclude that conferences may result in short-term (not long-term) reduction in risk.

The experimental literature from juvenile justice suggests that research on both distal and proximal outcomes of school-based RIs are needed. Moreover, research on the impact of RIs in educational settings is warranted because of the unique dynamics of school environments compared to community systems. For example, it is possible that the impact of RIs on youth outcomes could be stronger than what has been found from criminal justice studies. In schools, it is likely that RIs are used with a much more diverse group of young people with lower risk profiles than community-based offenders and could also lead to changes in relationships between youth, their peers, and the school adults with whom they interact on a more consistent basis than police officers or crime victims (Anyon, 2016).

School and student participation in RIs. Schools often vary tremendously in their adoption of new initiatives (Forman, 2015). One measure of program diffusion is the degree to which an intervention is used by practitioners (Durlak & Dupre, 2008; Schulte, Easton, & Parker, 2009). Several studies have shown that schools' use of restorative approaches to discipline can differ throughout a district, which may weaken their impact on student outcomes (Jain et al., 2014; McClusky et al., 2008). Lower use of RIs in response to discipline incidents can indicate practical barriers such as lack of training or staffing, poor alignment between a restorative philosophy and the norms or values of school personnel, and/or limited opportunities for practitioners to improve their skills (Anyon, 2016; Durlak & Dupre, 2008). Yet no studies have examined the relationship between school-level rates of RI use and individual student outcomes after receiving an RI.

Moreover, given consistent evidence that schools contribute to sorting and labeling students (Weinstein, 2002), a concern about the implementation of alternatives to suspension would be that participation in RIs would reflect typical dynamics related to power and privilege. This would parallel consistent patterns in the child welfare and juvenile justice systems whereby more disadvantaged youth and their families receive harsher consequences, and fewer support services, than their privileged counterparts (Chapin Hall Center for Children, 2008; Derezotes, Testa, & Poertner, 2005). In fact, using data from a national survey of principals in the late 1990s, Payne and Welch (2010) found that the proportion of Black students at the school level was negatively associated with the site's reported use of restorative practices in response to student misbehavior. Thus, there is a need for new research using actual discipline records to assess whether students from disadvantaged groups similarly participate in RIs when they receive a discipline referral compared to more advantaged student groups. Equitable participation is especially needed for Black, Latino, Native American, and male students as well as students in special education—all groups of students who tend to receive suspension at higher rates than more advantaged peers (Losen & Martinez, 2013).

The Current Study

RIs in school settings appear to be a promising response to discipline problems. However, to date, few studies have analyzed RI implementation results using statistical approaches that account for the hierarchical nature of these datasets, in which students are nested within schools, or control for confounds like the type of student offense (Schiff, 2013). The current study controls for a range of covariates, most notably students' socioeconomic status and their number/type of ODRs, to assess the relationship between participation in RIs and adverse student discipline outcomes during the spring semester. The study builds on prior analyses conducted in the

Anyon et al.

same school district (Anyon et al., 2014) by using time-ordered data from a subsequent school year linked to school-level rates of RI use. Specifically, current analyses examine whether receiving RIs in the first semester is associated with lower odds of disciplined students receiving additional ODRs or OSSs in the second semester. Scholars have argued that students from disenfranchised groups could benefit the most from RIs focused on building relationships, soliciting student voice, promoting an ethic of care, and reintegrating students back into the school community (Drewery, 2004). Therefore, an exploratory analysis also examines whether the association between RI participation and discipline is moderated by student racial background (Losen & Martinez, 2013).

The study also builds on prior findings indicating that the implementation of restorative programming varies widely across schools (Jain et al., 2014; McClusky et al., 2008). To this end, it examines whether school-level variation in RI use is associated with student-level discipline outcomes (Schulte et al., 2009). We postulate that the relationship between student participation in RIs and subsequent discipline outcomes will be stronger in schools that use RIs more often. Prior theory and research guides this hypothesis: Relative to schools relying more on exclusionary discipline and less on RIs, schools with higher rates of RI participation may reflect staff members' commitment to, preparation for, and/or skill in implementing high-quality RIs (Cross et al., 2011[AQ: 1]; Forman, 2015). Through an equity lens, the study also considers whether marginalized and disadvantaged groups have similar patterns of participation in RIs. To our knowledge, no studies have compared the sociodemographic characteristics of students who have or have not participated in RIs.

In summary, the following questions guide this study: (1a) Is a student's participation in one or more RIs in the first semester associated with lower odds of ODRs and/or OSSs in the second semester? (1b) Is the association between participation in RIs and later discipline incidents moderated by student racial background or school-level use of RI? (2) Do disciplined students from disadvantaged backgrounds have equitable participation in RIs?

Method

School District and Study Participants

Study site. The study site for this investigation is DPS (hereafter referred to as "the District"). The District is uniquely situated as a site to examine the influence of RIs on school discipline outcomes. First, following a major discipline policy reform in 2008 that aimed to reduce the use of exclusionary discipline sanctions, increase alternative approaches such as RIs, and eliminate racial disparities in suspension and expulsion, the District has witnessed

sustained reductions in rates of OSS, expulsion, and law enforcement referral (Anyon et al., 2014). These results are impressive because they have taken place during a time when the overall District population has increased by 14%, making the District among the fastest growing urban school districts in the nation (Department of Planning and Analysis, 2013). Second, despite these successes, District data indicate that reform goals have not yet been fully realized and that disparities in race, class, gender, and special education status persist in school discipline outcomes (Anyon et al., 2014). These trends prompted District leaders' interest in evaluating the impact of RIs on ODRs and OSSs to assess whether additional resources should be invested in this approach.

In the 2012–2013 school year, the District served a student population ($N = 90,546$) that was 57.31% Latino, 20.83% White, 14.47% Black, 3.35% Asian, 2.99% multiracial, 0.81% Native American, and 0.24% Pacific Islander. Forty-nine percent were female, and 51% were male. Close to half (44%) of District students were English Language Learners (ELLs). The District serves predominantly low-income students, as over two thirds of the students in the District were eligible for free and reduced lunch (68.8%) and 2.2% of students were identified as being homeless. In addition, 11.5% of District students participated in special education, and 1.1% were identified as having an emotional disability (ED).

School district discipline reform. After overhauling the District's school discipline policy reform in 2008, school officials began offering voluntary staff training in RIs. The training was, and continues to be, available to any employee of the District. Staff members can choose to sign up to participate via an online registration system where all professional development opportunities are listed. During monthly meetings with school-based staff, District leaders strongly recommended that principals, disciplinarians, teachers, and special service providers (such as social workers and psychologists) participate in the training. Two types of training are provided to staff. The first is an introductory training that is 4 hours long and focuses on preventive RIs (e.g., classroom community-building circles). Relevant to the current study, the second training is 2 days long and emphasizes RIs in response to discipline incidents. The following content is covered in the 2-day trainings: (a) overview of the origins and key principles of RIs (including their use in response to concerns about racial disparities in OSS), (b) review of empirical evidence of the effectiveness of restorative approaches in the District and beyond, (c) RIs as they relate to District discipline policy and schools' student codes of conduct, (d) brief introduction to prevention-oriented restorative practices (dialogues and proactive circles), (e) lengthy introduction to intervention-oriented restorative practices (reactive circles, mediations, and conferences), (f) overview of core features of all restorative practices (e.g., problem solving, paraphrasing, reframing), and (g) strategies

to monitor the implementation and success of restorative approaches. Each content area is supplemented with interactive role-play scenarios and case studies (DPS, 2012).

At the end of the training, participants are provided with a handbook that details all content from the training, including example codes of conduct and forms for implementing restorative approaches (e.g., agreements, action plans, parent letters, and evaluative surveys). On-site coaching and support from the district coordinator is available following the training. Since August 2008, more than 2,700 district educators have participated in the 2-day training. In the 2012–2013 school year (the focus of the current study), 126 staff members (37 teachers, 28 administrators, and 61 support service providers) represented 53 District schools at the trainings.

District policy strongly recommends that students be offered a RI for behavior that leads to a discipline action. Restorative conferencing is an option for Type 2 (e.g., severe defiance of authority/disobedience) through Type 5 (e.g., first degree assault) infractions (DPS, 2008). The policy suggests RIs may be provided independently (e.g., RI only), as alternatives to each other (e.g., RI or 1-day suspension), or in conjunction with each other (e.g., RI and in-school-suspension) (DPS, 2008). These decisions are made by school administrators and vary depending on their site's specific code of conduct. Therefore, it is not known if RIs were offered as options in a similar manner to students at all schools. This limitation is common in school discipline research because most discipline policies rely on the discretion of administrators in determining consequences, which are often inconsistently implemented (Hannon et al., 2013; Morris, 2005; Shaw & Braden, 1990).

If an administrator does decide to incorporate a RI as part of resolving a discipline incident, District protocol is that the student then meets with the trained staff member. If the student is willing to “take responsibility for his or her part of the situation” after reflecting on the incident, a restorative circle, mediation, or conference with all affected parties is held (DPS, 2012, p. 13). If not, the student is referred back to the school administrator for a different consequence. In the framework of a tiered system of support (RtI), circles, mediations, and conferences are considered targeted (Tier 2) and intensive (Tier 3) interventions, as opposed to Tier 1 universal supports (Berkowitz, 2012; Corrigan, 2012). Restorative conferences in the District typically involve those directly involved in the conflict (typically a two-party dispute). Reactive circles include individuals indirectly affected by an incident—an incident can indirectly affect others through disruptions to instruction or community well-being (González, 2015). At the end of a RI, participants develop an agreement or action plan for “making things right,” and all involved parties sign the agreement.

Student sample. The student sample included all youth ($n = 9,921$) in Grades K–12 across all District schools ($n = 180$ schools) who were issued

one or more ODRs in the 2012–2013 school year (see Table 1). These disciplined students comprised 11% of all youth in the District. Mirroring trends observed by other researchers, disciplined students were disproportionately Black, Latino, Native American, male, low-income, eligible for special education, and classified as ED. Findings from chi-square tests shown in Table 1 indicate that subgroup differences in discipline rates were statistically significant. For example, Black students comprised 14.5% of the general student population versus 25.2% of the population with at least one ODR. Students who identified as Asian, Pacific Islander, White, or ELL were issued ODRs at significantly lower rates than their enrollment. Suspension rates had similar patterns of significant over- and underrepresentation of varying student groups. Table 1 also indicates that RIs were most often utilized with students who were referred to the office for midrange offenses such as disobedience or defiance and detrimental behavior.

Table 2 presents disproportionality figures for ODRs and OSSs for all racial groups in the District that parallel the patterns evident in Table 1. Risk indices capture rates of suspension and referrals for all student racial groups in the district. They were calculated by dividing the number of one group of students who have been referred or suspended by the number of that same group in the population of the district (Skiba et al., 2008). Relative risk ratios (RRRs) were calculated as the ratio of the risk indexes of two groups (IDEA Data Center, 2014; Shaw, Putnam-Hornstein, Magruder, & Needell, 2008). In other words, the RRR is a ratio of ODR or suspension rates per 1,000 between two groups of students (Shaw et al., 2008). For example, in the case of ODRs, Black students had a risk index of 19.02% and RRRs of 3.41 compared to White youth and 1.99 compared to all other students. Among Latinos, the ODR rate was 11.22%, and the RRRs were 2.00 compared to White students and 1.06 compared to all other students. For OSSs, Black students had a suspension rate of 9.64% and the RRRs were 4.95 compared to White students and 2.55 compared to all other students. In contrast, the suspension rate for Latino students was 4.46%, whereas the RRRs were 2.29 compared to White students and .92 compared to all other students.

Measures

Sociodemographic and discipline records were downloaded from the District's student information system (Infinite Campus). Downloaded datasets included school-level characteristics (e.g., enrollment size), student background information, and student-level discipline records.

Student discipline records. The District's discipline database included information for each ODR issued to a student in 2012–2013; this information included the reason for each referral and related consequences. These data are entered by a school staff member trained to indicate the reason for the

Table 1

Sample Characteristics of Students Who Received One or More ODRs Across One School Year

	All Students (<i>N</i> = 90,546)	Disciplined Students ^{a,b} (<i>n</i> = 9,921)	Received an OSS ^c (<i>n</i> = 4,184)	Participated in RI ^d (<i>n</i> = 1,277)	Did Not Participate in RI ^e (<i>n</i> = 8,644)
Student sociodemographics (%)					
Race					
Native American (<i>n</i> = 735)	0.8	0.99*	1.0	.8	1.02
Asian (<i>n</i> = 3,036)	3.4	1.5***	1.3	1.2	1.5
Black (<i>n</i> = 13,098)	14.5	25.2***	30.2***	26.9	24.9
Latino (<i>n</i> = 51,893)	57.3	58.7***	55.3***	59.2	58.6
White (<i>n</i> = 18,858)	20.8	10.6***	9.0***	9.6	10.8
Pacific Islander (<i>n</i> = 216)	0.2	0.14*	0.1	0.1	0.2
Multiracial (<i>n</i> = 2,710)	3.0	2.9	3.1	2.2	3.0
Boys (<i>n</i> = 46,235)	51.1	65.8***	66.9***	67.3	65.6
Eligible for free/reduced lunch (<i>n</i> = 62,321)	68.8	84.8***	87.6***	85.4	84.7
ELLs (<i>n</i> = 39,871)	44.0	37.8***	33.5***	34.0**	38.4**
Students in special education (<i>n</i> = 10,422)	11.5	20.28***	25.7***	21.5	20.1
ED (<i>n</i> = 954)	1.1	4.0***	7.2***	4.0	4.1
Grade level					
Elementary (<i>n</i> = 35,916)	39.7	22.5***	15.9***	17.2***	23.33***
Middle (<i>n</i> = 12,924)	14.3	28.2***	28.2***	32.0***	27.6***
High (<i>n</i> = 19,034)	21.0	26.2***	25.9	20.2***	27.1***
Alternative grade configuration (<i>n</i> = 22,632)	25.0	23.1***	23.8	30.7***	22.0***
Referral reasons ^{e,f}					
Bullying (<i>n</i> = 823)		8.3	8.3	11.0***	7.9***
Destruction of school property (<i>n</i> = 200)		2.0	3.1***	2.4	2.0
Disobedient/defiant (<i>n</i> = 3,212)		32.4	34.6***	37.7***	31.6***
Other code of conduct violation (<i>n</i> = 2,474)		24.9	26.5**	32.6***	23.8***
Detrimental behavior (<i>n</i> = 5,415)		54.6	63.0***	69.0***	52.5***

(continued)

Table 1 (continued)

	All Students (<i>N</i> = 90,546)	Disciplined Students ^{a,b} (<i>n</i> = 9,921)	Received an OSS ^c (<i>n</i> = 4,184)	Participated in RI ^d (<i>n</i> = 1,277)	Did Not Participate in RI ^e (<i>n</i> = 8,644)
Third degree assault (<i>n</i> = 247)	2.5	<i>M</i> (<i>SD</i>)	4.4***	3.8**	2.3**
Unlawful sexual behavior (<i>n</i> = 23)	1.4		1.7	1.1	1.4
Drug possession or distribution (<i>n</i> = 730)	7.4		13.2***	4.5***	7.8***
Dangerous weapon (<i>n</i> = 173)	1.7		3.3***	2.1	1.7
Number of discipline consequences ^f		<i>M</i> (<i>SD</i>) ^g		<i>M</i> (<i>SD</i>) ^h	<i>M</i> (<i>SD</i>) ⁱ
OSS	.68 (1.12)		1.60 (1.25) ^{n/a}	.79 (1.31)**	.66 (1.10)**
In-school suspension	.61 (1.18)		.69 (1.45)***	.79 (1.41)***	.59 (1.14)***
Behavior contract	.03 (.20)		.07 (.30)***	.08 (.33)***	.02 (.18)***
RI	.17 (.52)		.18 (.56)	1.30 (.79) ^{n/a}	0 (0) ^{n/a}
Referred to law enforcement	.08 (.32)		.17 (.00)***	.11 (.01)***	.08 (.00)***
Expulsion	.01 (.09)		.02 (.00)***	.02 (.00)**	.01 (.00)***

^aDisciplined students received one or more ODRs during the 2012–2013 school year.

^bSignificant differences were determined by chi-square tests, compared with all students in the District who did not receive one or more office referrals.

^cSignificant differences were determined by chi-square tests, compared with all disciplined students not suspended.

^dSignificant differences were determined by chi-square tests, compared with all disciplined students who did not receive a RI.

^eSignificant differences were determined by chi-square tests, compared with all disciplined students who did receive a RI.

^fPercentages of referral reasons do not add up to 100 because 42% of students received more than one ODR over the course of a school year. ^gDue to space limitations, only the nine most common reasons for referral are presented in the table. The complete data for all referral reasons are available upon request from the authors. ^hAQ: 21.

ⁱSignificant differences were determined by independent-samples *t* tests, comparing students who did not receive an OSS to those who did.

^jSignificant differences were determined by independent-samples *t* tests, comparing students who did not receive a RI to those who did.

^kDiscipline consequences are not mutually exclusive. For one discipline incident, a student may receive multiple consequences as part of the resolution to his or her offense. Forty-eight percent of students also have more than one discipline incident, so they also can receive a consequence more than once.

p* < .05. *p* < .01. ****p* < .001.

Table 2
Risk Indices^a and R^b for Office Disciplinary Referrals and OSS

Student Race	Office Disciplinary Referrals All Year			OSS All Year		
	Risk Index	RRR Compared to White Students	RRR Compared to All Other Students	Risk Index	RRR Compared to White Students	RRR Compared to All Other Students
Black	19.08%	3.41	1.99	9.64%	4.95	2.55
White	5.60%	1.00	0.45	1.95%	1.00	0.37
Latino	11.22%	2.00	1.06	4.46%	2.29	0.92
Asian	4.74%	0.85	0.42	1.81%	0.93	0.38
Multiracial	10.55%	1.88	0.96	4.80%	2.46	1.04
Native American	13.33%	2.38	1.22	5.85%	3.01	1.27
Pacific Islander	6.48%	1.16	0.59	2.31%	1.19	0.50

^aThe risk index is the proportion of students from one racial group who have been suspended. It is computed by dividing the number of students suspended from one group by the total number of students from that group (Skiba et al., 2008).

^bThe risk ratio is computed by taking a ratio of the rates per 1,000 between two groups (IDEA Data Center, 2014; Shaw et al., 2008).

referral based on the categories in the District’s discipline policy (DPS, 2008). ODRs reflect a multistep process whereby educators assess student misconduct, complete the formal discipline referral paperwork, and record it in the school’s database. Despite educators’ varying approaches to addressing student behavior or misconduct, and their varying use of the formal discipline process, research has established that ODRs are consistent correlates of teachers’ perceptions of problematic behavior, poor teacher–student relationships, future misconduct, and future academic difficulties (e.g., Pas, Bradshaw, & Mitchell, 2011). For example, students’ receipt of one or more ODRs is associated with negative teacher behavioral ratings (Pas et al., 2011) and, years later, with being off-track for graduation (Tobin & Sugai, 1999)—evidence for the concurrent and predictive validity of ODRs.

The ODR referral categories represent the total number of times a student was referred to the office for each particular ODR reason (0 = student was not referred to the office for this reason, 1 or more = number of referrals a student received for this behavior). This coding method was necessary because almost half of disciplined students (42%) were referred to the office more than once during the school year, often for different behaviors. Therefore, ODR reasons at the student level are not mutually exclusive; as a result, the reference group for each referral category is all other reasons.

The nine most common reasons for an ODR (comprising 98% of all discipline incidents) were detrimental behavior (54.58% of disciplined students), disobedience or defiance (32.38%), other violations of the school's code of conduct (24.94%), bullying (8.30%), possession and/or distribution of drugs (7.36%), third degree assault (2.49%), destruction of school property (2.02%), unlawful sexual behavior (1.38%), and weapons possession (1.74%). An additional seven reasons for referral were included in the dataset but were infrequently issued (e.g., alcohol violation [.97%], robbery [0.60%], tobacco violation [0.50%], gang affiliation [0.32%], first degree assault [.23%], witness intimidation [0.14%], other felonies [0.14%]) (see Table 1).

For each student ODR, the District's discipline database indicated the type of consequences assigned by the administrator (one discipline incident may have multiple consequences). As shown in Table 1, among all disciplined students (defined as students who experienced an ODR at least once during the academic year), 13% received RIs and 43% received OSSs. In the current study, we distinguished between consequences received in the first or second semester. To answer Research Question 1, we created a dichotomous dependent variable to indicate whether, in the second semester, a student received one or more ODRs (1) or none at all (0). We also indicated whether a student in the second semester received an OSS (1) or not (0)—a group that included those students who never received another ODR. All analyses conducted to answer Research Question 1 included an independent variable that accounted for the total number of RIs a student received in the first semester (among those who participated in an RI; $M = 1.30$, $SD = .79$, minimum = 0; maximum = 8). To answer Research Question 2 regarding the equitable access to RIs, we created a dichotomous dependent variable to indicate whether, in the first semester as a consequence for an ODR, a student received one or more RIs (1) or no RIs (0). RIs included circles, mediations, and conferences; the dataset did not distinguish between these three practices.

Student characteristics. The school District's record system included sociodemographic information for each student issued a discipline referral in 2012–2013. Student racial categories used by the District are as follows: (1) Native American or Alaska Native, (2) Native Hawaiian or other Pacific Islander, (3) Asian or Asian American, (4) Black or African American (non-Hispanic), (5) Hispanic or Latino/Latina, (6) White or Caucasian, and (7) multiracial. Each racial category was recoded into dummy variables with White students as the reference group. Additional student-level variables available in the dataset were all dichotomous and included gender (male or not), free and reduced lunch eligibility (eligible or not), special education status (active Individualized Education Program or not), designation as seriously emotionally disabled (ED or not), and ELL or not.

School characteristics. School-level covariates were selected based on prior research consistently linking them to discipline outcomes (e.g., Arcia, 2007; Payne & Welch, 2010; Skiba et al., 2014), including proportion of the student body that was Black and the proportion eligible for free and reduced lunch, along with grade configuration and school size (divided by 100). Grade configuration was dummy coded with K–5 elementary schools as the reference group relative to middle schools, high schools, and schools with alternative combinations of grade levels (e.g., grades K–8). We also calculated school-wide use of RI by dividing the number of students who received RIs in each school by the number of students with ODRs from that site. The resulting RI rate ranged across schools from 0% to 75% ($M = 8.31\%$, $SD = 13.10\%$).

Data Analytic Plan

Multilevel logistic regression methods were used to assess study outcomes.¹ Using STATA 13 software, hierarchical modeling techniques accounted for the nested structure of the dataset (Level 1 = students; Level 2 = schools) and were used to estimate the relationships between (a) student sociodemographic characteristics, (b) participation in RIs during the first semester of the school year, and (c) discipline outcomes in the second semester (Rabe-Hesketh & Skrondal, 2008; Steenbergen & Jones, 2002). For research Questions 1a–c, we ran a set of models predicting the dependent variables (one or more second-semester ODRs and OSSs). The first models examined the overall association between student participation in RIs and subsequent discipline outcomes. The second models included interaction terms testing whether the correlation between participation in RIs and subsequent discipline outcomes varied by student race or by school-level rates of RI use. For Research Question 2, we ran a single model with socio-demographic characteristics predicting participation in one or more RIs.

In all analyses, we covaried number of referrals for each discipline category, as the District requires schools to implement a graduated discipline system in which consequences increase with the seriousness and number of student offenses. This practice is consistent with empirical evidence from other school districts indicating that ODR reason is related to the severity of consequence (Skiba et al., 2014). All possible reasons for referral were included as covariates in analyses, but regression estimates were only tabled for the nine most common reasons due to space constraints. The complete data and model output for all referral reasons are available upon request. To improve the precision of our estimates, we also covaried student characteristics (e.g., free and reduced lunch eligibility) and school characteristics (i.e., school size, proportion Black student enrollment).

Results

While reviewing the results below, the reader must keep in mind that this correlational study examines the association between RIs and students' future discipline outcomes, without accounting for all relevant confounds. Therefore, findings do not indicate causality or provide strong evidence of intervention efficacy or effectiveness. Moreover, because the study dataset does not include implementation process data, the circumstances under which certain groups came to participate in RIs more or less than others are not known. Results should therefore be interpreted to add depth to prior findings from descriptive, qualitative, and single case research on RIs and identify patterns and relationships that provide direction for future studies.

RIs and Second-Semester ODRs and OSSs

Tables 3 and 4 show the results of a multilevel logistic regression model predicting the odds of receiving at least one ODR or OSS in the second semester of the same year for students who participated in RIs during the first semester. The odds ratio (OR) for each predictor and the 95% confidence intervals for each OR represent the association of an individual predictor (e.g., number of RIs received) with the dependent variable (e.g., ODR), accounting for all other covariates. An OR larger (or smaller) than 1.00 indicates an increase (or decrease) in the odds of an RI participant receiving future discipline, compared to his or her referred peers who did not receive a RI. Findings reveal that, after accounting for students' reasons for and frequency of ODRs and a range of school and student characteristics, students who received RIs as consequences for referrals in the first semester had lower odds than their peers of being referred back to the office for misconduct in the second semester ($OR = .21, p < .001$; Model 1, Table 3); these students were also less likely to receive an OSS ($OR = .07, p < .001$; Model 3, Table 4) in the second semester.

The next statistical models included interaction terms to determine if the strength of this correlation varied by student or school characteristics. The results indicate that results were equivalent across racial groups for second-semester ODRs (Table 3, Model 2) and OSSs (Table 4, Model 4). Specifically, the interaction terms (e.g., Black \times Participation in RIs) testing whether student race moderated the link between participation in RIs and second-semester discipline outcomes were all nonsignificant.

Interaction terms were also included to examine whether the negative association between participation in RIs and adverse discipline outcomes was moderated by school-wide RI use (Participation in RI \times School RI Rate). As shown in Table 3, Model 2 indicates that RI participants had lower odds of receiving a second-semester ODR in schools that had higher school-wide RI rates ($OR = .86, p < .001$). These results suggest that school-level use of RI relates to the probability of a student receiving an ODR after

Table 3

Multilevel Logistic Regression Model of the Relationship Between Participation in RIs and Second-Semester Office Disciplinary Referrals (n = 9,921)

	Model 1		Model 2	
	OR	95% CI	OR	95% CI
Participation in RI				
First-semester RI	0.21***	.09, .51	.15***	.05, .43
Student sociodemographics				
Race (comparison group = White youth)				
Latino	0.95	.79, 1.14	.96	.80, 1.15
Black	1.09	.90, 1.32	1.09	.90, 1.32
Native American	1.22	.73, 2.04	1.22	.73, 2.04
Asian	0.93	.61, 1.41	.88	.57, 1.36
Multiracial	0.88	.64, 1.21	.89	.65, 1.22
Pacific Islander	2.71	.58, 12.79	2.65	.56, 12.49
Gender (male)	0.95	.86, 1.05	.96	.86, 1.06
Eligible for free/reduced lunch	0.97	.84, 1.12	.97	.84, 1.12
ELL	1.09	.97, 1.23	1.09	.96, 1.22
Special education	0.92	.81, 1.06	.92	.80, 1.05
ED	0.71*	.53, .95	.69*	.51, .92
Referral reason ^a				
Bullying	2.12***	1.76, 2.55	2.11***	1.78, 2.54
Destruction of school property	1.81**	1.25, 2.63	1.81**	1.25, 2.63
Disobedient/defiant	2.10***	1.93, 2.29	2.06***	1.90, 2.24
Other code of conduct violation	1.97***	1.77, 2.20	1.95***	1.75, 2.17
Detrimental behavior	1.89***	1.75, 2.03	1.87***	1.73, 2.01
Third degree assault	2.55***	1.84, 3.53	2.55***	1.85, 3.56
Unlawful sexual behavior	1.68*	1.10, 2.55	1.67*	1.10, 2.54
Drug possession or distribution	1.67***	1.40, 1.98	1.64***	1.38, 1.95
Dangerous weapon	2.07***	1.37, 3.12	2.04***	1.36, 3.08
School context				
% Black students	0.40*	.16, 1.00	.40 ⁺	.16, 1.01
% eligible free/reduced lunch	1.64	.89, 3.02	1.65	.89, 3.04
High schools (Grades 9–12)	0.43***	.29, .64	.43***	.29, .64
Middle schools (Grades 6–8)	0.87	.60, 1.26	.87	.60, 1.27
Other grade configurations	0.68**	.49, .95	.68*	.49, .94
School size	1.07**	1.03, 1.12	1.07**	1.03, 1.12
Rate of RI use	.98	.40, 2.36	1.08	.45, 2.65
Interactions				
Native American × Participation in RI			1.68	.31, 9.19
Black × Participation in RI			1.41	.70, 2.86
Latino × Participation in RI			1.08	.56, 2.09

(continued)

Table 3 (continued)

	Model 1		Model 2	
	OR	95% CI	OR	95% CI
Asian × Participation in RI			.36	.04, 3.93
Multiracial × Participation in RI			1.55	.45, 5.36
Pacific Islander × Participation in RI			— ^b	
Participation in RI × School RI Rate			.86***	.80, .93
Model statistics				
Constant	2.75***	1.48, 5.10	2.62***	1.41, 4.86
Variance component ^c	-.64	-.97, -.32	-.64	-.96, -.31
ICC ^d	.14	.10, .18	.14	.10, .18
Log likelihood	-5,148.02		-5,133.92	

Note. CI = confidence interval; OR = odds ratio.

^aAdditional low-frequency reasons for referral were included in the statistical model, but estimates are not presented in the table. The complete output for all models is available upon request from the authors.

^bNo Pacific Islander students received a RI in the first semester.

^cLog of the school-level random effect variance component.

^dResidual intraclass correlation, or the total variance contributed by the school-level random effect variance component.

+ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

participating in the intervention. Holding all other variables in the model constant, a referred student who did not participate in an RI during the first semester and attended a school with an average first-semester RI rate ($M = 8.31\%$) had a 72% probability of receiving one or more ODRs in the second semester (Huang, 2014; Huang, Invernizzi, & Drake, 2012). A referred student who did participate in at least one RI in the first semester and attended a school with an average first-semester RI rate had a much lower (28%) probability of receiving one or more ODRs in the second semester. A referred student who participated in at least one RI but attended a school with a school-wide RI rate that was 1 SD above the mean ($SD = 13.10\%$) had an even lower (18%) probability of receiving one or more ODRs in the second semester. In contrast, moderation by school-level rate of RI use was not statistically significant when predicting second-semester OSSs (Table 4, Model 4, Participation in RI × School RI Rate = ns). That said, the school-level RI rate was negatively correlated with receiving OSSs for all students (OR = .13, $p < .01$).

The models also revealed that even after accounting for RI participation at the student and school level, Black students and those eligible for free

Table 4
Multilevel Logistic Regression Model of the Relationship Between Participation
in RIs and Second-Semester OSS (n = 9,921)

	Model 3		Model 4	
	OR	95% CI	OR	95% CI
Participation in RI				
First-semester RI	0.07***	.01, .31	.10**	.02, .50
Student sociodemographics				
Race (comparison group = White youth)				
Latino	1.16	.95, 1.43	1.16	.94, 1.42
Black	1.33**	1.08, 1.65	1.32**	1.07, 1.64
Native American	1.54	.89, 2.66	1.47	.84, 2.55
Asian	1.31	.81, 2.12	1.33	.81, 2.16
Multiracial	1.31	.92, 1.85	1.27	.90, 1.81
Pacific Islander	1.14	.30, 4.35	1.17	.31, 4.44
Gender (male)	0.97	.87, 1.08	.97	.87, 1.08
Eligible for free/reduced lunch	1.28**	1.08, 1.51	1.27**	1.08, 1.50
ELL	0.91	.80, 1.04	.91	.80, 1.03
Special education	1.16*	1.00, 1.33	1.16*	1.01, 1.33
ED	1.39**	1.06, 1.83	1.37*	1.04, 1.81
Referral reason ^a				
Bullying	1.45***	1.24, 1.70	1.46***	1.25, 1.72
Destruction of school property	2.08***	1.50, 2.87	2.07***	1.50, 2.87
Disobedient/defiant	1.29***	1.23, 1.34	1.29***	1.23, 1.34
Other code of conduct violation	1.47***	1.35, 1.59	1.47***	1.36, 1.60
Detrimental behavior	1.62***	1.54, 1.70	1.62***	1.54, 1.70
Third-degree assault	3.05***	2.29, 4.06	3.10***	2.33, 4.13
Unlawful sexual behavior	1.86**	1.26, 2.74	1.88**	1.27, 2.77
Drug possession or distribution	3.26***	2.75, 3.86	3.27***	2.76, 3.87
Dangerous weapon	4.75***	3.32, 6.80	4.75***	3.32, 6.80
School context				
% Black students	2.07	.51, 8.17	2.06	.52, 8.22
% eligible free/reduced lunch	4.12**	1.61, 10.52	4.11**	1.61, 10.50
High schools (Grades 9–12)	0.80	.43, 1.49	.80	.43, 1.49
Middle schools (Grades 6–8)	1.86*	1.05, 3.29	1.86*	1.05, 3.29
Other grade configurations	1.49	.91, 2.43	1.49	.91, 2.43
School size	1.08*	1.01, 1.16	1.08*	1.01, 1.16
Rate of RI use	.13**	.03, .53	.13**	.03, .54
Interactions				
Native American × Participation in RI			.14	.01, 1.91
Black × Participation in RI			.80	.36, 1.80
Latino × Participation in RI			.62	.29, 1.34

(continued)

Table 4 (continued)

	Model 3		Model 4	
	OR	95% CI	OR	95% CI
Asian × Participation in RI			1.23	.13, 11.86
Multiracial × Participation in RI			.39	.09, 1.76
Pacific Islander × Participation in RI			— ^b	
Participation in RI × School RI Rate			.98	.91, 1.05
Model statistics				
Variance component ^c	.33	.05, .60	.32	.05, .60
ICC ^d	.30	.24, .36	.30	.24, .36
Log likelihood	−4,698.72		−4,695.00	

Note. CI = confidence interval; OR = Odds ratio.

^aAdditional low-frequency reasons for referral were included in the statistical model, but estimates are not presented in the table. The complete output for all models is available upon request from the authors.

^bNo Pacific Islander students received a RI in the first semester.

^cLog of the school-level random effect variance component.

^dResidual intraclass correlation or the total variance contributed by the school-level random effect variance component.

* $p < .05$. ** $p < .01$. *** $p < .001$.

lunch, classified as having an ED, or receiving special education services still had higher odds of receiving second-semester OSSs relative to their peers (ORs ranged from 1.16 to 1.37, $p < .05$). This finding held no matter the seriousness and frequency of ODRs or the type of school setting (e.g., grade level, school size), indicating that despite RI participation, disparities in exclusionary discipline remained for Black students, low-income students, and students in special education.

Equitable Participation in RIs

Table 5 presents the results of the multilevel logistic regression model predicting student participation in at least one RI. Results indicate that, accounting for students' number of ODRs in each referral category and school-level covariates, only one student group of interest—youth designated as ELL—had lower odds of participating in an RI (OR = .81, $p < .05$) compared to non-ELL students. On the other hand, many student groups that tend to be overrepresented in exclusionary discipline outcomes were equally likely to participate in a RI as their peers. Specifically, Native American students, males, low-income students, students in special

Table 5

Multilevel Logistic Regression Model Predicting Participation in a RI (n = 9,921)

	OR	95% CI
Student sociodemographics		
Race (comparison group = White students)		
Latino	1.40*	1.07, 1.83
Black	1.36*	1.02, 1.81
Native American	0.89	.40, 1.96
Asian	1.24	.65, 2.36
Multiracial	0.88	.53, 1.47
Pacific Islander	1.47	.29, 7.37
Gender (male)	1.04	.89, 1.21
Eligible for free or reduce price lunch	1.03	.83, 1.28
ELLs	0.81*	.69, .97
Special education	0.92	.76, 1.12
ED	0.82	.54, 1.23
Referral reason ^a		
Bullying	1.35**	1.11, 1.63
Destruction of school property	1.03	.65, 1.63
Disobedient/defiant	1.04 ⁺	.99, 1.09
Other code of conduct violation	1.31***	1.19, 1.44
Detrimental behavior	1.37***	1.30, 1.44
Third-degree assault	2.14***	1.47, 3.09
Unlawful sexual behavior	.86	.58, 1.28
Drug possession or distribution	0.72*	.55, .93
Dangerous weapon	1.24	.76, 2.03
School context		
% Black students	0.13	.01, 2.58
% eligible for free or reduced price lunch	1.26	.18, 8.81
High school	2.27	.60, 8.60
Middle school	3.29 ⁺	.94, 11.45
Alternative grade configuration	3.50**	1.21, 10.11
School size	1.10	.95, 1.27
Model statistics		
Variance component ^b	1.78	.143, 2.13
ICC ^c	.65***	.56, .72

Note. CI = confidence interval; OR = odds ratio.

^aAdditional low-frequency reasons for referral were included in the statistical model, but estimates are not presented in the table. The complete output for all models is available upon request from the authors.

^bLog of the school-level random effect variance component.

^cResidual intraclass correlation or the total variance contributed by the school-level random effect variance component.

+*p* < .10. **p* < .05. ***p* < .01. ****p* < .001.

education, and students classified as having an ED had similar odds of participating in a RI as other, more advantaged groups of students. Results also indicate that two marginalized and disadvantaged groups were more likely to participate in RIs. Of interest was that Latino (OR = 1.40, $p < .05$) and Black (OR = 1.36, $p < .05$) students had higher odds of participating in a RI relative to White students.

As shown in Table 5, findings from the statistical models also reveal important information about the type of offenses that are associated with RI participation. For example, students who were referred for a greater number of offenses involving interpersonal conflict had the highest odds of participating in RIs; these offenses included bullying, detrimental behavior, and third-degree assault (ORs ranged from 1.35–2.14, $p < .01$). In contrast, students referred for drug possession or distribution were less likely than other students to engage in a RI (OR = .72, $p < .05$). Other referral reasons, such as destruction of school property, disobedience or defiance, and weapons possession, were not statistically significant predictors of participation in RIs. Also noteworthy was the finding indicating that students in schools with alternative grade configurations (relative to elementary schools) had over three times higher odds of participating in RIs (OR = 3.50, $p < .05$).

Discussion

The study suggests that RIs may be a useful alternative to punitive, exclusionary consequences. Findings corroborate prior research (Anyon et al., 2014) and address methodological limitations by (a) controlling for a range of student and school characteristics using a multilevel modeling approach and (b) using time-ordered discipline records of individual students across a school year. Specifically, with each RI students received (circles, mediations, or conferences) during the first semester, their odds of receiving another ODR or OSS in the second semester were lower. This association held after accounting for sociodemographics (e.g., race, gender, free/reduced lunch eligibility), educational placements (e.g., general or special education), frequency or seriousness of office referrals (e.g., detrimental behavior, third-degree assault, dangerous weapon possession), and diverse school environments in terms of grade level (e.g., elementary school, high school), size of the student body, proportion of Black and low-income students, and school-level RI rate. The study also found that the negative association between participation in RIs and adverse discipline outcomes was similar across racial groups; in other words, student race did not have a moderating role.

Our ability to interpret these results or make claims about the impact of RIs on discipline outcomes is highly constrained by lack of random assignment (by school or student) to RIs and the limited covariates in our dataset. The associations between receipt of RIs in the first semester and fewer ODRs/OSSs in the second semester did not account for a range of student

and school characteristics that could influence RI participation and/or subsequent discipline incidents. For example, there is emerging evidence that students' likelihood of participating in a restorative conference is influenced by their trust or relationship with the person who will be implementing the intervention (Anyon, 2016). Other influences could include students' propensity to take responsibility for their actions, a disciplinarian's willingness to offer students the opportunity to participate in a RI, or a school leader's commitment to proactive or preventative approaches to addressing misbehavior (e.g., Payne & Welch, 2010; Skiba et al., 2014). Since we were not able to account for these confounding factors, study results cannot be interpreted to mean that first-semester RI participation caused a reduction in second-semester ODRs and OSSs. Nevertheless, the finding of a negative association between these two variables is promising and warrants further investigation.

Variability in School Use of RIs

The study demonstrated that schools varied considerably in the rates at which referred students in the school participated in circles, mediations, and conferences. For example, the full range of school-wide RI rates was 0% to 75% ($M = 8.3\%$, $SD = .13$), with 13% of all disciplined students in the district having participated in at least one RI. Referred students in schools with higher rates of RI use, in general, had lower odds of receiving an OSS than students in schools with lower RI rates. In the case of office referrals, school-wide RI rate also moderated the relationship between individual RI participation and subsequent ODRs. This might suggest that schools implementing circles, mediations, and conferences are generally seeking to steer students out of the discipline system and limit the use of suspension when they do. Perhaps they are engaging in a broad set of prevention and intervention initiatives to keep students in the classroom and the school building. This in and of itself is a worthy goal given that negative academic and behavioral trajectories of referred students are exacerbated when they are excluded from instruction (e.g., Balfanz et al., 2015).

Also noteworthy was that results at the student level held no matter the school's overall rate of RI use. Whether or not schools regularly or rarely engaged students in circles, mediations, or conferences, the negative correlation between RI participation in the first semester and exclusionary practices (ODRs and OSSs) in the second semester remained. Finally, the study found school-level use of RIs moderated the relationship between student-level participation in RIs and another discipline incident in the second semester. In other words, high program use at the school level (a facet of treatment delivery; Schulte et al., 2009) strengthened the negative relationship between RI participation and subsequent office referrals at the student level. This suggests that school-level participation rates relate to the

probability of a student experiencing an office disciplinary referral after participating in the intervention. These findings are consistent with arguments some RI scholars have made that the depth of school community engagement with these practices is critical to maximizing their benefits (Anyon, 2016). Indeed, as innovations spread and take hold through a school building, educators' attitudes, beliefs, and skills related to the new programming may actually strengthen the quality of implementation and resulting effects on student outcomes (Rogers, 2003). On the other hand, it is also plausible that school capacity to implement RIs meaningfully operates as a crucial driver of disciplinary outcomes.

RIs and Equity Issues

Results of this study demonstrate that in a large urban district many disadvantaged youth had similar rates of RI participation as more privileged students, with the noteworthy exception of ELL students. Relative to their peers, low-income students, Native American youth, males, and students with special education services or an ED classification participated comparably in RIs. Black and Latino students were more likely to participate in RIs than White students. Moreover, the only school-level predictor of student-level RI participation was the grade configuration of the school; students at sites with non-traditional formats (e.g., K-12, K-8) had significantly higher odds of participating in this alternative to suspension. These findings are surprising in light of experimental studies indicating that disadvantaged groups may consciously (or unconsciously) be issued harsher sanctions for similar behavior than more advantaged groups (Okonofua & Eberhardt, 2015). More specifically, results stand in contrast to previous research by Payne and Welch (2010) indicating that, at the school level, the proportion of Black students was a negative predictor of a school's use of student conferences, peer mediations, restitution, and community service in response to discipline incidents. It is possible that the unique dynamics of service access in this study reflect the District's focus on eliminating racial disparities in school discipline, as articulated by board policy, professional development trainings on RIs, and district officials' public statements (DPS, 2008, 2012). However, it is concerning that ELLs were less likely to participate in RIs. This finding is consistent with prior research demonstrating that students' access to school-based programs can be limited for students for whom English is not their native language (Anyon et al., 2013). Taken together, findings suggest that educators need to be vigilant in ensuring fair access to less punitive alternatives when implementing new discipline initiatives, such as RIs. In other words, the recent push for schools to disaggregate their ODR and suspension data should be extended to the use of RIs.

Despite a higher likelihood of participation in RIs relative to White students, Black students remained at heightened risk of being suspended in the

second semester. Likewise, despite having comparable participation in RIs, low-income students, youth in special education, and those with an ED classification also had higher odds of being issued a second-semester suspension than their more advantaged peers. Comparable participation in RIs for these groups did not correspond with reducing heightened risk among these populations for a future second-semester OSS across the district—a risk that persisted even after controlling for participation in RIs and frequency and seriousness of referral reasons. In other words, discipline disparities were not eliminated in the District despite its use of restorative alternatives to suspension. It therefore seems likely that additional forms of prevention and intervention, in addition to individual RIs, are needed to fully address equity concerns.

The persistence of disparities may be due to a number of reasons. RIs were issued as consequences for a wide range of referral reasons including detrimental behavior, bullying, and third-degree assault. That said, only 12.52% of all those referred for discipline in the first semester in the District ($n = 652$) received a RI. Implementation may need to be much more widespread and frequent to significantly reduce or eliminate disparities in discipline. The infrequent use of RIs in many schools reflects the challenge of integrating alternative disciplinary strategies in large urban school districts. Although the District offers voluntary training about RIs, additional resources like school-based RI coordinators may be necessary to increase implementation and reduce racial discipline gaps district-wide (Anyon, 2016; Durlak & Dupre, 2008).

In addition, the District may make additional gains in reducing discipline disparities by increasing their prevention efforts (while maintaining the focus on restorative approaches to intervention). Building community, creating positive social bonds, and fostering investment in school rules before conflict arises may be among the keys to reducing disparities, especially for students in groups who are alienated from school (Gregory, Bell, & Pollock, 2016). Moreover, preventive interventions can also occur by training teachers to strengthen the motivating and engaging qualities of instruction and by preventing negative teacher–student interactions from occurring in the first place (Gregory, Hafen et al., 2016). Finally, staff training about culturally responsive practices and racial justice may reduce the likelihood of misreading or mislabeling students' body language or speech and may decrease overly punitive responses to students of color (Davis, Lyubansky, & Schiff, 2015; Monroe, 2005).

Study Limitations and Implication for Future Research

Several limitations related to study design suggest that caution in interpreting study findings is warranted. First and foremost, we must reiterate that neither disciplined students nor school sites were randomly assigned

to participate in RIs. Instead, the District policy recommends that RIs be used in response to midrange offenses when a student is willing to accept some responsibility for his or her actions. This potential of individual and school sorting of students toward or away from a RI suggests multiple potential sources of selection bias. For example, at the student level, those who were able to acknowledge their role in an offense may have already been less likely to reenter the discipline system. In other words, a student's lack of ability to acknowledge his or her contributing role in a discipline incident is likely a risk for future discipline contacts but was not measured in this study. Moreover, the study was not able to directly measure and account for a school's or administrator's propensity to offer students RIs or punitive consequences in response to a discipline incident.

Whereas the current study found a significant association between RI participation and positive discipline outcomes, a crucial goal for future research is to move beyond a conditional analysis such as this to begin to identify design or analytic strategies to mitigate the influence of key selection processes. We anticipate that there may be barriers in implementing experimental designs to estimate the effects of RI (e.g., ethical considerations in differentially offering less exclusionary disciplinary options to students, given evidence of the harm of OSSs). If researchers must rely on observational designs, crucial first steps include more specific identification of (1) the RI intervention (e.g., what are the essential components of the intervention at school and student levels and at what intensity); (2) the student, educator, and school characteristics that are predictive of RI participation beyond basic sociodemographic characteristics (e.g., student verbal communication, emotional regulation, externalizing behaviors, history of disciplinary interactions, prior relationships with school staff, peer norms about the acceptability of RI, and a school or district's willingness to implement RIs); and (3) how such characteristics contribute to exclusionary discipline outcomes of interest (Hemphill, Plenty, Herrenkohl, Toumbourou, & Catalano, 2014; Skiba et al., 2014). If such factors can be observed and measured reliably, propensity score matching methods may be a viable option for understanding the impact of RIs (see Hong & Raudenbush, 2005, as an exemplar).

The low incidence of RI use among disciplined students also raises questions about the capacity for widespread dissemination of this approach in school settings. When the revised policy was passed, only 4% of all disciplined students participated in RIs. Four years later, the incidence of RI use increased to 13%, as reported in this study. Although this rate is low, it is important to note that district policy mandates were not accompanied with the financial incentives or staffing. The increased use of RIs is more impressive in this context, in which the only implementation supports offered to schools have been voluntary training and technical assistance led by one district coordinator. Future research should consider whether the diffusion of RIs in schools could be more widespread, with attention

to other factors that affect the implementation of interventions like work climate, staff norms, the support of organizational leaders, and the development of “program champions” (Durlak & Dupre, 2008).

Another study limitation is that our dataset provided no information to assess the process or quality of RI implementation. Schools did not track whether, for example, disciplinarians adhered to eligibility protocols when offering RIs to students with discipline incidents or if participants were motivated by threats for more severe repercussions if they did not engage in an RI. Moreover, the dataset did not indicate to what degree each school experienced pressure to implement this approach or the degree to which educators felt motivated or skilled to do so.

Future research would be substantially strengthened by the inclusion of indicators that measure multiple and multilevel (student/school) factors related to implementation fidelity. Of particular relevance to school-based RIs are compatibility or fit between school leaders’ discipline philosophies and the principles that guide RIs, the degree to which school staff buy in to the approach, staff capacity (in particular the presence or absence of a person trained and available to facilitate formal mediations or conferences), and school personnel’s participation in RI trainings (Anyon, 2016; Payne & Welch, 2010). Such process and quality characteristics and the development of measures thereof have great potential to enhance both observational and experimental designs.

To further elucidate the finding of moderation by school-level rate of RIs, future research might compare schools with high versus low rates to ascertain if high participation schools have (1) better skilled facilitators of circles, mediations, and conferences due to practice; (2) positive peer norms related to RIs that influence individual students’ commitment to the restorative process; and/or (3) positive staff attitudes and expectancies related to RIs that impact their commitment to providing behavioral supports and following through with RI participants. These studies should also explore in a more multifaceted way the quality, quantity, and the degree to which RI needs to be implemented school wide in order to be maximally effective. More broadly, study findings highlight the need to account for school-level contextual factors in future studies of RIs.

Similarly, there are potentially multiple decision-making points that could result in the finding of varying RI participation rates across diverse student groups. School staff may tend to refer to RIs students who are from certain racial groups more than from other racial groups. Once referred and oriented to the restorative process, students from certain groups may tend to consent to participate more than their peers from other groups. Given this complexity, the current study’s focus on participation in RIs is only a first step in understanding disparate or comparable patterns of use of a promising alternative to suspension. Future research might seek to explain why Black and Latino students were more likely to participate in RIs than White students in this study. It would be informative to know if school staff members

are aware of racial disparities in school discipline and are actively encouraging the use of less punitive practices among students of color. Studies might also seek to explain our finding that ELL students' RI participation was low relative to non-ELL students. Such research might examine whether schools inadvertently deny access to RIs for students whose first language is not English because of cultural barriers or limited language capacity on the part of school staff.

The study's singular focus on discipline outcomes also limits the scope of the findings. ODRs reflect school staff's use of formal discipline procedures and paperwork. Whether an ODR reflects student misconduct, staff's tendency to use the formal discipline procedures, or poor classroom management skills is unknown (Morrison, Redding, Fisher, & Peterson, 2006). Future studies should include other school outcomes such as academic engagement, attendance, and achievement. This much-needed research would build on findings from Oakland where schools implementing RI, along with a range of other interventions, had significantly greater improvements in reading proficiency and greater reductions in absenteeism and dropout than non-RI schools (Jain et al., 2014).

Conclusion

Study findings advance current knowledge about patterns and correlates of student participation in school-based RIs. Using discipline records from a large urban district, results indicate that students who received a RI in the first semester had lower odds of receiving another ODR or suspension in the second semester of the same school year. This finding held after accounting for student racial background, special education status, free or reduced lunch eligibility, and frequency and seriousness of disciplinary referrals. The study also showed that school-wide RI rates were negatively associated with exclusionary discipline outcomes. In fact, for ODRs, the strength of association between RI participation and adverse discipline outcomes was more pronounced in schools with high rates of RI use. Finally, participation in RIs was comparable across many disadvantaged groups, with the notable exception of ELL students. However, Latino and Black students, two groups with disproportionately high rates of suspension in many regions in the United States, had greater odds of receiving RIs than their White peers. This suggests that for most disadvantaged groups in the District, schools implemented RIs in a manner that provided them equal access to an alternative, problem-solving approach to conflict. These findings are encouraging but do not provide causal evidence of the utility of this approach to reducing students' risk of exclusionary discipline infractions. Experimental research with robust implementation process measures is sorely needed to identify the mechanisms underlying the patterns identified in this study.

Note

¹We considered implementing a propensity score matching strategy guided by the methods of Hong and Raudenbush's (2005) multilevel study of the effects of grade retention practices in schools. We ultimately ruled out such an approach for two reasons. First, we could not be completely certain, given the structure of the dataset, that key student-level predictors (ODRs) actually preceded receipt of RI, which is a fundamental condition for propensity score matching in intervention studies. Thus, the only available student matching variables were sociodemographic characteristics, which are unlikely to generate less biased results (Cook, Shadish, & Wong, 2008). As a comparison, Hong and Raudenbush's (2005) multilevel propensity methods benefited from an extensive body of prior research on the student- and school-level attributes that predict grade retention. Moreover, their rich dataset had over 200 multilevel covariates to use in generating propensity scores. There is no comparable prior research on correlates of RI participation, and the dataset employed in this study only included a limited number of control variables.

References

- American Academy of Pediatrics. (2013). Out of school suspension and expulsion. *Pediatrics*, *131*, 1000–1007. Available at <http://pediatrics.aappublications.org/content/131/3/e1000.full.pdf+html>
- Amstutz, L., & Mullet, J. H. (2005). *The little book of restorative discipline for schools*. Intercourse, PA: Good Books.
- Anyon, Y. (2016). *Taking restorative practices school-wide: Insights from three schools in Denver*. Denver, CO: Denver School-Based Restorative Justice Partnership.
- Anyon, Y., Horevitz, L., Moore, M., Whitaker, K., Stone, S., & Shields, J. (2013). Risk behaviors, race, and use of school-based health centers: Policy and service recommendations. *Journal of Behavioral Health Services and Research*, *40*, 457–468.
- Anyon, Y., Jenson, J., Altschul, I., Farrar, J., McQueen, J., Greer, E., Downing, B., & Simmons, J. (2014). The persistent effect of race and the promise of alternatives to suspension in school discipline outcomes. *Children and Youth Services Review*, *44*, 379–386.
- Arcia, E. (2007). Variability in schools' suspension rates of black students. *The Journal of Negro Education*, *76*(4), 597–608.
- Balfanz, J., Byrnes, V., & Fox, J. (2015). Sent home and put off track: The antecedents, disproportionalities, and consequences of being suspended in the 9th grade. In D. J. Losen (Ed.), *Closing the school discipline gap: Equitable remedies for excessive exclusion* (pp. 17–30). New York, NY: Teacher's College Press.
- Berkowitz, K. (2012). *Restorative practices whole-school implementation guide*. San Francisco, CA: San Francisco Unified School District.
- Blood, P., & Thorsborne, M. (2005). *The challenges of culture change: Embedding restorative practice in schools*. Paper presented at the Sixth International Conference on Conferencing, Circles and other Restorative Practices: Building a Global Alliance for Restorative Practices and Family Empowerment, Sydney, Australia.
- 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*, 508–520.
- Braithwaite, J. (1989). *Crime, shame and reintegration*. New York, NY: Cambridge University Press.

- Braithwaite, J. (2001). Restorative justice and a new criminal law of substance abuse. *Youth and Society*, 33, 227–248.
- Buckley, S., & Maxwell, G. (2007). *Respectful schools: Restorative practices in education. A summary report*. Wellington, New Zealand: Office of the Children's Commissioner and the Institute of Policy Studies, School of Government, Victoria University.
- Carter, P., Fine, M., & Russell, S. (2014). *Discipline disparities series: Overview*. Bloomington, IN: The Equity Project at Indiana University. Available at <http://rtpcollaborative.indiana.edu/briefing-papers/>
- Chapin Hall Center for Children. (2008). *Racial and ethnic disparity and disproportionality in child welfare and juvenile justice: A compendium*. Chicago: Chapin Hall Center for Children at the University of Chicago.
- Coates, R. B., Umbreit, M., & Vos, B. (2003). Restorative justice circles: An exploratory study. *Contemporary Justice Review*, 6, 265–278.
- Cook, T. D., Shadish, W. R., & Wong, V. C. (2008). Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *Journal of Policy Analysis and Management*, 27(4), 724–750.
- Costello, B., Wachtel, J., & Wachtel, T. (2010). *Restorative circles in schools: Building community and enhancing learning*. Bethlehem, PA: International Institute for Restorative Practices.
- Corrigan, M. (2012). *Restorative practices in New Zealand: The evidence base*. Wellington, New Zealand: Ministry of Education.
- Davis, F. E., Lyubansky, M., & Schiff, M. (2015). Restoring racial justice. In R. Scott & S. Kosslyn (Eds.), *Emerging trends in the social and behavioral sciences* (pp. **IAQ: 3**). Hoboken, NJ: John Wiley & Sons, Inc.
- Denver Public Schools. (2008). *Policy JK-R—Student conduct and discipline procedures*. Available at <http://sts.dpsk12.org/dps-school-board-policies/>
- Denver Public Schools. (2012). *Restorative approaches training manual*. Denver, CO: Mental Health and Assessment, Division of Student Services.
- Department of Planning and Analysis. (2013). *5 year enrollment trends*. Denver, CO: Denver Public Schools.
- Derezotes, D., Testa, M., & Poertner, J. (Eds.). (2005). *Race matters in child welfare: The overrepresentation of African American children in the system*. Washington, DC: Child Welfare League of America.
- Drewery, W. (2004). Conferencing in schools: Punishment, restorative justice, and the productive importance of the process of conversation. *Journal of Community & Applied Social Psychology*, 14, 332–344.
- Drewery, W. (2013). Restorative justice practice: Cooperative problem-solving in New Zealand's schools. *FORUM*, 55, 209–216.
- Durlak, J., & Dupre, E. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41, 327–335.
- Encarnacao, J. (2013, September 3). Sharp drop in suspensions as Boston schools try 'restorative' approach. *Boston Herald*.
- Fabelo, T., Thompson, M. D., Plotkin, M., Carmichael, D., Marchbanks, P., III, & Booth, E. A. (2011). *Breaking schools' rules: A statewide study of how school discipline relates to students' success and juvenile justice involvement*. Council of State Governments Justice Center Publications. Accessible at <http://justicecenter.csg.org/resources/juveniles>
- Forman, S. G. (2015). *Implementation of mental health programs in schools: A change agent's guide*. Washington, DC: American Psychological Association.

- González, T. (2015). Socializing schools: Addressing racial disparities in discipline through restorative justice. In D. J. Losen (Ed.), *Closing the school discipline gap: Equitable remedies for excessive exclusion* (pp. 151–165). New York, NY: Teachers College Press.
- Gregory, A., Bell, J., & Pollock, M. (2016). How educators can eradicate disparities in school discipline: Issues in intervention. In R. J. Skiba, K. Mediratta, & M. K. Rausch (Eds.), *Inequality in school discipline: Research and practice to reduce disparities* (pp. 39–58). New York, NY: Palgrave MacMillan.
- Gregory, A., Hafen, C. A., Ruzek, E. A., Mikami, A. Y., Allen, J. P., & Pianta, R. C. (2016). Closing the racial discipline gap in classrooms by changing teacher practice. *School Psychology Review, 45*, 171–191.
- Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher, 39*, 59–68.
- Hannon, L., DeFina, R., & Bruch, S. (2013). The relationship between skin tone and school suspension for African Americans. *Race and Social Problems, 5*, 281–295.
- Hemphill, S. A., Plenty, S. M., Herrenkohl, T. L., Toumbourou, J. W., & Catalano, R. F. (2014). Student and school factors associated with school suspension: A multi-level analysis of students in Victoria, Australia and Washington State, United States. *Children and Youth Services Review, 36*, 187–194.
- Hong, G., & Raudenbush, S. W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis, 27*(3), 205–224.
- Huang, F. L. (2014). Further understanding factors associated with grade retention: Birthday effects and socioemotional skills. *Journal of Applied Developmental Psychology, 35*, 79–93.
- Huang, F. L., Invernizzi, M. A., & Drake, E. A. (2012). The differential effects of pre-school: Evidence from Virginia. *Early Childhood Research Quarterly, 27*, 33–45.
- IDEA Data Center. (2014). *Methods for assessing racial/ethnic disproportionality in special education: A technical assistance guide* (revised). Rockville, MD: Westat, Julie Bollmer, Jim Bethel, Tom Munk, & Amy Bitterman. [AQ: 4]
- Jain, S., Bassegy, H., Brown, M. A., & Kalra, P. (2014). *Restorative justice in Oakland Schools. Implementation and impact: An effective strategy to reduce racially disproportionate discipline, suspensions, and improve academic outcomes*. Retrieved from www.ousd.k12.ca.us/
- Jeong, S., McGarrell, E. F., & Hipple, N. K. (2012). Long-term impact of family group conference on re-offending: The Indianapolis restorative justice experiment. *Journal of Experimental Criminology, 8*, 369–385.
- Kane, J., Lloyd, G., McCluskey, G., Riddell, S., Stead, J., & Weedon, E. (2007). *Restorative practices in Scottish schools*. Edinburgh, UK: Scottish Executive.
- Karp, D. R., & Breslin, B. (2001). Restorative justice in school communities. *Youth & Society, 33*, 249–272.
- Larsen, J. J. (2014). *Restorative justice in the Australian criminal justice system*. Canberra, Australia: Australian Institute of Criminology. Available at <http://www.aic.gov.au/publications/current%20series/rpp/121-140/rpp127.html>
- Latimer, J., Dowden, C., & Muise, D. (2005). The effectiveness of restorative justice practices: A meta-analysis. *The Prison Journal, 85*, 127–144.
- 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>
- Losen, D. J., & Martinez, T. E. (2013). *Out of school and off track: The overuse of suspensions in American middle and high schools*. Los Angeles, CA: University of California, Los Angeles.

- McClusky, G., Lloyd, G., Kane, J., Stead, J., Riddell, S., & Weedon, E. (2008). Can restorative practices in schools make a difference? *Educational Review*, *60*, 405–417.
- McGarrell, E. F., & Hipple, N. K. (2007). Family group conferencing and re-offending among first-time juvenile offenders: The Indianapolis experiment. *Justice Quarterly*, *24*, 221–246.
- Monroe, C. R. (2005). Understanding the discipline gap through a cultural lens: Implications for the education of African American students. *Intercultural Education*, *16*, 317–330.
- Morgan, E., Salomon, N., Plotkin, M., & Cohen, R. (2014). *The school discipline consensus report: Strategies from the field to keep students engaged in school and out of the juvenile justice system*, New York, NY: The Council of State Governments Justice Center.
- Morris, E. W. (2005). 'Tuck in that shirt!' Race, class, gender and discipline in an urban school. *Sociological Perspectives*, *48*, 25–48.
- Morrison, G. M., Redding, M., Fisher, E., & Peterson, R. (2006). Assessing school discipline. In S. R. Jimerson & M. J. Furlong (Eds.), *Handbook of school violence and school safety: From research to practice* (pp. 211–220). Mahwah, NJ: Lawrence Erlbaum Associates.
- Nathanson, D. (1997). Affect theory and the compass of shame. In M. Lansky & A. Morrison (Eds.), *The widening scope of shame* (pp. 339–354). Hillsdale, NJ: Analytic Press.
- Okonofua, J. A., & Eberhardt, J. L. (2015). Two strikes: Race and the disciplining of young students. *Psychological Science*, *26*, 617–624.
- Osher, D. M., Poirier, J. M., Jarjoura, G. R., & Brown, R. C. (2014). Avoid quick fixes: Lessons learned from a comprehensive districtwide approach to improve conditions for learning. In D. J. Losen (Ed.), *Closing the school discipline gap: Equitable remedies for excessive exclusion* (pp. 192–206). New York, NY: Teachers College Press.
- Pas, E. T., Bradshaw, C. P., & Mitchell, M. M. (2011). Examining the validity of office discipline referrals as an indicator of student behavior problems. *Psychology in the Schools*, *48*, 541–555.
- Payne, A. A., & Welch, K. (2010). Modeling the effects of racial threat on punitive and restorative school discipline practices. *Criminology*, *48*, 1019–1062.
- Public Schools of North Carolina. (2013). *Annual report on the use of corporal punishment session law 2011-282*. Raleigh, NC: State Board of Education Department of Public Instruction. Available at <http://www.ncpublicschools.org/>
- Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and longitudinal modeling using stata*. College Station, TX: Stata Press.
- Rausch, M. K., Skiba, R. J., & Simmons, A. B. (2004). *The academic cost of discipline: The relationship between suspension/expulsion and school achievement*. Bloomington, IN: Center for Evaluation and Education Policy, School of Education, Indiana University.
- Reyes, J. A., Elias, M. J., Parker, S. J., & Rosenblatt, J. L. (2013). Promoting educational equity for disadvantaged youth: The role of resilience and social-emotional learning. In S. Goldstein & R. B. Brooks (Eds.), *Handbook of resilience in children* (pp. 349–370). New York, NY: Springer.
- Riestedberg, N. (2013). Challenges to education: Restorative approaches as a radical demand on conservative structures of schooling. In H. Cremin, G. McCluskey, & E. Sellman (Eds.), *Restorative approaches to conflict in schools: Interdisciplinary perspectives on whole school approaches to managing relationships* (pp. 207–216). London: Routledge.

- Rogers, E. M. (2003). *Diffusion of innovations*. New York, NY: Free Press.
- Schiff, M. (2013). *Dignity, disparity and desistance: Effective restorative justice strategies to plug the "school-to-prison pipeline."* Paper presented at the Center for Civil Rights Remedies National Conference "Closing the School to Research Gap: Research to Remedies Conference," Washington, DC.
- Schulte, A. C., Easton, J. E., & Parker, J. (2009). Advances in treatment integrity research: Multidisciplinary perspectives on the conceptualization, measurement, and enhancement of treatment integrity. *School Psychology Review, 38*, 460–475.
- Shaw, S. R., & Braden, J. P. (1990). Race and gender bias in the administration of corporal punishment. *School Psychology Review, 19*, 378–384.
- Shaw, T. V., Putnam-Hornstein, E., Magruder, J., & Needell, B. (2008). Measuring racial disparity in child welfare. *Child Welfare, 87*, 23–36.
- Skiba, R. J., Chung, C., Trachok, M., Baker, T. L., Sheya, A., & Hughes, R. L. (2014). Parsing disciplinary disproportionality: Contributions of infraction, student, and school characteristics to out-of-school suspension and expulsion. *American Educational Research Journal, 51*, 640–670.
- Skiba, R. J., Simmons, A. B., Ritter, S., Gibb, A. C., Rausch, M. K., Cuadrado, J., & Chung, C. G. (2008). Achieving equity in special education: History, status, and current challenges. *Exceptional Children, 74*(3), 264–288.
- Smith, N., & Weatherburn, S. (2012). Youth justice conferences versus children's court: A comparison of re-offending. *Contemporary Issues in Crime and Justice, 160*, 1–23.
- Steenbergen, M. R., & Jones, B. S. (2002). Modeling multilevel data structures. *American Journal of Political Science, 46*, 218–237.
- Steinberg, M. P., Allensworth, E., & Johnson, D. W. (2014). What conditions support safety in urban schools? The influence of school organizational practices on student and teacher reports of safety in Chicago. In D. J. Losen (Ed.), *Closing the school discipline gap: Equitable remedies for excessive exclusion* (pp. [AQ: 5]). New York, NY: Teachers College Press.
- Stinchcomb, J. B., Bazemore, G., & Riestenberg, N. (2006). Beyond zero tolerance: Restorative justice in secondary schools. *Youth Violence and Juvenile Justice, 4*, 123–147.
- Strang, H., Sherman, L., Mayo-Wilson, E., Woods, D., & Ariel, B. (2013). Restorative justice conferencing (RJC) using face-to-face meetings of offenders and victims: Effects on offender recidivism and victim satisfaction. A systematic review. *Campbell Systematic Reviews, 9*(12). Available at <http://www.campbellcollaboration.org/lib/download/3066/>
- Teachers Unite. (2014). *Growing fairness toolkit*. Available at <http://www.teachersunite.net/toolkits>
- Tobin, T. J., & Sugai, G. M. (1999). Using sixth-grade school records to predict school violence, chronic discipline problems, and high school outcomes. *Journal of Emotional and Behavioral Disorders, 7*, 40–53.
- U.S. Department of Education. (2014). *Guiding principles: A resource guide for improving school climate and discipline*, Washington, DC: Author.
- Wachtel, T., Costello, B., & Wachtel, J. (2009). *The restorative practices handbook for teachers, disciplinarians and administrators*. Bethlehem, PA: International Institute of Restorative Practices.
- Wachtel, T., O'Connell, T., & Wachtel, J. (2010). *Restorative justice conferencing: Real justice and the conferencing handbook*. Bethlehem, PA: International Institute for Restorative Practices.
- Weinstein, R. S. (2002). *Reaching higher: The power of expectations in schooling*. Cambridge, MA: Harvard University Press.

- Wong, D. S., & Mok, L. W. (2011). Restorative justice and practices in China. *Journal of Community Justice*, 8, 23–35.
- Zehr, H. (2002). *The little book of restorative justice*. Intercourse, PA: Good Books.
- Zehr, H., & Toews, B. (2004). *Critical issues in restorative justice*. Monsey, NY: Criminal Justice Press.

Manuscript received November 23, 2014

Final revision received September 25, 2016

Accepted September 26, 2016