2011

Disciplinary Outcomes by Race and Gender in Schools Implementing Positive Behavior Support: Does Fidelity of Implementation Reduce Disproportionality?

Therese Sandomierski
University of South Florida, ThereseSando@gmail.com

Disciplinary Outcomes by Race and Gender in Schools Implementing Positive Behavior Support:

Does Fidelity of Implementation Reduce Disproportionality?

by

Therese Sandomierski

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Psychological and Social Foundations
College of Education
University of South Florida

Major Professor: Harold R. Keller, Ph.D.
Linda Raffaele-Mendez, Ph.D.
Don Kincaid, Ed.D.
Heather George, Ph.D.

Date of Approval:
July 5, 2011

Keywords: elementary behavior, referral, suspension, culturally responsive

Dedication

I dedicate this dissertation to my parents, Thomas and Patricia Sandomierski, whose own love of knowledge and tremendous work ethic gave me the drive and determination to tackle a lifetime of education. You've taught me by example, and continue to be my most influential instructors. If it weren't for your belief in me, I would not have known this was possible.

Acknowledgements

I would like to first thank my committee for their time, support, and expertise. Dr. Keller, you have been an ideal Chairperson, providing expert guidance and structure in an incredibly supportive way. Dr. Raffaele-Mendez, your emotional support and broad perspective kept me motivated and helped me to extend my thinking to a wider range of relevant issues. Dr. Kincaid, your expertise in SW-PBS and quick understanding of the data gave me new insights into significant issues impacting measurement and implementation. My understanding of the data set is considerably deepened because of your mentorship. Dr. George, your facilitation, encouragement and connection to current trends in our field has been a huge motivator, and inspires me to try to make an impact through my work. Dr. Brenda Walker, my outside Chairperson, your expertise in educational disproportionality provided validation for the core concepts and conclusions of the study, and I am grateful that you made the time to contribute to the proposal and defense. Collectively, my committee has been an outstanding source of knowledge and motivation, and I am deeply grateful for everything you've done.

I would like to thank Dr. Jennifer Frank from the University of Oregon, whose insights and assistance in securing the data set laid the foundation for a successful study.

To my colleagues at Florida's Positive Behavior Support: Response to Intervention for Behavior Project, your interest, questions, and gentle prodding bolstered me through this entire process. I could not ask for better colleagues and friends, and am grateful for your insights and encouragement. Stephanie Martinez, thank you for the extra cheerleading, structure and understanding…I hope I can repay the favor.

And last but not least, I would like to thank my partner and best friend, Ed Sherman. Your patience, love and support mean the world to me, and I'm thrilled to move into this next phase of life with you.

Table of Contents

List of Tables

List of Figures

Abstract

Disparities in behavioral outcomes for minority students are a decades-old problem. Recently, the systems-level approach of school-wide positive behavior support (SW-PBS) and its growing research base have garnered attention as a possible remedy. Although SW-PBS has been shown to be effective in reducing a school's overall level of office discipline referrals (ODRs) and suspensions (OSS), and its success has been replicated in schools with large populations of minority students, effective outcomes across all groups of students within a school are not guaranteed. Some reports document increases in the magnitude of disproportionality even when ODRs and OSS decrease for the school as a whole. However, studies of SW-PBS and disproportionality have overlooked the role of implementation fidelity as a potential mediator of student outcomes, allowing for the possibility that schools that fail to experience a reduction in ODRs and OSS across all groups of students are those in which few elements of SW-PBS have been implemented. The present study contributed to the current research base by investigating whether schools which implement SW-PBS with higher levels of fidelity were more likely to have lower levels of disproportionate ODRs and OSS for African American and Hispanic students. Drawing from online databases which record schools' implementation and ODR information, this study provided detailed school-level descriptive analyses of ODRs and OSS for African American, Hispanic, and White students. Additionally, risk ratios for receiving an ODR and for receiving an OSS were calculated for African American and Hispanic students, and then compared to each school's reported level of SW-PBS implementation as measured by their Benchmarks of Quality score. The descriptive analyses and follow-up Chi-Square analyses revealed that there was no significant relationship between a school's level of implementation fidelity and their magnitude of disproportionality for these groups of students. Implications for professional development, record keeping, and measuring disproportionality in schools are discussed.

Chapter One: Overview

A Note about Language

According to polls conducted by *Gallop,* the majority (61%) of black Americans report that they "have no preference" in the label used to describe their racial category (Newport, 2007). Of those respondents who reported a preference, the term "African American" was selected almost twice as often as the term "black," (24% to 13%; Newport, 2007). Therefore, the term "African American" will be used throughout this paper.

Introduction

The 1954 Supreme Court ruling in *Brown v. Board of Education* was a turning point in American educational history (Zirkel & Cantor, 2004). In this ruling, the Supreme Court acknowledged the detrimental effects of state-sanctioned racial segregation for students of color, and found that the resulting educational inequalities deprived students of color equal protection of the laws guaranteed by the Fourteenth Amendment (*Brown v. Board of Education*, 1954). Although more than 50 years have passed since the *Brown* ruling, the issue of educational inequalities faced by students of color continues today.

Since the *Brown* ruling, numerous studies have revealed a persistent pattern of lower academic achievement, higher rates of special education placements, and higher rates of suspension and expulsion for minority students than what is typically found in populations of White students (Baker, Hendricks, McGowan & McKechnie, 2004; Farkas, 2003; Talbert-Johnson, 2004; Zirkel, 2005). Nationally, the 2007 National Assessment of Educational Progress (NAEP) test scores show African American and Hispanic students' achievement scores in Reading and Math were noticeably lower than achievement scores for White students. For example, in 2007, only 46% of African American fourth-grade students, and 50% of Hispanic fourth-grade students scored at or above the Basic proficiency level in Reading, compared to

78% of White fourth-grade students (National Assessment of Educational Progress, 2009). Similar results were found in Reading among our nation's eighth-grade students, and among fourth- and eighth-grade students in Math. With regards to placement in special education programs, previous research at the state and national levels has shown that African American students were overrepresented in most disability categories, with the greatest disparities occurring in categories that were based on more subjectively-determined placement criteria (Green, 2005). These categories included mild/moderate mental retardation, emotional disturbance, and specific learning disability, where rates of placement for African American students were one-and-a-half to four times higher than rates of placement for White students (Losen & Orfield, 2002). Reviews of schools' use of disciplinary measures have revealed a highly consistent finding of racial disproportionality in schools' application of suspension for more than 30 years, with rates of suspension for African American students revealed as being two to three times higher than rates for White students (Skiba, 2000). African American students were also found to be more likely to experience harsher punishment than White students, for less-severe offenses (Skiba, 2000). Research on rates of discipline for Hispanic students has been mixed, with some studies finding that Hispanic students experienced higher rates of disciplinary measures than would be expected (Gordon, Piana & Keleher, 2000), while other studies found that Hispanic students experienced similar or lower rates of disciplinary measures than White students (Skiba, Poloni-Staudinger, Simmons, Feggins-Azziz & Chung, 2005).

The endurance of disproportionality among minority students, both academically and with regards to discipline, has been attributed to a wide array of causes, many of which are beyond the control of educators. However, two areas of focus offer promise both for increasing our understanding of the reasons for disproportionality, as well as remedying the problem: academic engaged time and cultural competency. Academic engaged time - the amount of time students spend involved in learning activities (Johns, Crowley & Guetzloe, 2008) - has been shown to directly impact academic achievement as well as student behavior. There are several avenues, both historically as well as current-day, through which the academic engaged time for minority students has been negatively impacted. These include shortened school years (Horsford &

McKenzie, 2008), compressed school days (Smith, 2000), tracking (Mickelson, 2001), special education placement and service provision (Blanchett, 2006), and discipline (Fenning & Rose, 2007). Interventions aimed at increasing students' academic engaged time are likely to improve academic and behavioral outcomes (Brophy, 1988).

Other researchers point towards improving educators' cultural competence as an avenue for intervention. Noting the difference between the cultures of White, middle-class teachers and their minority, (and often low-income) students (Ukpokodu, 2002), researchers have described how misinterpretations of culturally-based behavior can lead to interpersonal conflict and disciplinary removal (Vavrus & Cole, 2002). Culturally responsive education and culturally responsive classroom management, which aim to convey recognition and respect for all cultures while acknowledging the systemic inequalities that individuals from some groups may face (Ladson-Billings, 1995), have been proposed as approaches which offer promise in improving outcomes for minority students. However, there is a lack of quantitative research on the effectiveness of culturally-responsive strategies for improving students' behavioral outcomes (Bondy, Ross, Gallingane & Hambacher, 2007), and researchers have turned instead to the research base on positive behavior support for evidence.

As defined by the OSEP Center on Positive Behavioral Interventions and Support, PBS is "a general term that refers to the application of positive behavioral interventions and systems to achieve socially important behavior change" (Sugai, Horner, Dunlap, Hieneman, Lewis, Nelson, et. al, 2000). Originally applied to individual students with significant disabilities, the research and practice of PBS has expanded to include applications to entire school buildings and other complex organizations. At the school-wide level (SW-PBS), multidisciplinary teams of school personnel meet on a regular basis to design proactive, educative, and reinforcement-based strategies that will create an environment where appropriate behavior is more effective than inappropriate behavior (FL PBS Project, 2008). Similar to strategies recommended in culturally responsive classroom management, SW-PBS focuses on teaching students the behaviors that are expected of them, making the cultural code within a school building more explicit. The reinforcement-based strategies employed in SW-PBS achieve the dual outcomes of creating

positive interactions between students and staff, while teaching and encouraging appropriate behavior in students. Implementation of SW-PBS has been associated with decreases in office discipline referrals and suspensions, and improvements in school climate and academics (Barrett, Bradshaw, & Lewis-Palmer, 2008; Horner et al., 2009; Netzel & Eber, 2003; Putnam, McCart, Griggs, & Hoon Choi, 2009). Many of the outcomes of PBS, such as improved climate, more positive teacher-student relationships, and decreased problem behavior, have been proposed as ways of increasing students' academic engaged time, as well (Greenwood, Horton, & Utley, 2002).

Rationale for the Study

While positive behavior support and culturally responsive classroom management share many common strategies, the overlap between the two is not perfect. Proponents of culturally responsive classroom management recommend additional strategies that are not typically emphasized in the SW-PBS approach (Utley, Kozleski, Smith & Draper, 2002). Researchers have also documented how implementation of SW-PBS may produce a decrease in office discipline referrals and suspensions, but actually *increase* levels of disproportionality in those outcome measures (Skiba, 2007).

In light of research showing that schools which implement SW-PBS with high levels of fidelity experience stronger behavioral and academic outcomes (Cohen, Kincaid & Childs, 2007; Florida's Positive Behavior Support Project, 2009; Putnam, McCart, Griggs, & Hoon Choi, 2009), the current study investigated whether schools which implement SW-PBS with higher levels of fidelity tend to have lower levels of disproportionality in their office discipline referrals and out-of-school suspensions.

Purpose of the Study

The purpose of the present study was to investigate the rates of disproportionality in office discipline referrals (ODRs) and instances of out-of-school suspensions (OSS) for high-implementing and low-implementing PBS schools. Drawing from online databases which record and report schools' implementation and ODR information, this study examined the risk ratios for receiving an ODR and receiving an out-of-school suspension (OSS) for African American and

Hispanic students. This information was then compared to each school's reported level of SW-PBS implementation as measured by their Benchmarks of Quality score (BoQ; Cohen, et. al, 2007). The ensuing analysis provided insight into the relationship between the level of SW-PBS implementation fidelity and levels of disproportionality in ODRs and OSS for African American and Hispanic students.

Research Questions and Hypothesis

Research question one. What are the risk ratios for office discipline referrals and incidents of out-of-school suspensions for African American and Hispanic students in schools that implement SW-PBS?

Hypothesis one. Risk ratios for office referrals and suspensions for African American students will generally indicate rates of referrals and suspensions that are higher than would be expected given this group's distribution in their school's population.

Hypothesis two. Risk ratios for office referrals and suspensions for Hispanic students will be mixed, sometimes reflecting higher and sometimes reflecting lower rates of referrals and suspensions than would be expected given this group's distribution in their school's population.

Research Question two. Is there a relationship between the level of implementation of SW-PBS and levels of disproportionality in office discipline referrals and suspensions?

Hypothesis three. Schools that implement SW-PBS with higher levels of fidelity (as measured by a Benchmarks of Quality score of 70 or higher) will tend to have lower levels of disproportionate referrals and/or suspensions for African American and/or Hispanic students.

Significance of the Current Study

This study built upon existing research by examining rates of disproportionality for African American and Hispanic students in conjunction with schools' self-reported levels of SW-PBS implementation. Previous research involving SW-PBS and disproportionality (Skiba, 2007) has shown that levels of disproportionality may decrease following implementation, but has also shown that levels of disproportionality may actually *increase* following implementation, even as overall rates of referrals and suspension decrease for the school as a whole. The case studies described in prior research did not report to what degree SW-PBS was implemented, leaving the

question open as to whether improved implementation fidelity resulted in lower levels of disproportionality more frequently.

This study also added to the current field of knowledge by examining the range of disproportionality in schools that implement SW-PBS. With some proponents of culturally responsive classroom management turning to the literature base of SW-PBS for quantitative support for their strategies, it is critical to know whether SW-PBS implementation (and/or the quality of implementation) tends to be associated with more equitable distributions of referrals and suspensions. The in-depth descriptive analysis of the current study provided school-level data that allowed for a detailed picture of disproportionality in SW-PBS schools. This in-depth analysis also provided information on how the picture of disproportionality may be impacted by the availability and type of data used in schools' calculations.

Chapter Two: Review of the Literature

Introduction

Conversations about differences between racial groups can be difficult to negotiate, and conversations about racial inequities in education are no exception. It is not unusual for individuals to have feelings of anxiety, defensiveness, resistance, or hostility when discussing the topic (Ukpokodu, 2002), and these feelings can intensify when individuals misunderstand the intent of another's statements, or the background that gave rise to them. The following chapter attempted to address this by briefly noting the background leading up to the *Brown v. Board of Education* ruling and the actions school personnel engaged in following the decision. These events gave rise to a brief review of the current levels of disproportionality in academic achievement, special education placement and outcomes, and disciplinary actions and effects. Next, educators' attempts to explain the differences between racial groups were identified. This was followed by a discussion of strategies that are currently thought to offer promise in remedying the problem, including school-wide positive behavior support (SW-PBS).

Setting the Context: *Brown v. Board* and Desegregation

The 1954 *Brown v. Board of Education* ruling was a result of long-standing inequalities in public school education. Prior to the ruling, separate schooling and the disparate conditions that accompanied the practice existed in all regions of the United States, although many states from the North denied that segregation was ever tolerated (Douglas, 2005). Described as "grossly inadequate" (Fairclough, 2004), the school buildings assigned to African American students were often inadequately equipped and in poor repair. Accounts of post Civil War era public schooling reflect a long list of inadequacies in African American schools: hand-me-down textbooks, limited curricula, overcrowded classrooms, a shortened school year, overage students, inadequate transportation, unqualified teachers, significantly lower per-pupil expenditures for African

American students, and a system that supported the idea that by virtue of the color of their skin, some students were inherently inferior to others and could not compete in wider society (Fairclough, 2004; Horsford & McKenzie, 2008; Walker, 2001). For example, it was common for attendance at African American students' schools to vary according to whether a crop needed to be planted or harvested, and crop cycles have been cited as a reason for the shortened school year typical of some African American public schools (Horsford & McKenzie, 2008). Enrollment records show that in 1942, 30% of the African American students in the state of Georgia were enrolled in the first grade, evidencing the preponderance of overage students in the state at that time (Walker, 2001). Historians have extensively documented that African American teachers received less training, received substantially less pay than their White counterparts, were unable to participate in timely professional development, and were responsible for a variety of tasks that were unrelated to curriculum or instruction, such as maintaining their school's facilities (Walker, 2001).

At the same time, however, post Civil War era schooling for African American students was not without accomplishment. Private academies funded by the Freedman's Bureau, the American Mission Association, churches, philanthropists, and by the efforts of the local African American communities, created pockets of success for their students, as evidenced by dramatically higher rates of literacy across the South (Durham, 2003). Researchers have also described how separate schooling afforded a degree of community and positive climate among the students and their families (Durham, 2003). A report from a retired African American superintendent who attended a segregated school as a child and presided as an administrator over schools in the process of desegregation characterizes the difference between the two conditions:

> Within the segregated environment, it seems that there [was] a bit more nurturing going on …more ways where students would be inspired to achieve, rather than being relegated to some back room, or down in the basement, or becoming an untouchable within that environment with low expectations. And when people have low expectations,

they blame the victim rather than assume responsibility for their learning. (Horsford & McKenzie, 2008, p. 450)

In regards to teacher preparation, Walker (2001) describes the ebb and flow of professional training for teachers who were African American in the era leading up to the *Brown v. Board* decision. By 1950, college training and certification of African American teachers exceeded levels of training and certification of White teachers, with teachers who were African American receiving 4.1 years of college education compared to teachers who were White receiving 3.8 years of college education. In the years leading up to *Brown*, many districts increased the pay of African American teachers, which some scholars have interpreted as a gesture that was an attempt to appease advocacy groups and avoid federally legislated desegregation. Still, in spite of these gains and the success of some African American schools, the educational experience for an overwhelming majority of African American students was one of substantially less quality than that experienced by White students.

The 1954 *Brown v. Board of Education* ruling examined the notion of "separate but equal" as it might apply to public education, evaluating whether separate schools for African American and White students - given comparable facilities, curricula, teacher qualifications and the like - could ever provide for equal educational opportunities. Ultimately, the Supreme Court decided that they could not, citing a lower court's ruling that separate schooling affected a student's "…ability to study, to engage in discussions and exchange views with other students, and, in general, to learn his profession" (*Brown v. Board,* 1954). As a result of the ruling, school districts across the United States engaged in activities designed to blend the schools of African American and White students into one cohesive educational system.

In the first decade following *Brown*, a great deal of attention was given to "first-generation" desegregation issues, such as the racial distribution of schools' student populations. Sometimes by way of court order, districts instituted a range of strategies designed to improve the racial balance of their student bodies, including neighborhood rezoning and mandatory busing. Consequently, many African American schools were closed, African American teachers were laid

off or relocated, and African American students were bused to formerly all-White schools that were located outside of their immediate neighborhoods. Many times, African American students' arrival at formerly all-White schools was met with resistance and resentment by other students, teachers, administrators, and parents – bringing to light some of the historical "dirty laundry" of racial relations in the United States (Horsford & McKenzie, 2008, p. 447; Nelson, Palonsky, & McCarthy, 2004). Several states developed pupil placement laws, which some have argued gave local school boards a seemingly legitimate way to perpetuate separate schooling for African American students (Ferri & Conner, 2005). Pupil placement laws instructed local school districts to consider factors such as student preparation, moral character, conduct, and/or health in their decisions to place students in one school or classroom over another. In this way, many African American students did not receive the opportunity to be schooled with White students as ordered in *Brown v. Board*. Even in states that did not use pupil placement laws, many African American students were "tracked" into classes or programs that were intended for students with lower abilities and/or skill sets, limiting their access to the general education curriculum and instruction in higher-order thinking skills – a practice which continues today for many minority and low-income students (Mickelson, 2001). In fact, the use of tracking within the general education classroom and placement in special education has been applied so disproportionately with students of color that researchers have referred to this phenomenon as "within-school resegregation" (Blanchett, Mumford, & Beachum, 2005; Horsford & McKenzie, 2008).

The Current Landscape: Differences in Achievement, ESE Placement, and Punishment

More than fifty years after the *Brown v. Board* ruling, national achievement data reflect long-standing differences between White and minority students. National achievement measures illustrate a persistent pattern where African American and Hispanic students score lower than their White counterparts on standardized tests. Citing results from the National Assessment of Educational Progress (NAEP) that dates back to 1969, Nelson et. al (2004) described a consistent pattern of lower achievement for minority students, starting the first year the test was administered. During the early 1970's, African American and Hispanic students scored from 12 to 20 percent lower than White students at all grade levels, and across all subject areas. From the

1970's through the 1980's, the achievement gap between White and minority students narrowed. Since the 1990's, although all students experienced gains in test scores, the differences between White and minority students' scores have remained fairly stable (Nelson et. al, 2004). As of 2007 (the last year NAEP results are available), the gap between African American and White students was at least 26 points in all subject levels for 4th graders as well as 8th graders (Vanneman, Hamilton, Baldwin, & Rahman, 2009). The gap between Hispanic and White students is similar in magnitude, translating into a 9-12% difference between the average scaled score of White students and the average scale score of minority students at both grade levels (see Table 1; Lee, Grigg, & Dion, 2007; Lee, Grigg, & Donahue, 2007). Given these findings, it is clear that in spite of educators' attempts to integrate their schools, differences between White and minority students' educational experiences continue today.

Table 1

2007 NAEP Average Scale Scores

	Mean Scale Score Reading *(Percent Difference compared to White Students)*			Mean Scale Score Math *(Percent Difference compared to White Students)*		
	White Students	African American Students	Hispanic Students	White Students	African American Students	Hispanic Students
4th Grade	230	203 *(-12%)*	205 *(-11%)*	248	222 *(-10%)*	227 *(-9%)*
8th Grade	270	244 *(-10%)*	247 *(-9%)*	290	259 *(-10%)*	265 *(-9%)*

Note. Reading scores contained in this table may be found in Lee, Grigg, & Donahue (2007), and math scores may be found in Lee, Grigg, & Dion (2007).

Special education placements also may be examined as a domain of experiential discrepancy. Reports from the Office of Civil Rights show that minority students have been placed in some disability categories at a disproportional rate since the 1970's, and literature exists that notes the presence of disproportionality in special education (ESE) placements even further back

in our nation's history (Ferri & Conner, 2005). Currently, data from the National Center for Education Statistics (NCES) show that overall ESE placements (aggregated to account for all disability types) differ only slightly for each racial group when compared to each group's makeup in the K-12 population (see Figure 1). This is in line with what could be expected, in the absence of data to indicate that one racial population is more likely to have a disability than another. Of all students who received ESE services in 2004 (the most current year for which data are available), approximately 60 percent were identified as White, approximately 20 percent were identified as African American, and approximately 16 percent were identified as Hispanic (approximations were used due to limitations in reporting; see Appendix A). These approximations are similar to each group's representation in the K-12 student population, which for that year was calculated to be 58% White, 17% African American, and 19% Hispanic. However, differences between racial groups emerge when membership in specific disability categories is considered (see Figure 2). The percentage of students who were identified as White and receiving services for Specific Learning Disabilities (SLD), Emotional and Behavioral Disorders (EBD), and Visual Impairment (VI) was close to what would be expected given their percentage in the K-12 population for that year, while the percentage of students identified as White and receiving services under Mental Retardation (MR) was noticeably less. For instance, approximately 57% of the students receiving services for SLD were identified as White, which is similar to the percentage of White students in the K-12 population (58%). However, in the category of Mental Retardation (MR), only 51% of students receiving services under that label were identified as White, which falls noticeably short of White students' composition in the K-12 population. For students who were identified as African American, membership in each of the main disability categories was higher than would be expected, with the exception of the category of Visually Impaired – a category which is defined by specific, objective, physiologically-determined criteria. For students who were identified as Hispanic, membership in each of the listed disability categories was less than would be expected, with the exception of the category of SLD, which was similar to Hispanic students' makeup in the K-12 population (19%). It is noteworthy to consider that disproportional placement is most evident in the disability categories for Mental Retardation (MR), and Emotional/Behavioral Disordered

(EBD), which have been criticized as relying on relatively subjective criteria (Blanchett, 2006; Green, 2005; Skiba, Michael, Nardo, & Peterson, 2000). The category of Visually Impaired (VI), for which membership is decidedly objective, shows much less deviation from what would be expected given the different groups' distribution in the K-12 population.

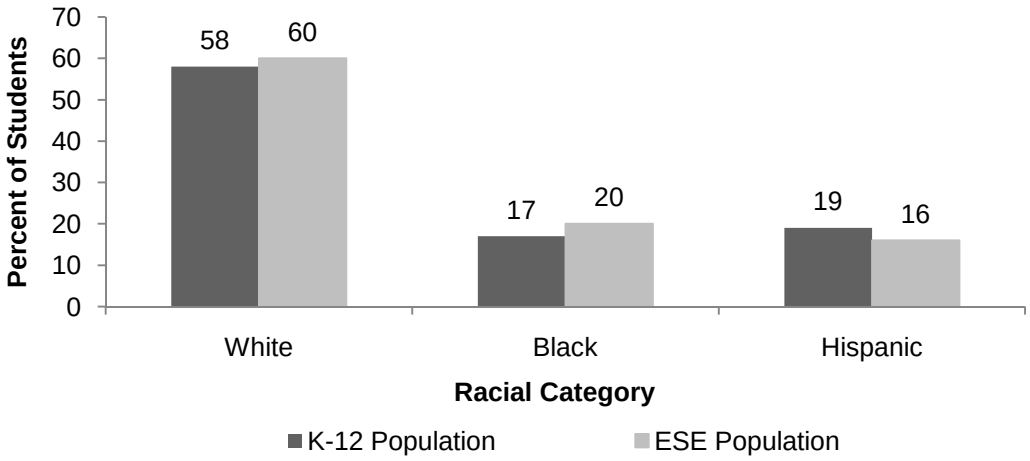

Figure 1. Comparison of percent of students enrolled in Special Education by race to percent of students in K-12 general education by race, 2004.

Note. Based on KewalRamani, Gilbertson, Fox, & Provasnik (2007). Percentages represent an estimation of the ESE population. See Appendix A for additional information.

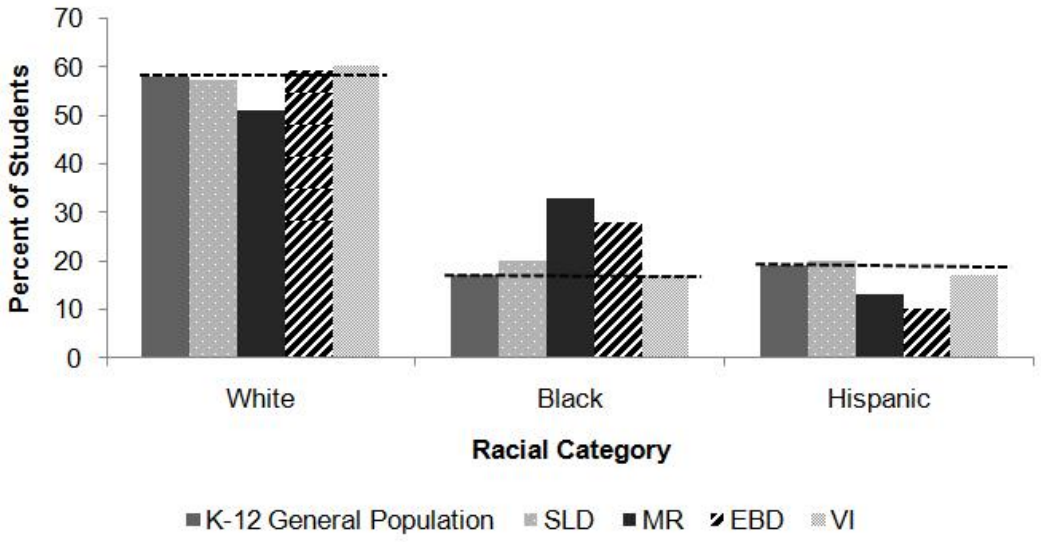

Figure 2. Comparison of enrollment in specific ESE categories by race, compared to racial composition in the general student population by race, 2004.

Note. Based on KewalRamani, Gilbertson, Fox, & Provasnik (2007). Percentages represent an estimation of the ESE K-12 population. See Appendix A for additional information.

The type of ESE label a student receives can be significant in that it may impact the amount of time s/he spends in the general education environment - the setting that is most likely to focus on the core academic competencies reflected in standardized tests and graduation requirements. A recent government report documented the amount of time students with different disabilities spent in receipt of ESE services outside of the general education classroom (U.S. Department of Education, 2007). More than half of students with the label of Mental Retardation (MR) spent more than 60% of their school day receiving services outside of their general education classroom, the highest of any ESE category. Students identified as having multiple disabilities were the next most likely to be removed for the majority of their school day (45%), followed by students with Autism (43%), students with Deaf-Blindness (33%), and students with Emotional or Behavioral Disturbances (30%). Perhaps in part due to their higher rate of identification in the categories of MR and EBD, the report also found that African American students were more likely than any other racial group to be removed from the general education

classroom for the majority of the school day (28%), or educated in a separate environment all together (5%; Blanchett, 2006; Losen & Orfield, 2002).

The likelihood of receiving a suspension or expulsion has been shown to vary by type of special education label as well, particularly for disability types that reflect a high degree of behavioral issues. While it is likely that a student's underlying behavioral needs, rather than the label itself, was responsible for variations in suspensions and expulsions, it is nevertheless noteworthy that special education students with the label of EBD are the most likely of students with any other disability type to be suspended or expelled from school. Achilles, McLaughlin, and Croninger (2007) examined data from the Special Education Elementary Longitudinal Study (SEELS), and found that even after controlling for different economic and social factors (such as socioeconomic status and family structure), students with the label of EBD were more likely than children labeled with Other Health Impairments (with a history of ADHD) or with the label of LD to have been suspended or expelled from school. The authors also found that African American ethnicity, age, gender, low socioeconomic status, multiple school changes, urban schooling, and low parent satisfaction were also significantly correlated with exclusionary discipline – findings which were also reported in reviews from other studies (Christle, Nelson, & Jolivette, 2004). Thus, students with a label of EBD – especially African American students who have this label – may be at greater risk of being removed from the general education classroom not only for instructional purposes, but also as a result of disciplinary actions.

The over-identification and removal of students from the general education classroom for the provision of ESE services might be less problematic if these actions resulted in improved outcomes for students. Unfortunately, this has not been the case. A recent report from the U.S. Department of Education examined outcomes for students with disabilities during the first four years after high school, and revealed that less than half of all special education students (45%) continued on to postsecondary education, compared to 53% of their non-disabled peers (Newman, Wagner, Cameto & Knokey, 2009). Youths who were labeled as having Emotional or Behavioral Disturbances (EBD) were found to have some of the poorest post-high school outcomes, with only 34% continuing their education beyond high school, 60% being arrested, and

39% spending a night in jail. When looking at outcomes for students with disabilities from different racial groups, the report concluded that there were no significant differences in the likelihood of "being engaged in school, work or preparation for work" between different racial categories, but also reported the seemingly contradictory finding that White students with disabilities were almost twice as likely to have been employed as were African American students with disabilities (80% versus 47%). Thus, depending on ESE category, and in some cases race, a student who receives an ESE label may be more likely to experience poorer post-high school outcomes than students who do not receive special education services.

Unfortunately, placement in special education is not the only way students of color are removed from the general education environment. Disproportionality in school punishment has been a highly consistent finding for more than 30 years, particularly for African American students (Fenning & Rose, 2007). Current national data supports this finding, showing that African American students are suspended or expelled more often than students of any other racial category (KewalRamani, Gilbertson, Fox, & Provasnik, 2007). Data from 2003 showed that almost 20% of African American students across the country were suspended, compared to almost 9% of White students, and just over 10% of Hispanic students.

Numerous studies have reported rates of suspension and expulsion that were two to more than three times higher for African American students than their White peers (Baker, Hendricks, McGowan & McKechnie, 2004; Gordon, Piana & Keleher, 2000; Skiba, Michael, Nardo & Peterson, 2000). In a summary of 12 separate studies which spanned over 20 years, Skiba et al. (2000) examined the overrepresentation of minority students in school suspensions and expulsions. In each study, percentages of African American students who received suspensions or expulsions exceeded their distribution in the schools' population by anywhere from 27 to 114 percent. In an analysis of statewide data from Maryland, Krezmien, Leone, and Achilles (2006) found that not only were rates of suspensions for African American students disproportionate, but their odds of being suspended *increased* over the eight years for which data was available, while the odds of being suspended for White and Hispanic students remained approximately the same. In addition, the authors found that African American students identified

as having any disability except for Other Health Impaired had the greatest risk for suspension out of any other group of students from any racial or disability category. In other words, in these authors' sample, a African American student who had reading difficulties (labeled SLD) was more likely than a White or Hispanic student with emotional or behavioral challenges (labeled EBD) to be suspended.

It appears that a disproportionate rate of suspension may begin early in a student's school career, as well. Raffaele-Mendez and Knoff (2003) analyzed out of school suspensions in one large Florida district, and found that overrepresentation of African American students began at the elementary school level and peaked at the middle school level, where almost 50% of African American male students experienced at least one suspension. This imbalance continued in high school, where although rates of suspensions for all groups of students declined compared to middle school, rates for African American students were still more than twice as high as for White students. Notably, the authors in this study also revealed that the rates of suspension for African American female students at the middle and high school levels (31.88% and 20.69%, respectively) were higher than the rates of suspensions for White *males* at those levels (25% and 18.9%). This pattern was also found in national data from 2003, where 15.2% of African American females across all grade levels were suspended, compared to 12.7% of White males (KewalRamani et al., 2007), and is particularly striking given the customary finding that male students experience disciplinary actions (including suspensions) at a much higher rate than female students (Cartledge, Tillman, & Talbert Johnson, 2001; Skiba et al., 2000).

Continuing their analysis, Raffaele-Mendez and Knoff (2003) noted that the percentage of suspensions accounted for by White male students was similar to their percentage in the overall student population (30%) for 11 out of 15 different rule violations. The offenses that did not fit that pattern included Tobacco Possession (64%), Weapons Possession (48%), Narcotics Possession (55%), and Alcohol Possession (48%). For African American male students, the findings were reversed: although they accounted for only 12% of the total student population, African American male students received 24-48% of all suspensions for every offense *except for* Tobacco Possession (5%), Narcotics Possession (15%), and Alcohol Possession (14%). These data

indicate that depending on the offense, students from different racial groups received a higher percentage of suspensions than would be expected given their makeup in the general student population. However, these findings do not reveal whether students from different racial groups committed offenses at a rate or intensity that was different than their peers (i.e., if African American male students engaged in "Sexual Harassment" more frequently or more severely than White male students). With this in mind, it is interesting to note that the greatest degrees of disproportionate suspensions for African American students were identified for offenses that are defined by relatively subjective criteria (Sexual Harassment [48%], and Threat/Intimidation [43%]) – and that where White students were over-represented, it was for offenses that were much more objectively defined.

Earlier studies have investigated the question of whether African American students engage in higher rates of rule violations or more severe problem behaviors than other groups of students. Analyzing office discipline referral data from middle schools in a Midwestern school district, Skiba et al. (2000) determined that African American students not only received higher rates of more severe consequences (such as out-of-school suspensions), but African American students were also more likely to receive these consequences for less severe behavior. According to their analysis, White students were more likely to be referred to the office for smoking, leaving without permission, and vandalism; while African American students were more likely to be referred for disrespect, excessive noise, threat, and loitering. Skiba and his colleagues point out that not only were these offenses less severe than the offenses White students were most likely to be referred for, but the offenses African American students were referred for were decidedly more subjective in nature. For example, when a student is referred for smoking, there is clear evidence of the offense (such as a cigarette). However, when a student is referred for disrespect, there is ample room for personal interpretation of ambiguous or neutral behaviors – one person's perception of "disrespectful" behavior could be another person's perception of a failed attempt at humor or mild expression of disgust. Based on this pattern, it seems that when there is more room for school personnel to interpret behavior, there may also be greater likelihood of disparities in discipline – a theme that echoes patterns of disproportionality in

different ESE categories. When concrete and objective criteria play a smaller role in school personnel's decision making, the risk for inequitable outcomes is heightened.

The consequences of disproportional punishment can be far-reaching. Numerous studies have documented decreases in school attachment and increases in grade retention, dropout, and contact with the criminal justice system for students who have been suspended, especially for repeat offenders (Christle et al., 2004; Costenbader & Markson, 1998; Gordon et al., 2000; Skiba & Rausch, 2006). Just as research shows that African American students are suspended at a higher rate than students from other racial categories, recent national data showed that African American students were also more likely than students from any other racial category to experience grade retention and had the second-highest rates of high school dropout (10.4%; Hispanic students had the highest rates of dropout, at 22.4%; KewalRamani et al., 2007). The extensive documentation of the overrepresentation of minority adults in the prison population and in the percentage of minority adults who live at or below the poverty line, further illustrates the potential lifelong consequences of allowing disparities in education to go unchecked.

Responding to the Data: Educators' Attempts to Explain Disproportionality

Numerous studies have investigated the potential causes of these academic and discipline-related disparities, many with an eye towards identifying strategies that could improve them. Farkas (2003) provides an extensive review of studies that have investigated the different variables proposed to account for educational inequities. Among the different explanations, the far-reaching effects of poverty are frequently hailed as accounting for numerous causational and maintaining factors. To start with, children from low socioeconomic status (SES) families typically start school with fewer language, pre-reading, pre-math, and behavioral skills than children from middle class or higher-SES families. These early differences translate into what Farkas refers to as a "Matthew effect," where early differences in skills compound in a cumulative fashion across a student's educational career, resulting in large differences in skills in later grades. The effects of poverty can further add to this process through summertime "fall back," where lower SES students may experience a decline in academic skills during the summer months due to a less-enriched vacation environment. Community variables in poverty-stricken neighborhoods –

violence, arson, and change of housing, to name a few – inflict additional challenges that can affect achievement and behavior. Schools in poverty-stricken neighborhoods also are frequently faced with deteriorating facilities and out-of-date academic equipment and materials (computers, school books, etc…), which further impact the learning process. When all of these challenges are combined, it's not surprising that many teachers would be reluctant to work in this environment – a sentiment which reflects the reality facing many urban, low-SES schools today where "highly-qualified" teachers are in short supply. Given the wide range of variables impacted by poverty and the high degree of overlap between socioeconomic status and minority group membership, it is not surprising that conversations about deficits caused by socioeconomic status are commonly used as a substitute for conversations about the disparate educational experiences of minority students (Kunjufu, 2006). In other words, educators seem to feel more comfortable talking about issues associated with poverty that can cut across any race, rather than confront issues that deal directly with differences between two or more races.

Certainly, poverty and minority status are intertwined. Historically, minority families in the United States have been more likely than White families to live at or below the poverty line. This holds true today: the latest figures from the U.S. Census Bureau reveal that the poverty rates for African American and Hispanic families were 24.7% and 23.2%, respectively; the poverty rates for White and Asian families were 8.6% and 11.8% (De-Navas-Walt, Proctor, & Smith, 2009). Loosely translated, this indicates that a person who is African American or Hispanic is more likely to live at or below the poverty line in the United States than is a person who is White or Asian. Considering, too, that segregated housing patterns (e.g., "White flight" to suburban neighborhoods) have resulted in high concentrations of minority families in poverty-stricken neighborhoods, many educators – and some researchers – look to poverty to explain existing educational disparities between students of different races. According to that view, differences in behavior and achievement that are found between students of different races can be better accounted for by SES status than by the students' race. This approach has been popular with many educators, who seem to be more comfortable talking about behavioral and learning differences they observe in "students from a low-SES background" than they are in talking about

behavioral and learning differences they observe in students from African American or Hispanic families (Kunjufu, 2006). Unfortunately, this displacement creates a “red herring” for effective intervention. While poverty can certainly play a role in racial differences in educational experiences, it does not fully explain the academic or behavioral gap. Skiba and his colleagues (2005) outline the shortcomings of the poverty-disproportionality link along several fronts, among them the finding that risk of placement in special education may *increase* as a student’s level of poverty *decreases*. For example, disproportionate representation in special education was more extreme in schools that drew from high socioeconomic status neighborhoods than it was in schools that drew from low socioeconomic status neighborhoods. Skiba et al. also note that disproportionality in special education exists in the “soft” disability categories of LD, MR, and EBD, but not in the more objectively defined categories of visual or hearing impairment. Further, Skiba et al. point out that although poverty rates for African American and Hispanic students are similar, findings of disproportionality in special education placement and disciplinary actions vary substantially between the two races.

In addition to their logical rebuttal of the poverty-disproportionality link, Skiba and colleagues offer statistical evidence of the weakness of this explanation. Employing a logistic regression to analyze district-level data, Skiba et al. showed that poverty was a “weak and inconsistent” predictor of African American students’ placement in special education, and only served to magnify existing racial disparities. Skiba’s group also revealed that the strongest predictor of disproportionality in special education was the district’s rate of suspension/expulsion, such that high district rates of suspension and expulsion were accompanied by disproportional rates of African American students’ placement in special education. In light of those results, it seems that time spent out of class for disciplinary action may be a more viable avenue for intervention than a student’s socioeconomic status. It is much more realistic for educators to impact how much time a student spends in their classroom than it is for educators to impact a student’s poverty level.

The time that a student spends outside of class for disciplinary (or other) reasons reduces the amount of time a student may spend engaged in the learning process – a variable

which directly impacts academic achievement as well as student behavior. Unlike socioeconomic status, academic engaged time (AET) can be directly impacted by an educator's actions, which provides opportunities for intervention. Academic engaged time can be thought of as the "portion of instructional time that students spend directly involved in learning activities," (Johns, Crowley & Guetzloe, 2008), and can include activities such as listening attentively to instruction, working on assigned materials, asking for assistance, or having an instructionally-relevant discussion with one's peers. When students are academically engaged, they are more likely to benefit from instruction (i.e., they are more likely to learn the material), and are less likely to engage in behaviors that lead to disciplinary consequences (such as removal from the classroom). Interventions that aim to increase students' AET have been shown to be effective at improving achievement (Brophy, 1988; Greenwood, Horton, & Utley, 2002), as well as student behavior (Gilbertson, Duhon, Witt, & Dufrene, 2008; Johns, Crowley, & Guetzloe, 2008; McComas, Hoch, Paone, & El-Roy, 2000), leading some researchers to conclude that instructional strategies which are based on increasing students' AET are more likely to be effective in improving outcomes than instructional strategies which are based on a student's perceived deficit or difference (see Brophy, 1988, for a review).

As outlined in previous sections, it is clear that there are several avenues through which African American and Hispanic students' academic engaged time has been, and continues to be, unevenly impacted. Historically, separate schools with shortened school calendars provide one example, while current day documentation of disproportionate tracking, ESE placement, and disciplinary action provides evidence that in many cases, minority students spend less time engaged in the general education curriculum than their White peers. In addition to these mechanisms, AET may also be impacted by systems-level variables that disproportionately affect minority students. In one example, Smith (2000) analyzed the amount of instructional time allocated to inner-city Chicago schools compared to the amount of instructional time that actually took place within their classrooms. In her analysis, Smith found that the inner-city schools operated on a daily schedule that was 30-45 minutes shorter than a majority of other schools in the country. The inner-city schools utilized a compressed schedule that ignored the amount of

time needed to address routine tasks and needs, such as taking attendance, lining up, and providing bathroom breaks. While these activities may not seem significant, other studies have found that they can account for significant portions of class time. For example, Fisher (2009) reported that 17% of class time in a suburban high school was used for activities such as these. Smith also noted that time needed to address community traumas (such as gang-related shootings), review previously learned material, accommodate special events, and deal with daily exceptions to routines (such as cut fingers, lost textbooks, etc…), was likewise unaccounted for by the master schedule. Classroom management practices were found to impact instructional time, with poorly managed classrooms experiencing approximately 20% less time for instruction than well-managed classrooms – a notable difference, but one that the author felt was dwarfed by the limitations imposed by the master schedule. In all, Smith estimated that in most cases, inner-city Chicago schools were able to utilize only 60% of the instructional time allocated to them by their district. Although the results of Smith's study may not be representative of every school system, it serves as an important illustration of why educators should consider the potential impact of district and school-level variables on student outcomes, a sentiment shared by other researchers (Skiba, Simmons, Ritter, Gibb, Rausch, Cuadrado & Chung, 2008).

In addition to the often-referenced arguments about poverty and the academic and behavioral gaps, cultural differences between students and teachers have also been argued to play a role in the diminished achievement and behavioral outcomes of minority students. In some cases, the difference that is cited is placed squarely on the shoulders of the students who are not making progress. Supporters of this argument say that minority students do not place a high value on education, and they reject behaviors that might make them appear to want to be successful in school. Using survey data from 10th grade students from the National Educational Longitudinal Study (NELS) and the National Assessment of Educational Progress (NAEP), Farkas, Lleras, and Maczuga (2002) found that in high-minority schools (schools where White students make up less than 25% of the student body), African American students were more likely than their White peers to report being "put down" for being a "very good" student. In an additional analysis, they also found that among fourth-graders, students in Title 1 schools were

also more likely to have peers who discouraged academic effort than were students in non-Title schools. Overall, citing findings from other studies in addition to their own, the authors concluded that there is evidence to support an "oppositional culture" towards school effort among African American students, particularly in high-minority and high-poverty schools. Although the framing of this issue is fairly one-sided (blaming the students for an "oppositional culture"), peer pressure for students to behave in ways that do not support academic achievement or a positive classroom atmosphere could be an additional barrier teachers and students must face along the path to equitable outcomes.

Other researchers take a more balanced approach to this argument, noting the differences between a teacher's culture and their minority students' culture ("cultural mismatch") as a source for conflict in the classroom. Vavrus and Cole (2002) provided a qualitative study of the interactions that took place in classrooms just prior to a teacher's decision to suspend a student. In their analysis, the researchers detail an interaction between a student and teacher in which the teacher uses common Hispanic names as part of a reference to the difficulty of identifying students' nameless papers ("You know how many Joses and Marias are there with no names?" [p. 107]). A student in the class took exception to the teacher's choice of words, challenging the teacher's use of "Spanish" words in a science class. Several students joined the conversation at this point, and the student who originally challenged the teacher was eventually removed from class. The authors argue that the decision to remove the student because of a relatively harmless statement ("I said we is not bilingual. We cannot understand what you say." [p. 107]) begins to seem justified only when considering the sociocultural context of the classroom – namely, the differences between the teacher's and students' background. In this example, the teacher reported being particularly upset by the student's statements because she had intentionally chosen to use names that would not be considered stereotypical of the African American students in her class – an effort that unfortunately resulted in misunderstanding and conflict.

Further, the researchers posit that specific interactions between students and teachers are complicated by the teacher's perception of their degree of control within the classroom. In this

study, interactions that occurred within a series of events that appeared to challenge the teacher's authority over their students were interpreted more severely than interactions – even disrespectful ones – that occurred in isolation. In other words, a disrespectful comment may be overlooked – or even responded to positively (with a smile, for example), if the teacher feels that they still have control over their students. The authors argue that teachers' perception of control is often influenced by "unspoken and unwritten rules of linguistic conduct" (p. 91), such as expectations for silence, formal address, or turn-taking during conversations. Because these rules (or expectations) are unspoken, there may be more opportunities for misunderstanding and potential disciplinary reaction (as in the example above) in situations where the backgrounds of students and teachers are very different.

This assertion is supported by a recent literature review on the cross-cultural issues involved in classroom discipline. Pane (2009) provides numerous examples supporting the miscommunications, misunderstandings, and conflicting expectations about classroom behavior that arise as a result of differing cultural backgrounds. Drawing from the body of research she described, Pane asserts that instances of classroom discipline are socially negotiated and heavily influenced by culture. This conclusion appears reasonable, given the high rates of disproportionate punishment for subjectively-defined behaviors (such as disrespect, threat/intimidation, etc…) noted in previous sections (Raffaele-Mendez & Knoff, 2003; Skiba et al., 2000). The key to reducing exclusionary discipline, Pane argues, is to improve the cultural competency of teachers.

Additional support for a theory of disproportional discipline as a result of cultural mismatch comes in the form of research examining disciplinary interactions between students and a teacher whose cultures are relatively similar. In their 2004 study, Monroe and Obidah noted the importance of the *context* of student-teacher interactions, maintaining that behaviors which might seem inappropriate in one setting might be considered appropriate in a different situation. The authors posited that when individuals' cultures were similar, the contextual subtleties of social interactions would be better understood by both parties, making it less likely that a comment would be misinterpreted as an act of disrespect or disruption. Certainly, the authors'

analysis seemed to support this hypothesis: of the 387 student actions researchers recorded as "disruptive," only two instances (0.5%) resulted in an office discipline referral. Conversely, the teacher used humor and culturally-based dialect to address 112 (29%) of those student actions. The authors asserted that the teacher's familiarity with her students' backgrounds enabled her to recognize behaviors that weren't intended to be troublesome, allowed her to feel secure and "in control" of her class, and respond in a manner that would promote a positive connection with her students. Teachers who were not familiar with their students' culture, on the other hand, could be more likely to misinterpret their students' actions, and potentially escalate a situation to the point that it requires removal from the classroom. Monroe and Obidah made a point of stating that teachers whose backgrounds were not similar to their students could achieve the same results as the teacher in their study, so long as they took care to "build cultural bridges" (p. 258) between themselves and their students.

Avenues of Hope: Promising Strategies to Reduce Disproportionality

The process of 'building cultural bridges' is reflected in the practice of culturally responsive education. Ladson-Billings (1995) built on Irvine's 1990 concept of "cultural synchronization" by defining culturally responsive pedagogy as "a theoretical model that not only addresses student achievement but also helps students to accept and affirm their cultural identity while developing critical perspectives that challenge inequities that schools (and other institutions) perpetuate" (Ladson-Billings, 1995, p. 469). Rather than dismiss cultural differences between individuals through a "colorblind" ideology, culturally responsive education aims to convey recognition and respect for all cultures while acknowledging the systemic inequalities that individuals from some groups may face. Teachers who approach instruction through a culturally responsive framework tend to produce higher levels of academic achievement than would be expected, as well as students who are adept at critical thinking and problem solving, and academically engaged (Ladson-Billings, 1995).

Applying the ideals of culturally responsive pedagogy to the realm of behavior management, Weinstein, Curran, and Tomlinson-Clarke (2003) described culturally responsive classroom management (CRCM) as being a "state of mind as much as a set of strategies and

practices" (p. 275). According to the authors, it requires an understanding of "the self" (the role an individual's own culture plays in behavior and understanding), "the other" (the differences that exist between individuals' cultures) and the context (how system-level variables and differing cultures may interact to impact behavior and understanding). Culturally-Responsive classroom managers must also have a willingness to reflect on the ways that personal choices for classroom management techniques can either facilitate or impede a student's engagement with the learning process. There are numerous examples of specific practices that are reflective of CRCM, most of which are consistent from expert to expert. Those practices are summarized in the following categories (Bazron et al., 2005; Bondy, Ross, Gallingane, & Hambacher, 2007; Cartledge & Kourea, 2008; Delpit, 1995; Ware, 2006; Weinstein et al., 2003; Zirkel, 2005):

- Ongoing analysis of academic and behavioral data coupled with responsive decision making and early intervention;
- Fostering positive and caring student-teacher relationships and school climates;
- Designing positive, preventative and proactive discipline strategies;
- Teaching clear expectations for behavior;
- Communicating high expectations for student achievement;
- Utilizing effective models of explicit instruction (which include high rates of academic responding, appropriate pacing, purposeful movement, cooperative/community structure, and timely feedback);
- Becoming familiar with students' backgrounds and values, and using this knowledge to foster positive relationships, improve curricula and instruction, and mediate interpretations of behavior;
- Promoting parent involvement;
- Considering how the personal culture and biases of "power-holders" (e.g., teachers, administrators, and community leaders) may impact those individuals' evaluations of behavior.

Although there is a scarcity of quantitative research on CRCM in the current literature (Bondy et al., 2007; Cartledge & Kourea, 2008), qualitative studies report that teachers who employ CRCM

strategies tend to have classrooms with positive atmospheres, positive relationships with their students, and high levels of academic engagement and achievement (Bondy et al., 2007; Ware, 2006).

Given the long-standing nature of educational disparities in both academic achievement and discipline, the current lack of empirical evidence for CRCM is troubling. While there are studies showing the effectiveness of isolated CRCM strategies for minority students' academic outcomes (such as the effect of class-wide peer tutoring on the DIBELS scores; Kourea, Cartledge, & Musti-Rao, 2007), there is little to no specific research on the effectiveness of CRCM strategies for behavioral outcomes. Instead, Cartledge and Kourea (2008) point to the literature base on positive behavior support (PBS) as being "particularly appropriate for CLD (culturally and linguistically diverse) learners" (p. 363); and indeed, many elements of PBS align with the practices recommended for CRCM. However, while the theoretical underpinnings of PBS would seem to lead practitioners to use CRCM strategies, the real-life application of PBS may not always result in a culturally responsive model of implementation.

As defined by the Office of Special Education Programs' (OSEP) Center on Positive Behavioral Interventions and Support, PBS is "a general term that refers to the application of positive behavioral interventions and systems to achieve socially important behavior change" (Sugai, Horner, Dunlap, Hieneman, Lewis, Nelson, et. al, 2000). Developing out of the field of behavior analysis and originally applied with individuals who had significant disabilities, PBS has expanded to include applications with entire school buildings and other complex organizations.

At its core, PBS encompasses four main features: (1) the application of behavioral science; (2) the use of practical, multi-component interventions that can be implemented by stakeholders in natural contexts; (3) a focus on person-centered, values-driven lifestyle changes (as opposed to the simple elimination of problem behavior); and (4) an emphasis on systems-level variables that affect stakeholders' ability to implement interventions effectively (Dunlap, Sailor, Horner & Sugai, 2009). When applied within school buildings, PBS utilizes a multi-tiered framework of supports: Tier 1 (primary, or universal) supports, designed for all students and staff in all settings; Tier 2 (secondary, or supplemental) supports, designed for targeted settings and

groups of students who need support beyond what is provided at the primary level; and Tier 3 (tertiary, or individual/intensive) supports, designed to meet more intensive individualized needs. The first tier of support, sometimes referred to as school-wide PBS (SW-PBS), is essential to creating an environment that supports effective teaching, and in providing a foundation for more intensive supports. Sugai and Horner (2009) identify six major features of SW-PBS:

1. An approach to discipline that is based on behavior analytic principles, reflects evidence-based practices, is culturally appropriate, and based on active instruction;
2. A few positively stated expectations are established to guide the behavior of all staff and students in all settings, which convey support for academic and behavioral outcomes and are culturally appropriate;
3. The school-wide expectations are explicitly and continually taught;
4. A system exists to encourage students and staff to follow the school-wide expectations (i.e., a reward system for appropriate behavior);
5. A continuum of consequences for inappropriate behavior exists that is clearly defined, includes prevention strategies and options for more intensive instruction and support, is regularly monitored for effectiveness, and is based on the function of the student's behavior;
6. A data management system exists that is accurate, up-to-date, easy to access/use, and useful in guiding decisions about environmental and behavioral supports.

Studies evaluating the effectiveness of SW-PBS have noted decreases in office discipline referrals (ODRs; Barrett, Bradshaw, & Lewis-Palmer, 2008; Bohanon et al., 2006; Lohrmann-O'Rourke et al., 2000; Warren et al., 2006), decreases in out-of-school suspensions (OSS; FLPBS Project, 2009; Netzel & Eber, 2003; Warren et al., 2006), increases in student time in school (Scott & Barrett, 2004), increases in attendance (Luiselli, Putnam, & Sunderland, 2002), improved school climate (Putnam, McCart, Griggs, & Hoon Choi, 2009), and improvements in academics (Horner et al., 2009).

School-Wide PBS has been successfully implemented in urban schools, as well. Urban schools are typically characterized by higher rates of poverty and crime, larger student enrollments with greater diversity and more risk factors among students, more staff turnover, less qualified staff, and are frequently considered lower-performing (Putnam et al., 2009). In spite of these challenges, implementation of SW-PBS in these settings has been followed by improvements in student outcomes. Netzel and Eber (2003) implemented SW-PBS in an elementary school where the 96% of the student body was identified as being of minority status, and 68% of the student body qualified for free-and-reduced lunch. After one year of implementation, the school reported a 22% reduction in suspensions. Putnam et al. (2009) reported a 2007 study undertaken by Rey, Their, Handler, and Putnam involving several urban elementary and middle schools. All schools had populations characterized by high percentages of minority students (approximately 90%) and approximately 75% of students qualified for free-reduced lunch. Schools that had higher percentages of students who could state the school-wide rules reported larger decreases in out-of-school suspensions than schools that had lower percentages of students who could state the rules. McCurdy, Mannella, and Eldridge (2003) described SW-PBS implementation in a school whose population was identified as being comprised of 44% Asian/Pacific Islander students, 33% African American students, 18% European American students, and 5% Latino American students. The authors in this study reported “a high percentage” of students who were eligible for free-reduced lunch (p. 160). By the end of the second year of implementation, the school reported a statistically significant 46% reduction in ODRs compared to their baseline year (the year prior to implementation); an even larger decrease (55%) was reported for the specific offense of “fighting.” Bohanon et al. (2006) studied SW-PBS implementation in an urban high school of 1,800 students representing more than 75 countries. The student body composition was 36% African American students, 36% Hispanic students, 16% Asian American students, 8% Caucasian students, 2% Native American students, and 2% “Other” students. Approximately 89% of the student body qualified for free-reduced lunch. In spite of the additional challenges inherent in implementing SW-PBS in a high school, Bohanon et al. reported a 20% reduction in average daily ODRs during the third year of

implementation, along with larger decreases in ODRs for the specific offenses of “Dress Code” and “Serious Disobedience of Authority.” Across the three years of the study, a statistically significant decrease in the proportion of students who received more than two ODRs was also documented. Lassen, Steele, and Sailor (2006) reported the results of a three-year study of SW-PBS implementation in an urban middle school. The school’s enrollment was made up of 623 students; 26% of the students were identified as African American, 40% Hispanic, 30% White, and 4% Asian Pacific Islanders. Approximately 80% of the students qualified for free-reduced lunch. Over the three years of the study, the authors reported a statistically significant decrease in ODRs per student and long-term suspensions compared to baseline (the authors used an alpha of .025). Mean scores on standardized math and reading tests also increased by year three of the study, although the increases in reading scores were not statistically significant.

Clearly, SW-PBS offers promise in improving behavioral and academic outcomes in schools with large percentages of minority students. Although it hasn’t yet been established whether one (or more) specific components of SW-PBS has a greater influence on student outcomes than others (Peshak-George, Kincaid, & Pollard-Sage, 2009), there are several avenues through which the SW-PBS approach, when implemented with fidelity, may impact ODRs, suspensions, and achievement. Perhaps most obviously, fewer ODRs and suspensions increase the likelihood that students will spend more time in their classrooms, where they may benefit from instruction. The focus on consistent, appropriate discipline, as well as the emphasis on developing systems to acknowledge appropriate behavior, helps to create a more positive school climate and student-teacher relationships, which may boost a student’s academic engagement and attendance. Ongoing teaching of school-wide expectations and rules helps students understand the kinds of behaviors that will help them be successful at school, which may cut down on the frequency of some culturally-based behaviors that can lead to misinterpretation and conflict. Ongoing data analysis and proactive intervention helps to ensure that behavior problems are caught early, before they become intensive or severe, and also helps school teams identify when interventions are not appropriate or effective for their students and change course. The collaborative nature of SW-PBS (Peshak-George et al., 2009; Sugai &

Horner, 2009) helps to ensure that multiple perspectives are taken into account during planning and implementation, increasing the potential for the school-wide plan to reflect values that are representative of the entire school and community. In short, many of the avenues through which SW-PBS practices may influence behavioral outcomes are the same or similar to strategies recommended as part of a CRCM approach to behavior management. Table 2 outlines the specific areas where CRCM strategies and features of SW-PBS overlap.

Table 2

Alignment between Culturally Responsive Classroom Management Strategies and Features of School-Wide Positive Behavior Support

Culturally Responsive Classroom Management Strategies		Features of School-Wide Positive Behavior Support *(*Sugai *& Horner, 2009)*
Ongoing analysis of academic and behavioral data coupled with responsive decision making and early intervention	↔	A data management system exists that is accurate, up-to-date, easy to access/use, and useful in guiding decisions about environmental and behavioral supports *(feature #6)*
Fostering positive and caring student-teacher relationships and school climates	↔	A system exists to encourage students and staff to follow the school-wide expectations *(feature #4)*
Designing positive, preventative and proactive discipline strategies	↔	A continuum of consequences for inappropriate behavior exists that is clearly defined, includes prevention strategies and options for more intensive instruction and support, is regularly monitored for effectiveness, and is based on the function of the student's behavior *(feature #5)*
Teaching clear expectations for behavior	↔	A few positively stated expectations are established to guide the behavior of all staff and students in all settings, which convey support for academic and behavioral outcomes and are culturally appropriate *(feature #2)*
Communicating high expectations for student achievement	↔	
Utilizing effective models of explicit instruction	↔	The school-wide expectations are explicitly and continually taught *(feature #3)*
Becoming familiar with students' backgrounds and values, and using this knowledge to foster positive relationships, improve curricula and instruction, and mediate interpretations of behavior	↔	An approach to discipline that is based on behavior analytic principles, reflects evidence-based practices, is culturally appropriate, and based on active instruction *(feature #1)*
Promoting parent involvement		*Not explicitly addressed*
Considering how the personal culture and biases of "power-holders" may impact those individuals' evaluations of behavior		*Not explicitly addressed*

In some cases, the consistency between CRCM strategies and the features of SW-PBS is evident (such as with ongoing data analysis for decision making and establishing a data system that guides decisions about supports). However, proponents of CRCM recommend additional strategies that are not explicitly outlined as a core feature of SW-PBS: parent involvement, and consideration of how the personal culture and biases of "power-holders" may impact those individuals' evaluations of behavior.

While not outlined as a core feature of SW-PBS, parent involvement is still considered a necessary part of implementation, as evidenced by its inclusion on measures used to evaluate the fidelity of SW-PBS (e.g., the *Benchmarks of Quality*; Cohen, Kincaid, & Childs, 2007). However, the CRCM strategy of considering how power-holders' personal culture and bias may impact their evaluations of behavior, is not reflected in the core features of SW-PBS, nor is it reflected in common SW-PBS evaluation measures. Utley, Kozleski, Smith, and Draper (2002) examine the role of multicultural education in SW-PBS, and while they note Sugai et al.'s (2000) assertion that "the use of culturally appropriate interventions also is emphasized in the PBS approach" (p. 134), they identify several key elements of multicultural education that nevertheless need to be directly addressed in order to design SW-PBS programs that are effective for minority students, and which respect the values and beliefs of differing cultures. Practitioners of PBS must go beyond Sugai et al.'s (2000) consideration of "the unique and individualized learning histories…of all individuals…who participate in the PBS process," and engage in a deliberate and thoughtful examination of the values and beliefs inherent in their school-wide plan, and how those values and beliefs might differ from those of the students and families for which their plan was designed. Sometimes, this includes a willingness on the part of educators to adjust what they consider to be "appropriate" behavior when selecting behaviors to reward and to punish. Based on the recommendations outlined in Utley et al.'s (2002) paper, it seems that PBS practitioners may need to move beyond basic SW-PBS implementation to ensure their approach is truly culturally responsive.

Other researchers seem to agree. In his 2007 keynote address for the PBS Implementer's Forum in Chicago, Illinois, Skiba presented information showing how even though

overall suspensions could be reduced for a school that implemented SW-PBS, levels of disproportionality could actually *increase* following implementation, in cases where referrals or suspensions were significantly reduced for one racial group but not another. However, the database and case example shared by Skiba did not include information about the degree to which those schools implemented SW-PBS with fidelity. It could be argued that the schools in that sample might have implemented very few PBS strategies, which produced benefits for only a portion of the student body – rather than the 80-90% that should benefit from a fully-implemented SW-PBS approach (Peshak-George et al., 2009). Other research (Cohen et al., 2007; Florida's Positive Behavior Support Project, 2009) has supported that schools which implement higher levels of SW-PBS experience stronger behavioral and academic outcomes. If a SW-PBS approach truly benefits 80-90% of students within a school's population – including minority students – then it may be that schools must reach a high level of implementation fidelity before experiencing equitable gains in behavioral outcomes across all of their students. The present study aimed to determine whether schools that implement SW-PBS with fidelity experience lower levels of disproportionality in ODRs and suspensions compared to schools that implement SW-PBS with less fidelity.

Summary

Disparities in academic and behavioral outcomes for minority students are decades-old problems, with roots dating back to post-civil war era schooling. Historical and modern-day practices of tracking, ESE placement, and exclusionary discipline negatively impact the amount of time minority students spend engaged in the general education curriculum, resulting in lower rates of achievement and poorer post-high school outcomes compared to their White peers. Exclusionary discipline in particular is troubling given the extent of the disproportionality (especially for African American students, who nationally receive suspensions at three times the rate of White students), the persistence of the findings (current levels of disproportionate suspension have been documented since the 1970's), and the resistance of educators to acknowledge a potential for their contribution to the problem (as evidenced through attempts to focus attention on socioeconomic status or cultural tendencies inherent only in the students).

Culturally responsive education is frequently hailed as a possible remedy for disproportionate outcomes, but a lack of empirical support for culturally responsive classroom management (CRCM) strategies has turned advocates' attention towards the systems-level approach of positive behavior support and its growing research base. While positive behavior support has proven to be effective in reducing overall levels of office discipline referrals and suspensions, and its success has been replicated in schools with high populations of minority students, the overlap with recommended CRCM strategies is not perfect, and effective outcomes across all groups of students is not necessarily guaranteed. The present study investigated whether schools that implement positive behavior support with high levels of fidelity are more likely to have lower levels of disproportionate office referrals and suspensions than schools which implement with less fidelity.

Chapter Three: Method

Introduction

The purpose of the present study was to investigate the rates of disproportionality in office discipline referrals (ODRs) and instances of out-of-school suspensions (OSS) for high-implementing and low-implementing PBS schools. Specifically, risk ratios for receiving an ODR and for receiving an OSS were calculated for African American and Hispanic students on a school-by-school basis, and then compared to each school's Benchmarks of Quality score. The ensuing analysis evaluated whether schools that implement SW-PBS with higher levels of fidelity tended to have lower levels of disproportionate referrals and suspensions for African American and Hispanic students.

Data Sources

In collaboration with the University of Oregon's College of Education/Educational and Community Supports PBIS Technical Assistance Center, school-level data from the *PBS Surveys* and the *School-Wide Information System* (SWIS) applications were utilized for this study. Both applications are implementation tools that assist schools in collecting and organizing data relevant for decision making in SW-PBS systems. Data entered into each system reflect "real-life" conditions such that school-level personnel enter the information they have available as they are able. Data from these databases reflect the reality of school-level implementation, rather than experimentally-controlled research conditions.

PBS Surveys is an internet application available to schools nationwide that implement SW-PBS. School personnel typically learn of PBS Surveys through consultation with technical assistance providers who are familiar with the application, through information shared at professional conferences, or through the website www.pbis.org. To establish an account, school personnel fill out a general application, which includes a data sharing agreement. Personnel from

participating schools may log in to their account to complete a variety of survey tools, including the Benchmarks of Quality, which are designed to assess components of their school-wide implementation (Educational and Community Supports, 2010). Personnel can then immediately view reports that summarize the survey information. All PBS Survey information is self-report information presumed to have been accurately entered by school personnel who are trained in SW-PBS; however, the accuracy of the information could not be established for this study. The *PBS Surveys* database was used to provide the Benchmarks of Quality score for all schools.

The *School-Wide Information System* (SWIS) is another internet application designed to "help school personnel to use office referral data to design school-wide and individual student interventions" (Educational and Community Supports, 2010b). The SWIS application is an internet application available to schools that implement SW-PBS. Schools that use the SWIS application decide with their administrator and/or district personnel whether they will utilize the application for a small annual fee. All schools using SWIS have been trained in SW-PBS, but information regarding the school's level of implementation is not collected in this database. Once schools are trained in SW-PBS, they can request paperwork from their PBS trainer and/or SWIS facilitator that, once completed, will establish a SWIS account for their school. Once two to three people from the school are trained by the SWIS facilitator in the specifics of how to use the application, those personnel may log in to their account. School personnel enter ODR data and view graphs and reports that summarize and/or disaggregate the ODR information to assist their SW-PBS teams in data-based decision making. All SWIS information is entered by school personnel based on information contained in each school's ODR form, and information in the SWIS database presumably reflects a running total of up-to-date, valid office discipline referrals; however, the accuracy and completeness of the information is not controlled. For the purpose of this study, the SWIS database was used to provide school demographic information (including enrollment data by race and ethnicity), ODR information, and OSS information.

Participants

Participants for this study consisted of a paired sample of 83 elementary schools from multiple states that utilized the both the PBS Surveys and SWIS applications during the 2008-2009 school year. More specifically, schools that participated in this study had: (1) completed the Benchmarks of Quality through PBS Surveys; (2) used SWIS to keep track of their ODR and suspension data; (3) reported their school's race and ethnicity data in SWIS; and (4) viewed their school's "Ethnicity Report" in SWIS at least once during the 2008-2009 school year. This last criterion was included to help establish some social validity for the sample – if schools looked at their Ethnicity Report, it seemed more likely that they attended to how implementation impacted students of different races than if they never looked at this report at all.

As part of the registration process, schools which sign up for an account with PBS Surveys and SWIS sign a data sharing agreement. This agreement permits the University of Oregon/PBIS Technical Assistance Center to use schools' data in an anonymous format for evaluation and research purposes, consistent with the proposed purpose and methods of this study. As the results of this study will be shared with the PBIS Technical Assistance Center with the intention of advancing the field of SW-PBS implementation, conflict arising from data sharing issues is not expected. Demographic information describing the overall population of schools utilizing PBS Surveys and SWIS, as well as the specific schools randomly selected for this study, are included in the final results. No individual student information, school name, or state will be collected for this study.

Variables

Level of implementation: Higher- and lower-implementing schools. The independent variable for this study was the school's level of implementation of SW-PBS, as summarized by their Benchmarks of Quality (BoQ) score. The Benchmarks of Quality (BoQ) is a team-based self-report tool designed to measure the level of fidelity with which a school implements SW-PBS (Cohen et al., 2007). The BoQ contains 53 items evaluating different activities involved in SW-PBS implementation, and has possible scores ranging from 0 to 100 points. School-Wide PBS team members assess the degree to which each item is in place at their school; different items

have varying degrees to which they may be implemented. A PBS Coach (who has additional expertise in SW-PBS) completes his/her own version of the BoQ independent of the SW-PBS team members, using a Scoring Guide to determine point value. Once all team members and their Coach have completed the tool, they meet to discuss any areas of disagreement and to identify implementation goals based on the combined results. Although the BoQ Scoring Form provides a summary sheet that helps guide teams and their Coach through the process of identifying disagreements, this information is not collected as part of the PBS Surveys database – only the final (agreed-upon) scores for each item are recorded. Therefore, the initial level of agreement between the team members and the PBS Coach cannot be assessed. The BoQ total score captured in the PBS Surveys database reflects the final score that is based on feedback and discussion from all SW-PBS team members, their Coach, and the Scoring Guide.

The BoQ is a valid tool for measuring SW-PBS implementation, with an internal consistency of .96, test-retest reliability of .94, and concurrent validity of .51 (Cohen et al., 2007). The BoQ is thought to be a more sensitive measure than the comparative research tool, the *School-Wide Evaluation Tool* [SET; Sugai, Lewis-Palmer, Todd, & Horner, 2005], and the BoQ's lower level of concurrent validity is thought to be reflective of its inclusion of implementation information not assessed by the SET (Cohen et al., 2007). The BoQ has been shown to differentiate schools with greater decreases in ODRs, such that schools scoring higher than 70 on the BoQ experienced greater decreases in office discipline referrals than schools that scored below 70 (Cohen et al., 2007). Informal evaluations have also supported this distinction; in Florida, schools scoring 70 or higher on the BoQ experienced greater decreases in ODRs, in-school suspensions, and out-of-school suspensions than schools that scored below 70 on the measure (Florida's Positive Behavior Support Project, 2009). In the current study, cut scores based on the overall distribution of BoQ scores where used to identify high- and low-implementing schools. The process for identifying these cut scores is described in detail in Chapter Four.

Realizing that the categorization of continuous variables such as the BoQ has the potential to obscure the exact nature of the relationships between variables, the analysis portion

of this study included a detailed descriptive analysis depicting each school's overall referral and suspension rates, referral and suspension rates by gender, and referral and suspension rates by race. This descriptive analysis was performed once using duplicated counts of referrals and suspensions (so that when a student received multiple instances of these consequences, all incidents were counted), and was performed a second time using unduplicated counts of referrals and suspensions (so that each student was counted only once, regardless of how many referrals and/or suspensions s/he received). This process is described in more detail in Chapter Four. To provide a descriptive analysis that included fidelity, each school's total BoQ score was plotted against the school's risk ratios for African American and Hispanic students for ODRs and OSS. It was hoped that this detailed descriptive analyses facilitated a more thorough understanding of the nature of disproportionality in schools that implement SW-PBS, and the potential relationship between reported implementation and these groups' risk for ODR and OSS.

Level of disproportionality: High-, moderate-, and low- disproportionality schools. The dependent variables for this study were the school's level of disproportionality for African American and Hispanic students, as summarized by their respective risk ratios for ODRs and OSS. While there are multiple statistics that have been used to summarize disproportionality in schools (composition, risk, and risk ratio are the most common), the risk ratio is the only measure that can be used on its own while still providing a complete picture of a group's level of risk (Bollmer, Bethel, Garrison-Mogren, & Brauen, 2007). Risk ratios are also the most commonly used measure in state definitions of disproportionality for special education and related services (Burdette, 2007). The risk ratio summarizes a group's risk for membership in a category compared to the risk of a comparison group (which in this study will be "all other students"), and is calculated as follows:

$$\text{Risk ratio} = \frac{\text{Risk for racial group}}{\text{Risk for comparison group}}$$

To better illustrate, consider a hypothetical school where African American students make up 30% of the school's population, and students from all other racial groups made up the

balance of the population (i.e., 70% of the students came from a combination of all other racial categories).If 75% of African American students in this school received an OSS, but only 25% of students from every other racial category received an OSS, the risk ratio would be 3.0 (.75/.25 = 3.0). In other words, in this hypothetical school, an African American student would be three times more likely to receive an OSS than would a student from any other racial category. In this study, risk ratios were calculated separately for ODRs and OSS for each racial category, resulting in four total risk ratios for each school (two racial categories [African American and Hispanic] for each of two risk categories [ODRs and OSS]).

Just as the BoQ scores for schools were divided into high- and low-implementing categories, schools' risk ratios also were categorized. While this grouping sacrificed a level of specificity in the results, this allowed for an easier conceptualization of the overall relationship between implementation and disproportionality, and provided flexibility in terms of obtaining a specific sample size for statistical analysis (which will be described in the *Procedures* section). School's risk ratios for ODRs and OSS were grouped into three categories:

High Disproportionality: Schools with risk ratios of 2.0 or higher for African American or Hispanic students were included in the "High Disproportionality" category;

Moderate Disproportionality: Schools with risk ratios of 1.20-1.99 for African American or Hispanic students were included in the "Moderate Disproportionality" category; and

Low Disproportionality: Schools with risk ratios below 1.20 for African American or Hispanic students were included in the "Low Disproportionality" category.

Schools which have no disproportionality (or under-representation) in ODRs or OSS for African American or Hispanic students were classified under the "Low Disproportionality" category. While this may not be an entirely accurate description, the research questions centered around *levels* of disproportionality and *levels* of SW-PBS implementation. The emphasis on the more general "level" of disproportionality allowed for a consistent interpretation of the results (in other words, schools with "low" levels of disproportionality tended to have "X" levels of SW-PBS implementation).

The proposed breakdown of disproportionality levels was not without foundation: at the request of the Office of Special Education Programs (OSEP), Project Forum at the National Association of State Directors of Special Education (NASDSE) published a brief policy analysis in 2007 that summarized state definitions of "Significant Disproportionality" in special education placements (Burdette, 2007). Each state is allowed to set its own limit on what constitutes "significant," and as a result the cut-off values for significant disproportionality varies from state to state. Sample definitions included one standard deviation more or less than a 3-year average baseline; 20% more or less than a district's total education population; a risk ratio of 3.0 or higher for three consecutive years; a risk ratio of 2.5 or higher in certain ESE categories; and a risk ratio of 2.0 or higher (no additional specifications were provided). Other guidelines offered by advocates and researchers have specified that risk ratios over 1.0 signal disproportionality, with indexes of 1.2 and higher being important enough for a school, district, or state to take action (Kozleski, 2005). The proposed breakdowns for this study reflected components of both the variation in state definitions and advocates' recommendations for action.

The specific information collected for each school was as follows:

1. Benchmarks of Quality total score
2. Total enrollment (number)
    a. Total enrollment reflects the total number of students enrolled at a school. All enrollment information is entered into the SWIS application's database by school personnel, and is assumed to be accurate to within 10% of the total population.
3. School-Level demographics by race/ethnicity (number)
    a. This included the total number of students identified as African American, American Indian, Asian, Hispanic, and White enrolled at the school. Students are identified as belonging to a particular racial and/or ethnic category by their parents upon enrollment in the school district.
    b. This also included other descriptive information, such as the school locale, Title 1 eligibility, Free/Reduced Lunch enrollment, classroom teacher full-time equivalents (FTE).

4. Referral-Level information, including
    a. School number
    b. Student number
    c. Referral number
    d. Student race/ethnicity
    e. Student gender, and
    f. Administrative decision

All data will reflected information from the 2008-2009 school year.

Procedure

The University of Oregon's PBIS Technical Assistance Center pulled the specified data on a school-by-school basis, removing any identifying information prior to sending the information to this author for analysis. Therefore, specific numbers of ODRs, OSS, Benchmarks of Quality scores, and risk ratios were not attributed to an identifiable school. Upon receiving the raw data, this author performed the descriptive analyses (described above), and calculated risk ratios for African American and Hispanic students for both ODRs and OSS. The data set then was categorized to identify high- and low-implementing schools, and then to identify high- and low-implementing schools with high- moderate- and low levels of disproportionality.

Analysis

For the descriptive analysis of fidelity and disproportionality, scatter plots were developed to depict the distribution of schools' BoQ scores in relation to their risk ratios for ODRs and OSS for each racial group (African American students and Hispanic students). A total of four scatter plots were developed (BoQ scores and African American ODR risk ratios; BoQ scores and African American OSS risk ratios; BoQ scores and Hispanic ODR risk ratios; and BoQ scores and Hispanic OSS risk ratios). These graphs provided a detailed visual representation of the relationship between the independent and dependent variables that enabled a descriptive analysis of the data. The descriptive analysis addressed the first research question, "What are the risk ratios for office discipline referrals and incidents of out-of-school suspensions for African American and Hispanic students in schools that implement SW-PBS?"

The second research question “Is there a relationship between the level of implementation of SW-PBS and levels of disproportionality in office discipline referrals and suspensions?” was addressed through a chi-square test of independence. A chi-square test is appropriate when the variables under consideration are independent categorical variables, as was the case in the current study.

Four chi-square tests were performed: one test to determine the relationship between level of implementation and level of disproportionality for African American students in ODRs, another to determine the relationship between level of implementation and level of disproportionality for Hispanic students in ODRs, a third to determine the relationship between the level of implementation and level of disproportionality for African American students in OSS, and a fourth to determine the relationship between the level of implementation and level of disproportionality for Hispanic students in OSS. Schools were classified as “High Disproportionality” schools if they had a risk ratio of 2.0 or higher for African American or Hispanic students. Similarly, if a school had a risk ratio for African American or Hispanic students that fell between 1.20 and 1.99, they were classified in the “Moderate Disproportionality” category. Finally, if a school had a risk ratio for African American or Hispanic students of 1.19 or lower, they were classified in the “Low Disproportionality” category. Each school was classified only once; if they had a high level of disproportionality for one racial group, and moderate or low disproportionality for another racial group, they were only classified as having “High Disproportionality.”

The results of the chi-square analysis determined whether there was a statistically significant relationship between level of SW-PBS implementation and level of disproportionality in African American or Hispanic ODRs and OSS.

Chapter Four: Results

This chapter provides a thorough and systematic investigation of disproportionality in office referrals and suspensions for African American and Hispanic students in elementary schools that implement SW-PBS. Beginning with an examination of the characteristics of the original sample, rates of referrals were examined that were based on duplicated and unduplicated counts as a way of establishing a background for the analysis of disproportionate referrals for different groups of students. Differences resulting from the two types of counts prompted a continuation of this dual analysis for referrals by gender and race/ethnicity (for African American students, Hispanic students, and White students), leading to a demonstration of the advantages of the risk ratio in calculating disproportionality across a large number of schools. The dual comparison and descriptive risk ratio analyses were repeated for suspensions, and culminated in an examination of the relationship between SW-PBS implementation and levels of disproportionality. The chapter concludes with a statistical test of significance for each of the groups.

Establishment of the Sample

School- and referral-level data were received for 131 schools. School-Level data included total student enrollment, enrollment by race/ethnicity, school locale (e.g., city, suburb, rural, etc.; National Center for Educational Statistics [NCES], no date), Title 1 eligibility, Free/Reduced Lunch enrollment, classroom teacher full-time equivalents (FTE, which represents the number of full-time positions allotted to instructional positions), and Benchmarks of Quality (BoQ) scores. Referral-Level data included de-identified student and referral ID numbers, student ethnicity, student gender, and administrative decision. A little more than one third (37%) of the initial data set contained information that could not be included in the study. Specifically, the initial set included 34 middle, high, and/or combination schools (e.g., grades K-8), seven schools that were

missing key information (such as grade levels served, or school-level race/ethnicity enrollment); and seven schools that were duplicate entries. All 48 of these schools were subsequently omitted from the analysis, resulting in a total sample size of 83 elementary (K-6) schools.

Regional representation. A total of eight states were represented in the final sample. Table 3 lists the number of districts, the number of schools, and the extent to which each locale was represented in each state. Two states contributed a substantially larger number of schools to the sample than the others (F and E, with 31 and 21 schools respectively). In fact, the number of schools from these two states together was greater than the number of schools from the other six states combined (52 schools from states F and E, versus 31 schools from all of the other states). The majority of the schools in this sample were located in cities (47%) or suburbs (30%), with only 15% in rural locations and 8% located in towns.

Table 3

Regional Representation of Elementary School Sample (n=83)

State	Number of Districts	Number of Schools	Schools by Region			
			Rural	Town	Suburb	City
A	1	1	--	--	--	1
B	1	15	--	--	1	14
C	2	2	1	--	1	--
D	2	8	1	2	3	2
E	10	21	5	5	3	8
F	4	31	3	--	15	13
G	2	2	--	--	2	
H	1	3	2	--	--	1

School characteristics. Table 4 lists descriptive characteristics for the 83 schools included in the sample. The distributions of most of the school characteristics were approximately normal, with the exception of the minority rate, which at more than twice the standard error of skewness was significantly positively skewed (skewness = .69; *SE* of skewness = .264). The majority of the schools in the sample (78%, or 65 schools) had minority rates of 50% or less, with half of the sample containing schools with even lower minority rates (28% or less). With

exceptions in only 22% of the sample, it appears that most commonly, students who were identified as white made up the majority of each school's population. In addition, every school reported having students who participated in the Free/Reduced Lunch program (FRL), ranging from 7% to 99% of their populations. Every school reported a different FRL rate, so a mode could not be calculated. The distributions of the sample's descriptive characteristics are presented in more detail in Appendix B.

Table 4

School-Level Characteristics (n=83)

	Min.	Max.	Median	Mode	*M*	*SD*	Skew	Kurtosis
Total Population	88	769	406	451	424	133	.17	-.15
Minority Rate	.04	.92	.28	.25	.33	.22	.69	-.43
Free/Reduced Lunch Rate	.07	.99	.51		.52	.23	.16	-.68
Student to Teacher Ratio	8:1	24:1	17:1	15:1	17:1	3.5	-.24	-.52

Note. Mode could not be calculated for Free/Reduced Lunch Rate

School-Level Referral Distributions

Overall referral rates (duplicated count). To help provide a context for understanding schools' risk ratios, the total referrals for each school included in the final sample were examined along several dimensions. First, the total number of referrals for each school was divided by their total enrollment, producing an overall referral rate for each school. The referral rate is preferred over the total number of referrals as a way of understanding differences between schools because it takes the size of the school's enrollment into account. For example, one may expect a school with 1,000 students to write more referrals than a school with 200 students, even under ideal circumstances. Therefore, differences between schools in the total number of referrals are limited in their usefulness, unless differences in school enrollment are also taken into account.

The overall referral rate was initially based on a duplicated count of referrals, meaning that every referral written in the school was counted, even when multiple referrals were written for

the same student. This method of counting referrals may result in a relatively small number of students (those who receive multiple referrals) having a large impact on a school's overall rate; for example, a summary of national SWIS data showed that students who received six or more referrals made up only three percent of elementary schools' populations, but they accounted for 33% of schools' total referrals (Sugai & Horner, 2007). While basing a referral rate on a duplicated count of referrals may seem to some to be a somewhat unfair representation of a school, the duplicated count provides an indication of how frequently referrals were written, and may provide some insight into a building's general capacity for addressing problem behavior. The distribution of referral rates based on a duplicated count of referrals is presented in Figure 3.

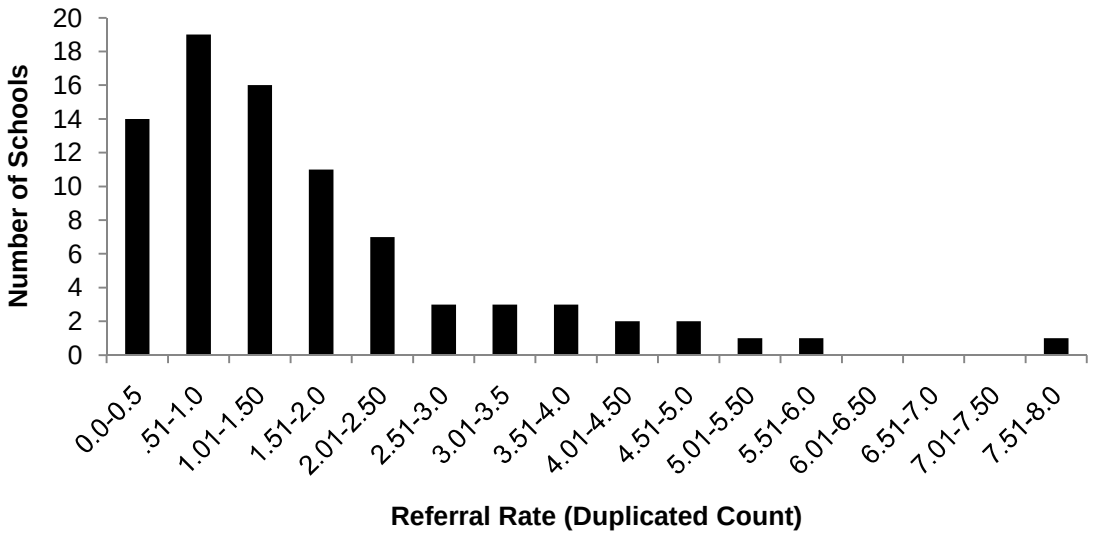

Figure 3. Distribution of schools' overall referral rates.

The distribution of duplicated referral rates ranged from .09 to 7.76 referrals per student (M = 1.65, SD = 1.45). Loosely interpreted, the school with the highest referral rate generated enough office discipline referrals so that every student in the school would receive at least seven referrals over the course of the school year. While this is not likely to be the actual scenario, it serves as a demonstration of the volume of referrals handled by the school's personnel. Overall, the volume of referrals for schools in this sample was somewhat high: 60% of the schools in the

sample had duplicated referral rates of 1.0 or higher (or at least one office discipline referral per student), while the national average for schools (K-12) using SWIS for 2007-2008 was just under one referral per student (.90; national average calculated by the author based on summary information available through the SWIS website: http://www.swis.org/index.php?page=resources;rid=10117).

The distribution of referral rates was significantly positively skewed (skewness = 1.71; *SE* of skewness = .264), with a significant tendency for rates to fall close to the mean (kurtosis = 3.44; *SE* of kurtosis = .523). It should be noted that one school had a referral rate (7.76) that placed them more than four standard deviations away from the mean, and two additional schools had rates that were more than two standard deviations from the mean. It is likely that these schools' referral rates impacted the overall mean for the sample, causing it to be somewhat larger than might be expected with a random sample of schools from a normal distribution.

Overall referral rates (unduplicated count). The duplicated referral rate may lead some observers to believe that the school personnel wrote referrals on a frequent basis, when in practice referrals were written infrequently for most students, but much more often for a small handful of "frequent flyers," or students who received multiple referrals. In order to address this influence within the current sample, referral rates were calculated a second time for each school based on an *unduplicated* count of referrals. The unduplicated count reflects the number of students who received at least one referral, and does not count a student more than once if they received multiple referrals. In other words, if a single student received 90 ODRs, that student would only contribute one referral to the overall unduplicated rate. The distribution of the unduplicated referral rates is presented in Figure 4.

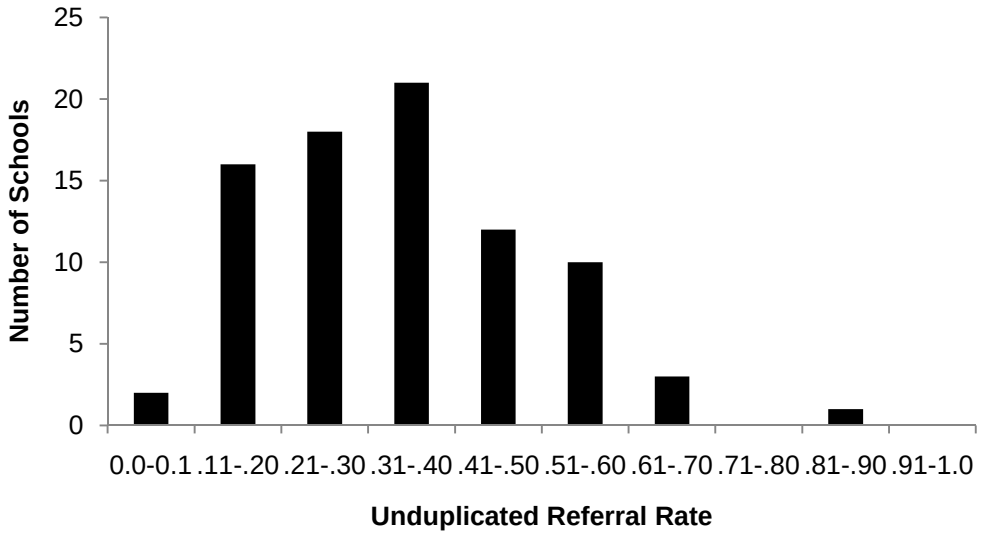

Figure 4. Distribution of schools' unduplicated referral rates.

Compared to the distribution of referral rates based on duplicated counts, the distribution of referral rates based on unduplicated counts was closer to resembling a normal distribution. The range of referral rates for the unduplicated counts reflected a minimum of .05 and a maximum of .83 (*M* = .34, *SD* = .15). Loosely translated, the school with the lowest unduplicated referral rate issued a referral for one out of every 20 students, schools with an average unduplicated referral rate issued approximately one referral for every three students, and the school with the highest unduplicated referral rate issued a referral for almost every student. In spite of the inclusion of a school whose rate was more than three standard deviations from the mean, the skewness and kurtosis of the overall distribution were both within an acceptable range (skewness = .50, kurtosis = .27).

Differences in the duplicated and unduplicated counts (overall referral rate). Changing the basis for the referral rates from duplicated to unduplicated counts resulted in greatly reduced referral rates for the schools, a smaller range of rates for the sample, and reduced variation between schools. Table 5 outlines these differences. Not surprisingly, the minimum, maximum, and average referral rates were smaller when based on the unduplicated count of

referrals, and all rates fell under one referral per student – which, since each student who received a referral is counted only once, should be expected. The small standard deviation (*SD* = .15, or one referral for every 6-7 students), indicated that in general, when the influence of students who frequently receive office referrals is removed from the picture, schools in this sample were rather consistent in the degree to which they issued referrals. Figure 5 further illustrates the impact of duplicated and unduplicated counts on schools' overall referral rates.

Table 5

Characteristics of Referral Rates (referrals per student) Based on Duplicated and Unduplicated Counts of Referrals

	Duplicated Count	Unduplicated Count
Minimum	.09	.05
Maximum	7.76	.83
Mean	1.65	.34
Standard Deviation	1.45	.15

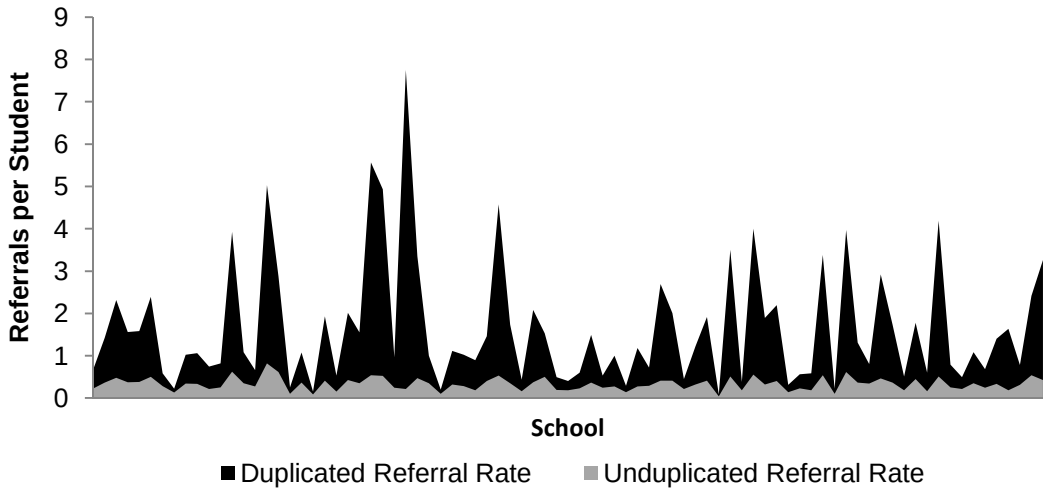

Figure 5. Differences between duplicated referral rates and unduplicated referral rates.

Referrals by gender (duplicated count). The percentage of each school's total referrals accounted for by both genders is reflected in Figure 6. Every school in the current study generated more office discipline referrals for boys than they did for girls. The percentage of referrals accounted for by boys ranged from 66% to 94% (*M* = 80.27, *SD* = 6.39), while the percentage of referrals accounted for by girls ranged from just under 6% to 33% (*M* = 19.73, *SD* = 6.39). Stated another way, at minimum boys received twice as many referrals as girls, and on average received four times more referrals than girls. In the most extreme case, boys received more than 16 times as many referrals as girls.

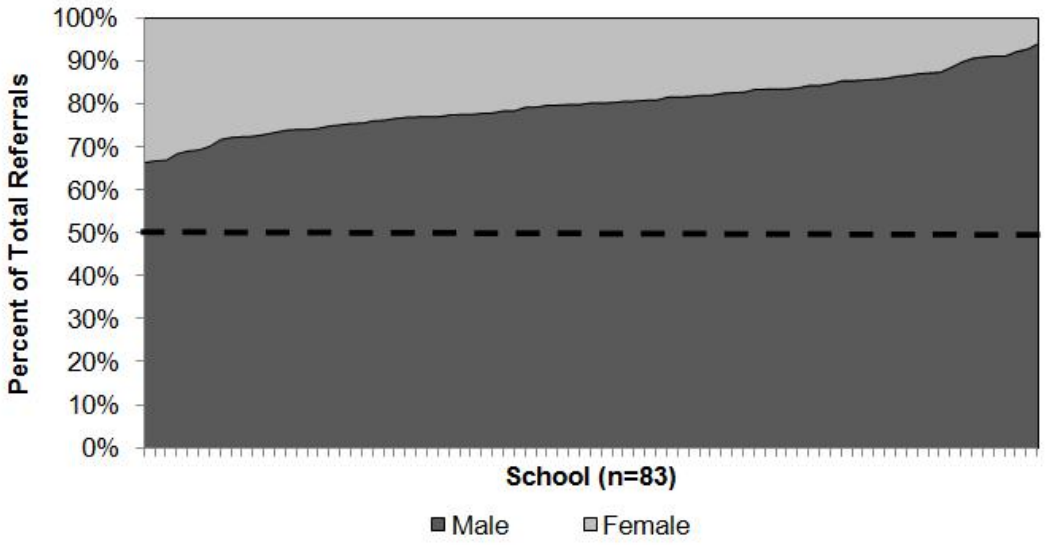

Figure 6. Percentage of referrals accounted for by gender (duplicated count).

Students receiving referrals by gender (unduplicated count). The percentage of students who received an office referral who were male and female is reflected in Figure 7. Although the actual school enrollment by gender was not available for this study, one might assume that the current sample mimicked the general population such that each gender made up approximately 50% of the student body. In the absence of data to indicate otherwise, one would therefore expect each gender to contribute an approximately equal proportion of students (50%) to the total number of students who received referrals. This was not the case for schools in the

current study: boys were represented among students who received referrals to a greater extent than girls in every instance. Boys' composition in the group of students who received referrals ranged from 56.6% to 90% (*M* = 72.4, *SD* = 6.5), while girls' composition ranged from 10% to 43.4% (the inverse of male students; *M* = 27.6, *SD* = 6.5). Stated another way, the percent of students who received an office referral who were male was at a minimum about 13% higher than expected; on average 45% higher than expected, and in the worst case, 80% higher than expected (or almost double, and accounted for almost all of the referrals in the school). Based on these findings, regardless of whether one examines the actual number of referrals by gender, or the proportions of each gender represented among students who receive referrals, boys consistently experience more referrals than girls – at frequently alarming rates.

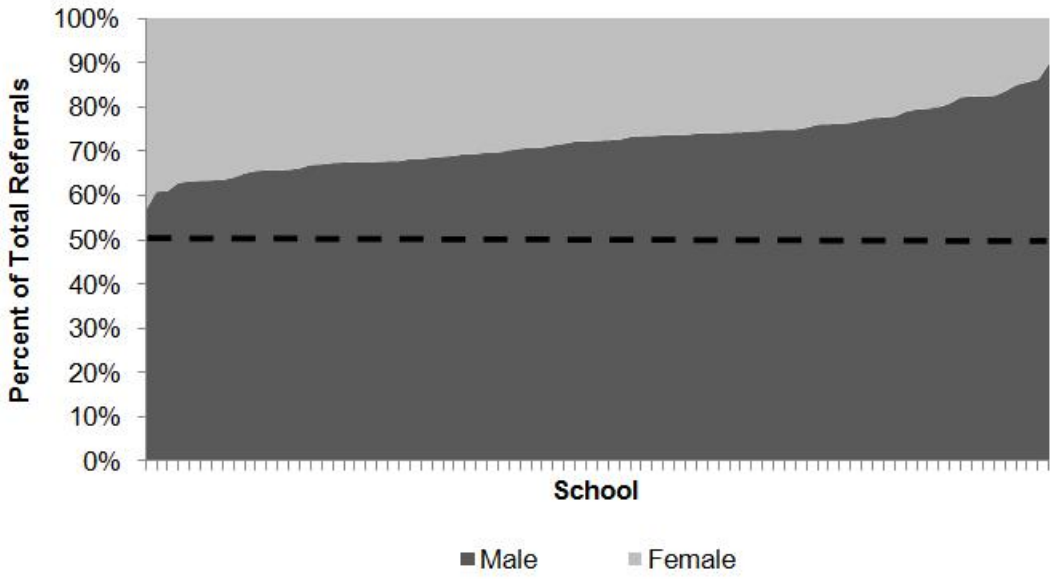

Figure 7. Percent of students who receive referrals by gender (unduplicated count).

Differences between the duplicated and unduplicated referral counts (gender).

Table 6 summarizes the differences in the gender analysis findings based on duplicated and unduplicated counts of referrals. In all but four schools, boys' duplicated count was higher than their unduplicated count, suggesting that most of the time boys were more likely than girls to receive multiple referrals. In four schools, the duplicated count *lowered* the percentage of

referrals accounted for by boys, which then in turn increased the percentage of referrals accounted for by girls. Although this analysis doesn't indicate whether these exceptions are due to a general pattern of more girls receiving multiple referrals than boys, or whether the exceptions are due to a single female student with many more referrals than other students, the rarity of this exception seems noteworthy.

Table 6

Characteristics of Referrals by Gender Based on Duplicated and Unduplicated Counts of Referrals

	Duplicated Count		Unduplicated Count	
	Boys	Girls	Boys	Girls
Minimum	66.7	5.8	56.6	10.0
Maximum	94.2	33.3	90.0	43.4
Mean	80.3	19.7	72.4	27.6
Standard Deviation	6.4	6.4	6.5	6.5

Impact of referral counts on school-level interpretations (gender). As might be expected, the change from duplicated to unduplicated count impacted schools to varying degrees. For example, the school that appeared to have the most proportionate distribution of referrals under the duplicated count (school #182) was not the same school that had the most proportionate composition of students who received referrals in the unduplicated count (school #345). This finding suggests that the unit of analysis used to determine gender disproportionality may be important, and that "frequent flyer" students may influence interpretations of gender disproportionality within a school for the better or for worse. A list of schools rank-ordered according to the percent of referrals accounted for by boys can be found in Appendix C.

Referrals by Race/Ethnicity

Missing data. Of the 83 schools in the sample, only six schools kept track of race/ethnicity information on all of their office discipline referrals. In fact, it was common for schools in this sample to overlook this information as they entered referrals into the SWIS database: 21% of the sample (17 schools) did not keep track of race/ethnicity on any of their

referrals, and an *additional* 31% of the sample (26 schools) was missing the information on more than 10% of their referrals. Figure 8 provides a visual representation of schools that were missing race/ethnicity data. In addition to the lack of consistency in recording race/ethnicity information on referrals, one school reported impossible numbers for their school-level race/ethnicity enrollment, where they listed a total of six African American students in their student body, but recorded referrals for 22 different African American students. Overall, over half of the schools were missing more than 10% of their referral-level race/ethnicity data, and as a result were removed from the sample. The remainder of the analysis focused on the 39 schools with race/ethnicity information on at least 90% of their referrals. It should be noted that although these schools had more complete records for race/ethnicity, it could not be determined how frequently school teams analyzed or applied the information.

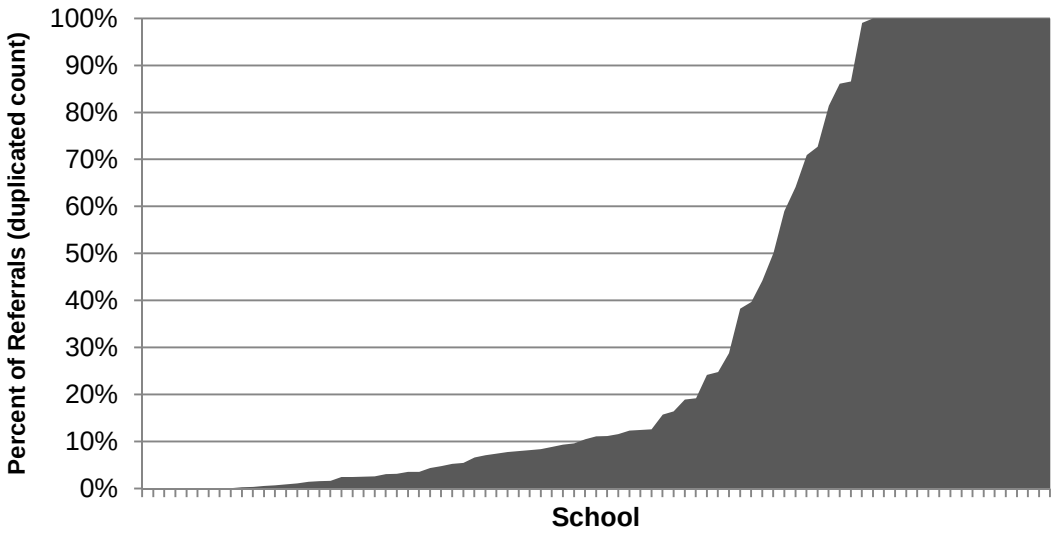

Figure 8. Percent of referrals missing race/ethnicity data at each school (n = 83).

Referrals by race/ethnicity: Students identified as African American (duplicated count). While there are several ways to measure disproportionality (Bollmer et al., 2007), comparing a group's percent that they are represented in a target category to the percent that they are represented within a larger population seems to be a method that is readily understood

by many people. For example, many people would recognize that if African American students were represented in 75% of a school's referrals, but only made up only 10% of the school's population, something may be out of order. In the absence of evidence to indicate otherwise, groups' representation in a target category should be similar to their representation in the population, and significant deviations from this pattern would indicate over- or under-representation. To supplement the descriptive analysis in the current study, the percent of referrals and the percent of students who received referrals was calculated for African American, Hispanic, and White students, and then compared to each group's percentage of their school's population. Other racial groups were not included in this portion of the analysis because the main research questions focus on referral patterns for African American and Hispanic students, and because other racial groups had much smaller numbers across the sample. Referral patterns for White students were included as a point of reference to a majority race. It should be noted that the total number of schools displayed in this part of the analysis was only 37, as two schools did not report any referrals for African American students. The comparison of the percent of referrals accounted for by African American students (the duplicated count) to their percent of the school's population is presented in Figure 9.

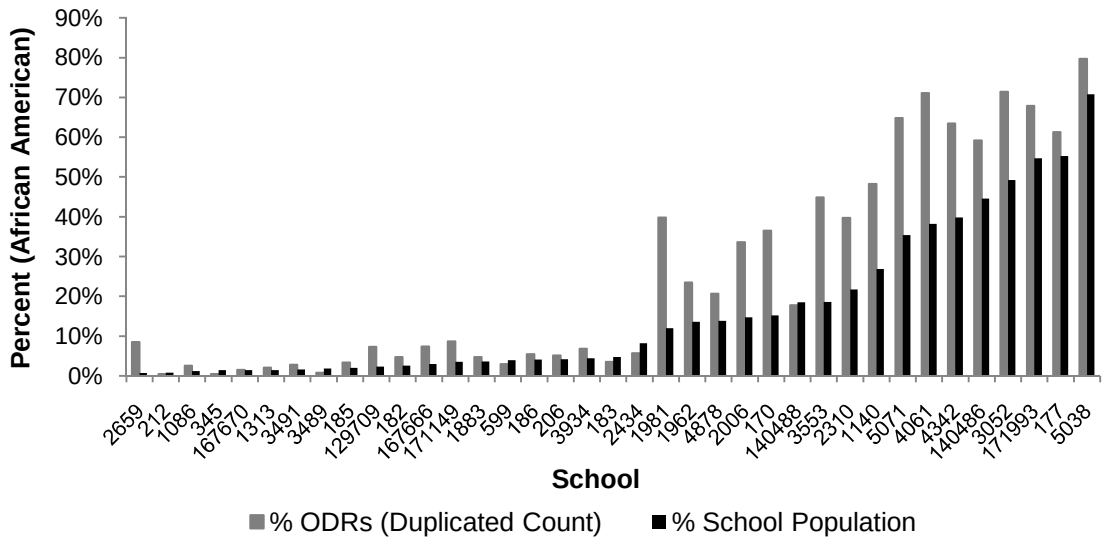

Figure 9. Percent of referrals accounted for by African American students compared to their proportion of the school's population.

Under the duplicated count, African American students were over-represented in 30 of the 37 schools (81% of this sample). The percentage of referrals accounted for by African American students ranged from 0.5% to 79.7%, while African American students' percent of their school population ranged from 0.7% to 70.8%. The school-by-school differences between percent of referrals and percent of population ranged from -64.3% to 1,114%, meaning that at the lowest end of the distribution African American students accounted for about 60% fewer referrals than would be expected, while at the highest end of the distribution African American students accounted for more than 11 times as many referrals as would be expected.

Students receiving referrals by race/ethnicity: Students identified as African American (unduplicated count). Figure 10 compares the percent of students who received referrals who were identified as African American (unduplicated count) to their percent of the school's population. With the unduplicated count, African American students were over-represented in 31 of the 37 schools (84% of this sample). The percent of students who received referrals who were identified as African American ranged from 0.5% to 81.0%, while (as stated before) African American students' percent of their school population ranged from 0.7% to 70.8%. The school-by-school differences between percent of students who received referrals and percent of population ranged from -72.2% to 278.8%, meaning that at the lowest end of the distribution the percent of students who received referrals who were African American was about 70% less than expected, while at the highest end of the distribution the percent of students who received referrals who were African American was almost three times higher than expected. Taken together with the findings from the duplicated count, it seems that while there are exceptions, African American students commonly experience more referrals than would be expected given their proportion of their school's population, regardless of the type of count used in the calculation.

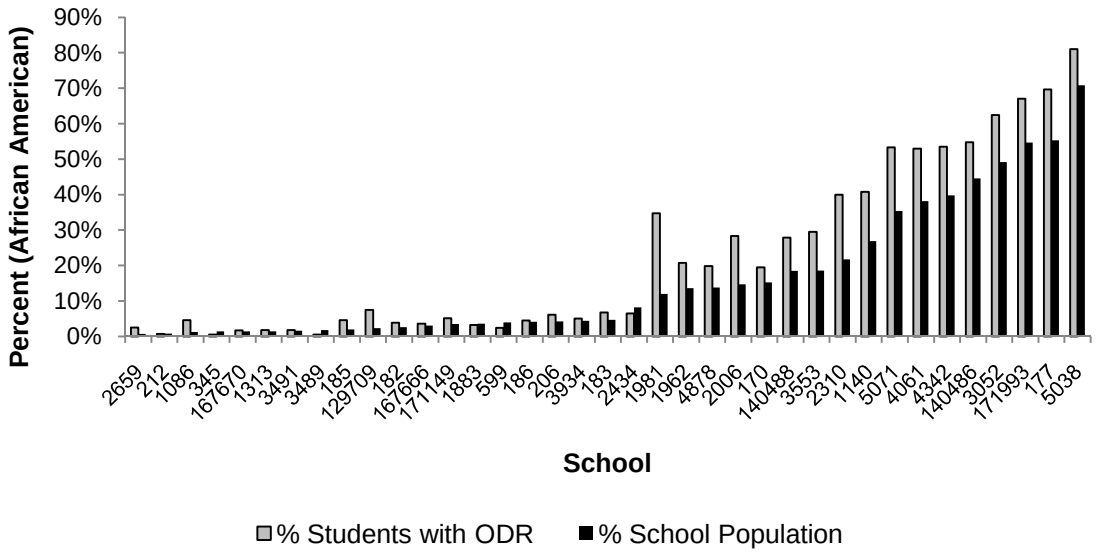

Figure 10. Percent of students who received an office referral who were identified as African American compared to their proportion of the school's population.

Magnitude of over-representation of African American students. In schools where there was over-representation, the average difference between the percent of referrals accounted for by African American students and their percent of the student population was 114.9% (*SD* = 197.2), meaning that on average, in the 30 schools with over-representation in their duplicated count, African American students accounted for about double the amount of referrals as would be expected given their proportion of the student population. The range of over-representation in the duplicated count was large, beginning at just 7% more referrals than would be expected, and maxing out at more than 11 times the number of referrals than would be expected. The range of over-representation with the 31 schools in the unduplicated count was smaller, beginning at 8.4% more African American students receiving referrals as would be expected and maxing out at almost three times as many African American students receiving referrals as would be expected. In schools with over-representation, the average percent of students who received referrals who were African American was 66.6% higher than their proportion in the student population (*SD* = 73).

Magnitude of under-representation of African American students. There were a few instances in which the duplicated and unduplicated counts resulted in *under*-representation for African American students. The differences between the percent of referrals and percent of African American students' population in these schools ranged from -3.8% to -64.3% for the seven schools under the duplicated count (*M* = -34.3%, *SD* =20), or from 3% less than expected to just over 60% less than expected. With the six schools under the unduplicated count, differences ranged from -9.4% to -72.2% (*M* = -35.7%, *SD* = 26.3), or from about 10% less than expected to about 70% less than expected. For each school with under-representation, African American students made up less than 20% of the school's population, and in seven of the eight instances the African American student population was even smaller (less than 10% of the school's total population). At the same time, however, a small African American student enrollment did not automatically lead to under-representation in referrals – in fact, the opposite pattern was discovered. The schools with the highest magnitudes of over-representation were all schools where African American students made up less than 20% of the school population, regardless of whether the duplicated or unduplicated count was used.

Overall, African American students were much more likely to experience over-representation than under-representation, and the degree of disproportionality they experienced in schools where they were over-represented (i.e., the percent difference between what was expected and what was observed) was much more severe than the degree of disproportionality they experienced in schools where they were under-represented.

Differences between the duplicated and unduplicated counts (African American students). Table 7 lists the school-by-school breakdowns of African American students' composition within a school building, their percent of referrals, and their composition within the body of students at that school who received referrals. As was found in the other analyses, the duplicated count tended to produce larger values than the unduplicated count – in most schools, the percent of referrals accounted for by African American students was higher than their composition in the group of students who received referrals. However, there were several instances where the opposite was true: in 12 of the 37 schools (32% of the sample), the

duplicated count for African American students was lower than their unduplicated count. This may be taken to mean that in those 12 schools, there were relatively fewer African American students who received multiple referrals.

Table 7

Differences in Composition of School Enrollment, Referrals, and Students who Received Referrals for African American Students

School ID	Percent School Population	Percent of Referrals, Duplicated Count	Difference from Population *(Duplicated)*	Percent of Students Receiving Referrals, Unduplicated Count	Difference from Population *(Undupl.)*
2659	0.7%	8.5%	1114.3%	2.5%	261.7%
212	0.8%	0.5%	-37.5%	0.7%	-11.3%
1086	1.2%	2.6%	116.7%	4.5%	278.8%
345	1.4%	0.5%	-64.3%	0.5%	-60.8%
167670	1.4%	2.1%	7.1%	1.8%	19.0%
1313	1.4%	1.5%	50.0%	1.7%	29.9%
3491	1.6%	2.8%	75.0%	1.8%	9.6%
3489	1.8%	0.8%	-55.6%	0.5%	-72.2%
185	2.0%	3.4%	70.0%	4.6%	130.8%
129709	2.3%	7.3%	217.4%	7.5%	224.5%
182	2.6%	4.7%	80.8%	3.8%	46.8%
167666	3.0%	7.4%	146.7%	3.6%	19.0%
171149	3.5%	8.7%	148.6%	5.1%	46.5%
1883	3.6%	4.7%	30.6%	3.3%	-9.4%
599	3.9%	3.0%	-23.1%	2.4%	-38.9%
186	4.1%	5.5%	34.1%	4.4%	8.4%
206	4.2%	5.1%	21.4%	6.1%	44.9%
3934	4.4%	6.8%	54.5%	5.1%	15.1%
183	4.7%	3.5%	-25.5%	6.8%	44.0%
2434	8.2%	5.7%	-30.5%	6.5%	-21.3%
1981	12.0%	39.8%	231.7%	34.7%	189.1%
1962	13.6%	23.5%	72.8%	20.7%	52.2%
4878	13.8%	20.7%	50.0%	19.8%	43.5%
2006	14.7%	33.6%	128.6%	28.3%	92.4%
170	15.2%	36.5%	140.1%	19.5%	28.4%
140488	18.5%	17.8%	-3.8%	27.9%	50.8%
3553	18.6%	44.9%	141.4%	29.5%	58.5%
2310	21.7%	39.7%	82.9%	40.0%	84.3%
1140	26.9%	48.3%	79.6%	40.8%	51.6%
5071	35.4%	64.8%	83.1%	53.3%	50.7%
4061	38.2%	71.1%	86.1%	52.9%	38.6%
4342	39.8%	63.5%	59.5%	53.5%	34.3%
140486	44.6%	59.2%	32.7%	54.8%	22.8%
3052	49.2%	71.4%	45.1%	62.4%	26.8%
171993	54.7%	67.9%	24.1%	67.1%	22.6%
177	55.3%	61.3%	10.8%	69.6%	25.9%
5038	70.8%	79.7%	12.6%	81.0%	14.4%

Impact of referral counts on school-level interpretations (African American students). The change from duplicated to unduplicated count resulted in a different interpretation of disproportionality in only three cases. In two of those cases, the duplicated count resulted in a finding of under-representation while the unduplicated count resulted in a finding of over-representation. So although the finding of over-representation in office discipline referrals for African American students is overwhelmingly consistent, the type of count used may occasionally influence the finding.

Referrals by race/ethnicity: Students identified as Hispanic (duplicated count). The comparison of the percent of referrals accounted for by Hispanic students (the duplicated count) to their percent of the school's population is presented in Figure 11**.** It should be noted that the 37 schools displayed in this part of the analysis were not all of the same 37 schools displayed in the comparison analysis for African American students. In the earlier analysis, two schools reported referrals for African American students but no referrals for Hispanic students, and in this analysis two schools reported referrals for Hispanic students but no referrals for African American students. Between the two comparison analyses, 35 of the 37 schools reflected information from the same schools.

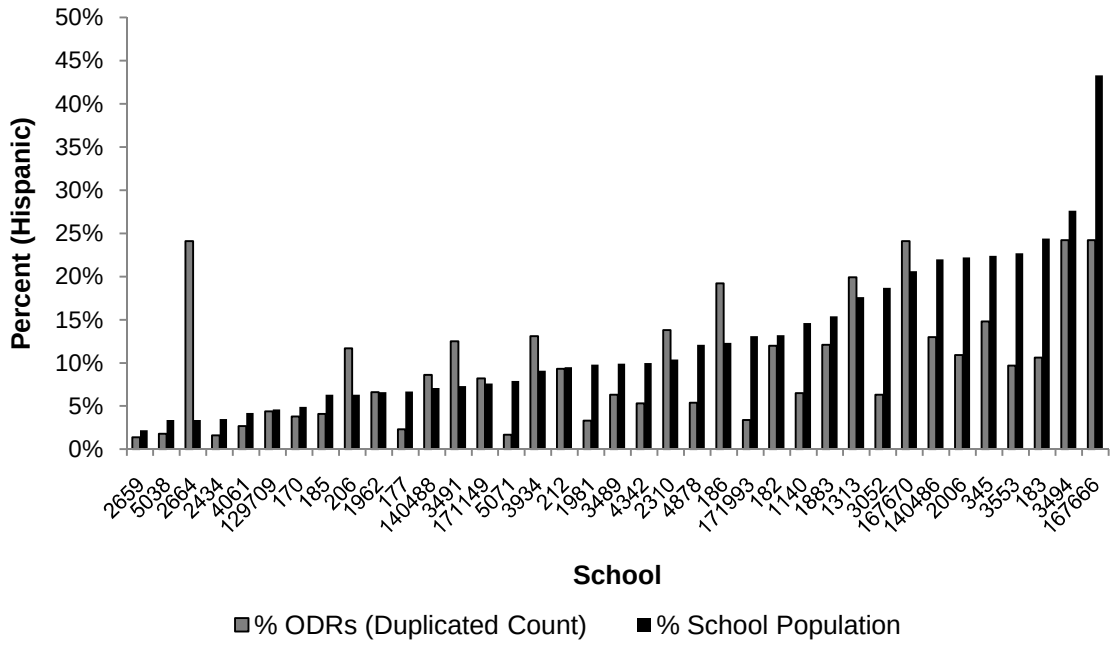

Figure 11. Percent of referrals accounted for by Hispanic students compared to their proportion of the school's population.

Under the duplicated count, Hispanic students were over-represented in just 10 of the 37 schools (27% of this sample). The percentage of referrals accounted for by Hispanic students ranged from 1.4% to 24.2%, while Hispanic students' percent of their school population ranged from 2.2% to 43.3%. The school-by-school differences between percent of referrals and percent of population ranged from -78.5% to 608.8%, meaning that at the lowest end of the distribution Hispanic students accounted for close to 80% fewer referrals than would be expected, while at the highest end of the distribution Hispanic students accounted for more than six times as many referrals as would be expected.

Students receiving referrals by race/ethnicity: Students identified as Hispanic (unduplicated count). Figure 12 compares the percent of students who received referrals who were identified as Hispanic (unduplicated count) to their percent of the school's population. With the unduplicated count, Hispanic students were over-represented in 16 of the 37 schools (43% of this sample). The percent of students who received referrals who were identified as Hispanic ranged from 2.2% to 33.9%, while Hispanic students' percent of their school population ranged

from 2.2% to 43.3%. The school-by-school differences between the percent of students who received referrals who were identified as Hispanic and the percent of their school population ranged from -73.1% to 73%, meaning that at the lowest end of the distribution the percent of students who received referrals who were Hispanic was about 70% less than expected, while at the highest end of the distribution the percent of students who received referrals who were Hispanic was 70% higher than expected. Overall, with the presence of over-representation for Hispanic students identified in less than 50% of the schools in this sample, disproportionality in office referrals appears to be less common for students from this ethnic group. Additionally, the likelihood of finding over-representation in office referrals for Hispanic students may depend on how the comparison is calculated, with the unduplicated count being more likely to result in over-representation than the duplicated count.

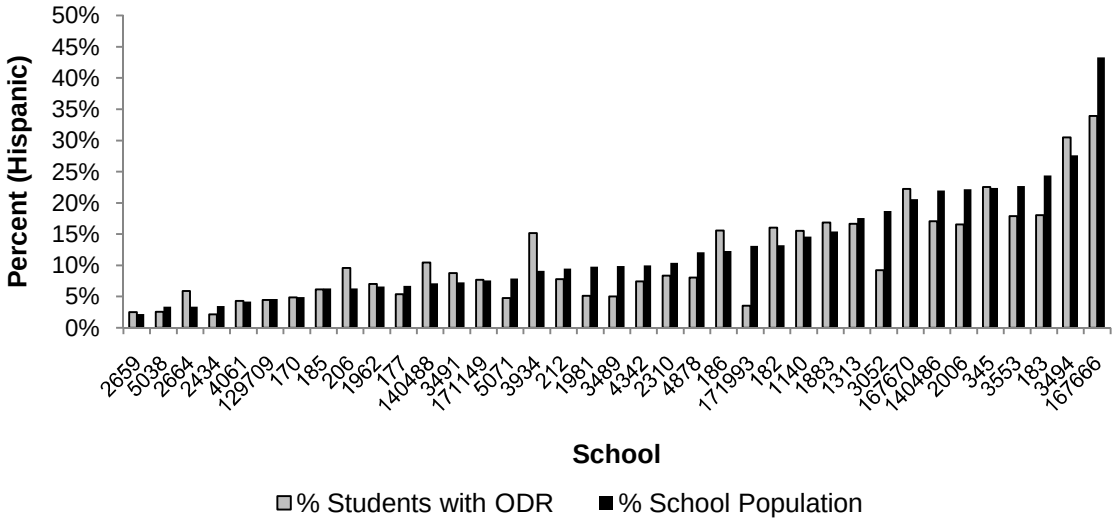

Figure 12. Percent of students who received an office referral who were identified as Hispanic compared to their proportion of the school's population.

Magnitude of over-representation of Hispanic students. In the ten schools where Hispanic students received more referrals than expected (using the duplicated count), the average difference between the percent of referrals accounted for and the percent of the student population was 95.8% (*SD* = 182.1), meaning that on average, when there was over-

representation for Hispanic students in referrals, they accounted for almost double the amount of referrals as would be expected given their proportion of the student population. The range of over-representation in the duplicated count spanned from just about 8% more referrals than would be expected, up to 6.1 times the number of referrals that would be expected. It should be noted, however, that with the exception of one school, differences ranged from 7.9% to 85.7%. In other words, in all but one case, the amount of referrals accounted for by Hispanic students was no more than 85% higher than would be expected.

The range of over-representation with the unduplicated count was smaller, beginning at less than 0.6% and maxing out at 73%, so that in the worst instance there were just over 70% more Hispanic students receiving referrals as would be expected. On average, in the 16 schools with over-representation in their unduplicated count, the percent of students who received referrals who were Hispanic was 23% higher than their proportion in the school population (*SD* = 23.7).

Magnitude of under-representation of Hispanic students. In all, there were 26 schools (70% of the sample) with under-representation in their duplicated count for Hispanic students, with differences between the percent of referrals accounted for by Hispanic students and their percent of population that ranged from -2.1% to -78.5%. On average, schools with under-representation for Hispanic students had 42.6% fewer referrals than expected given their proportion of the student population (*SD* =21.1). In the 21 schools (57% of the sample) with under-representation in their unduplicated count, the differences in the percent of students who received referrals who were Hispanic and their percent population ranged from -0.4% to -73.1% (*M* = -27.1%, *SD* = 18.2). Overall, under-representation was much more common for Hispanic students than was over-representation, and with one exception the degree of disproportionality (the percent difference between what was expected and what was observed) was about equal, regardless of direction (over-representation or under-representation).

Differences between the duplicated and unduplicated counts (Hispanic students). Table 8 lists the school-by-school percentages for Hispanic students' composition within a school building, their percent of referrals, and their composition within the body of students at that school

who received referrals. In stark contrast to earlier analyses, the duplicated count of referrals frequently produced *smaller* values than the unduplicated count: in 27 of the 37 schools (73% of this sample), Hispanic students' composition within the body of students who received referrals was higher than the proportion of referrals attributed to them. This suggests that it was relatively less common for Hispanic students to receive multiple referrals.

Table 8

Differences in Composition of School Enrollment, Referrals, and Students who Received Referrals for Hispanic Students

School ID	Percent School Population	Percent of Referrals, Duplicated Count	Difference from Population *(Duplicated)*	Percent of Students Receiving Referrals, Unduplicated Count	Difference from Population *(Undupl.)*
2659	2.2%	1.4%	-36.4%	2.5%	15.1%
5038	3.4%	1.8%	-47.1%	2.6%	-24.6%
2664	3.4%	24.1%	608.8%	5.9%	73.0%
2434	3.5%	1.6%	-54.3%	2.2%	-38.6%
4061	4.2%	2.7%	-35.7%	4.3%	2.7%
129709	4.6%	4.4%	-4.3%	4.5%	-2.7%
170	4.9%	3.8%	-22.4%	4.9%	-0.4%
185	6.3%	4.1%	-34.9%	6.2%	-2.3%
206	6.3%	11.7%	85.7%	9.6%	51.8%
1962	6.6%	6.6%	0.0%	7.0%	6.5%
177	6.7%	2.3%	-65.7%	5.4%	-20.0%
140488	7.1%	8.6%	21.1%	10.5%	47.4%
3491	7.3%	12.5%	71.2%	8.8%	20.2%
171149	7.6%	8.2%	7.9%	7.7%	1.2%
5071	7.9%	1.7%	-78.5%	4.8%	-39.7%
3934	9.1%	13.1%	44.0%	15.2%	66.9%
212	9.5%	9.3%	-2.1%	7.8%	-17.9%
1981	9.8%	3.3%	-66.3%	5.1%	-47.9%
3489	9.9%	6.3%	-36.4%	5.0%	-49.5%
4342	10.0%	5.3%	-47.0%	7.4%	-25.7%
2310	10.4%	13.8%	32.7%	8.3%	-19.9%
4878	12.1%	5.4%	-55.4%	8.1%	-33.4%
186	12.3%	19.2%	56.1%	15.6%	26.5%
171993	13.1%	3.4%	-74.0%	3.5%	-73.1%
182	13.2%	12.0%	-9.1%	16.0%	21.4%
1140	14.6%	6.5%	-55.5%	15.5%	6.4%
1883	15.4%	12.1%	-21.4%	16.8%	9.4%
1313	17.6%	19.9%	13.1%	16.7%	-5.3%
3052	18.7%	6.3%	-66.3%	9.2%	-50.8%
167670	20.6%	24.1%	17.0%	22.2%	7.9%
140486	22.0%	13.0%	-40.9%	17.1%	-22.4%
2006	22.2%	10.9%	-50.9%	16.6%	-25.4%
345	22.4%	14.8%	-33.9%	22.5%	0.6%
3553	22.7%	9.7%	-57.3%	17.9%	-21.2%
183	24.4%	10.6%	-56.6%	18.0%	-26.0%
3494	27.6%	24.2%	-12.3%	30.5%	10.5%
167666	43.3%	24.2%	-44.1%	33.9%	-21.6%

Impact of referral counts on school-level interpretations (Hispanic students). In contrast to the analysis for African American students where school-level interpretations of disproportionality based on the duplicated and unduplicated counts were mostly similar, the results for Hispanic students revealed a little more difference between the two types of counts. In all, there were nine instances where the duplicated count and unduplicated count did not result in the same finding. In seven of those instances, the duplicated count reflected under-representation. This was similar to the findings for African American students, where in two of the three cases where there were discrepancies between counts, removing the effects of "frequent flyer" students (changing to the unduplicated count) resulted in a finding of over-representation when previously the findings indicated under-representation.

Differences between referral patterns between African American and Hispanic students. Generally, over-representation in office referrals was a much more common finding for African American students than it was for Hispanic students, and tended to be more intense. African American students had more extreme differences between their percent of duplicated referrals and their percent of school population, with maximum values for this measure reaching more than 11 times what would be expected. For Hispanic students, the difference reached just over six times what would be expected, but this was due to one outlier among the schools. After removing the outlier, the maximum difference between the percent of duplicated referrals and percent population for Hispanic students was a little more than 85% higher than would be expected.

For the percent of referred students (unduplicated count), African American students were found to be represented at almost three times their proportion of the student population. Hispanic students, on the other hand, were represented at a rate that was up to 70% higher than their proportion of the student population.

Hispanic students were found to be *under*-represented much more frequently than African American students, with more than half of the schools reporting this finding (70% of schools for the duplicated count, and 57% of schools for the unduplicated count). Only seven schools (19%) reported under-representation for African American students in the duplicated

count, and only six schools (16%) reported under-representation in the unduplicated count. Typically, under-representation of African American students was found in schools with small African American enrollment (less than 10% of the school population). Perhaps due to the higher number of schools with under-representation for Hispanic students, or perhaps due to the smaller range of Hispanic student enrollment (which at its largest was only 43% of the school's population) no such association was found for Hispanic students (i.e., under-representation and over-representation were found in schools with both small and large Hispanic student enrollment).

Differences between racial/ethnic groups were also found in the values of the different types of referral counts. While the duplicated count tended to produce higher values than the unduplicated count for African American students, the opposite was true for Hispanic students. Although there were exceptions in several schools, it seemed that it was less common for Hispanic students than African American students to be represented among those who receive multiple referrals.

Finally, changing from the duplicated to unduplicated count appeared to have more of an impact for school-level interpretations of disproportionality for Hispanic students than it did for school-level interpretations of disproportionality for African American students. There were nine instances in which the duplicated count and unduplicated count did not result in the same finding for Hispanic students. By contrast, changing from the duplicated to unduplicated count resulted in a different interpretation of disproportionality for African American students in just three schools.

Referrals by race/ethnicity: Students identified as White (duplicated count). The comparison of the percent of referrals accounted for by White students (the duplicated count) to their percent of the school's population is presented in Figure 13**.** All 39 schools in the final sample reported referrals for White students.

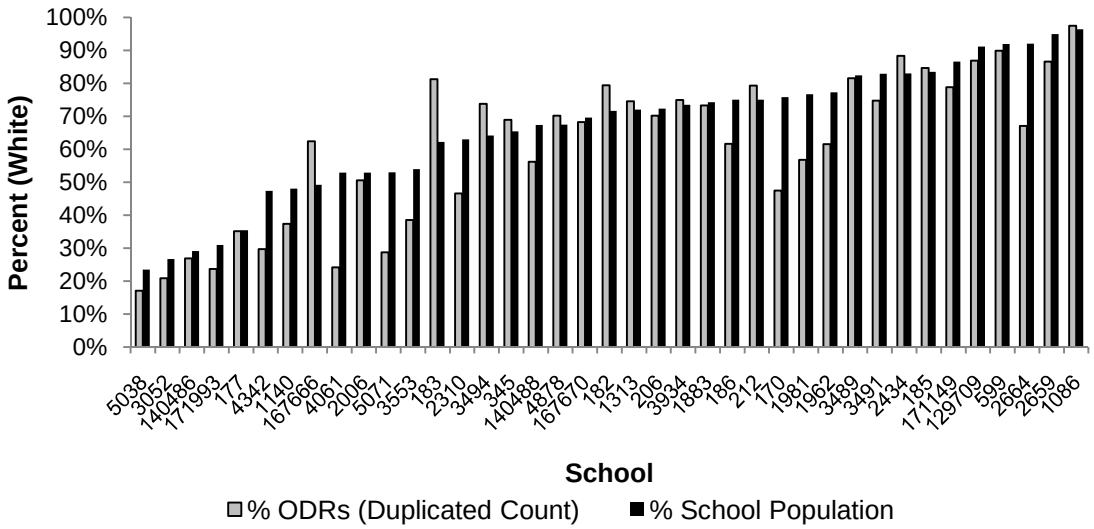

Figure 13. Percent of referrals accounted for by White students compared to their proportion of the school's population.

Under the duplicated count, White students were over-represented in 12 of the 39 schools (31% of the sample). The percentage of referrals accounted for by White students ranged from 17% to 97%, while White students' percent of their school population ranged from 24% to 96%. The school-by-school differences between percent of referrals accounted for by White students and their percent of school population ranged from -54.3% to 30.5%, meaning that at the lowest end of the distribution White students accounted for about 50% fewer referrals than would be expected, while at the highest end of the distribution White students accounted for just 30% more referrals than would be expected.

Students receiving referrals by race/ethnicity: Students identified as White (unduplicated count). Figure 14 compares the percent of students who received referrals who were identified as White (unduplicated count) to their percent of the school's population. With the unduplicated count, White students were over-represented in just seven schools (18% of the sample). The percent of students who received referrals who were identified as White ranged from 14% to 95%, while White students' percent of their school population ranged from 24% to 96%. The school-by-school differences between percent of students who received referrals who were White and percent of school population ranged from -38.9% to 23.4%, meaning that at the

lowest end of the distribution the percent of students who received referrals who were White was about 40% less than expected, while at the highest end of the distribution the percent of students who received referrals who were White was a little more than 20% higher than expected. Overall, with the presence of over-representation for White students identified in no more than 30% of the sample, disproportionality appears to be less common for White students than it is for either African American or Hispanic students, and when found it is more often based on the duplicated count.

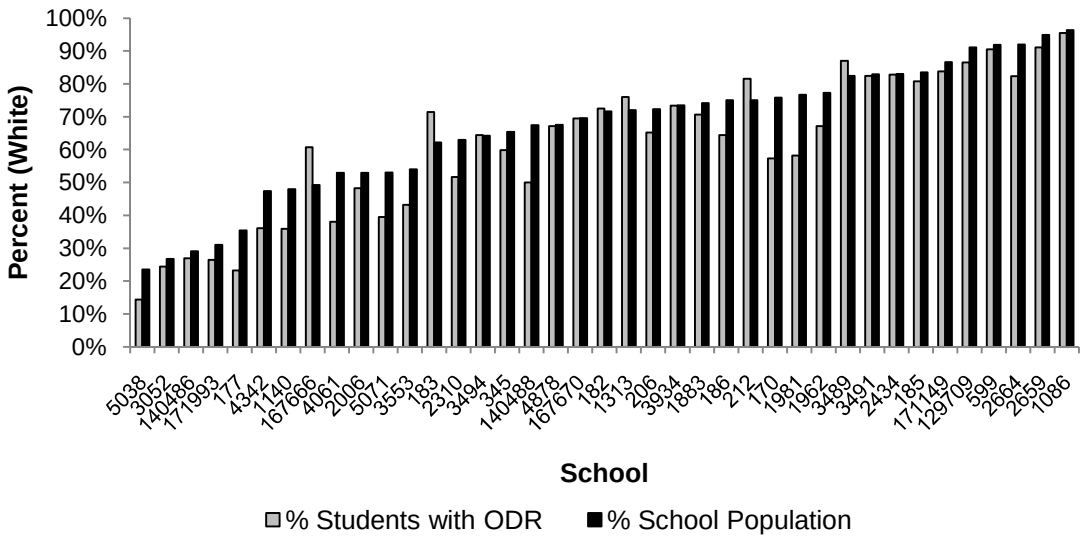

Figure 14. Percent of students who received an office referral who were identified as White compared to their proportion of the school's population.

Magnitude of over-representation of White students. In the twelve schools where White students received more referrals than would be expected (using the duplicated count), the average difference between the percent of referrals accounted for and the percent of the student population was 9.4% (*SD* = 9.9), meaning that on average, White students accounted for almost 10% more referrals than would be expected given their proportion of the student population. The range of over-representation in the duplicated count spanned from approximately 1% to 30%, while the range of over-representation in the unduplicated count was slightly smaller, spanning from *less than* 1% to 23.4% (*M* = 8.5%, *SD* = 8.1).

Magnitude of under-representation of White students. There were 27 schools (69% of the sample) with under-representation in their duplicated count for White students, with differences between the percent of referrals accounted for by White students and their percent of the school population that ranged from -0.8% to -54.3%. On average, schools with under-representation for White students had 18% fewer referrals than expected given their proportion of the student population (*SD* = 14.7). In the 32 schools (82% of the sample) with under-representation in their unduplicated count, the differences in the percent of students who received referrals who were White and their percent of the school population ranged from -0.1% to -38.9% (*M* = -12.7%, *SD* = 11.2). Overall, for White students under-representation was much more common, and in cases where White students were over-represented, it was to a much smaller degree than the magnitude of under-representation in referrals.

Differences between the duplicated and unduplicated counts (White students). Table 9 lists the school-by-school percentages for White students' composition within a school building, their percent of referrals, and their composition within the body of students at that school who received referrals. The differences between the duplicated and unduplicated counts for White students reflected mixed results. In 21 schools (54% of the sample), the duplicated count for White students was lower than their unduplicated count, suggesting that there was no clear pattern for which type of count would be more likely to produce higher levels of representation.

Table 9

Differences in Composition of School Enrollment, Referrals, and Students who Received Referrals for White students

School ID	Percent School Population	Percent of Referrals, Duplicated Count	Difference from Population *(Duplicated)*	Percent of Students Receiving Referrals, Unduplicated Count	Difference from Population *(Undupl.)*
5038	24%	17%	-27.2%	14%	-38.9%
3052	27%	21%	-21.7%	24%	-8.6%
140486	29%	27%	-7.6%	27%	-7.3%
171993	31%	24%	-23.5%	26%	-14.6%
177	35%	35%	-0.8%	23%	-34.4%
4342	47%	30%	-37.3%	36%	-23.8%
1140	48%	37%	-22.1%	36%	-25.2%
167666	49%	62%	26.8%	61%	23.4%
4061	53%	24%	-54.3%	38%	-28.1%
2006	53%	51%	-4.3%	48%	-8.7%
5071	53%	29%	-45.8%	40%	-25.4%
3553	54%	39%	-28.7%	43%	-20.1%
183	62%	81%	30.5%	71%	14.8%
2310	63%	47%	-26.0%	52%	-18.0%
3494	64%	74%	15.0%	64%	0.3%
345	65%	69%	5.4%	60%	-8.4%
140488	67%	56%	-16.6%	50%	-25.8%
4878	68%	70%	4.0%	67%	-0.6%
167670	70%	68%	-2.0%	69%	-0.2%
182	72%	79%	10.9%	73%	1.3%
1313	72%	75%	3.5%	76%	5.6%
206	72%	70%	-2.9%	65%	-9.8%
3934	74%	75%	1.9%	73%	-0.1%
1883	74%	73%	-1.2%	71%	-4.8%
186	75%	62%	-17.9%	64%	-14.1%
212	75%	79%	5.7%	82%	8.7%
170	76%	48%	-37.3%	57%	-24.4%
1981	77%	57%	-25.9%	58%	-24.2%
1962	77%	62%	-20.4%	67%	-13.1%
3489	82%	82%	-1.1%	87%	5.6%
3491	83%	75%	-9.9%	82%	-0.5%
2434	83%	88%	6.4%	83%	-0.2%
185	84%	85%	1.3%	81%	-3.3%
171149	87%	79%	-9.0%	84%	-3.3%
129709	91%	87%	-4.6%	87%	-5.0%
599	92%	90%	-2.2%	90%	-1.5%
2664	92%	67%	-27.1%	82%	-10.5%
2659	95%	87%	-8.7%	91%	-4.0%
1086	96%	97%	1.0%	95%	-1.0%

Observers may note that although the duplicated count of referrals for White students frequently produced smaller values than the unduplicated count, over-representation of White students was found more often when calculations were based on the duplicated count. This is because 10 of the 12 schools where the duplicated count resulted in over-representation of White students were also among the 18 schools where the duplicated count was higher than the unduplicated count. So although these findings may appear to be contradictory, the small number of schools involved supports both statements.

Impact of referral counts on school-level interpretations (White students). Similar to the findings for Hispanic students, interpretations of disproportionality sometimes depended on the type of count used. Overall, there were seven instances where the duplicated count and unduplicated count did not result in the same finding, and all but one of those instances were situations where the duplicated count resulted in an finding of over-representation. This stands in contrast to the findings for African American and Hispanic students, where most of the discrepancies in interpretations were instances where the duplicated count reflected under-representation.

Differences in referral patterns between African American, Hispanic and White students. Over-representation in office referrals was generally found less often for White students than it was for either African American or Hispanic students. When over-representation was found for White students it was less intense, with percent differences from their population no more than 30% higher than expected (for either the duplicated or unduplicated counts) – compared to maximum values that were 11 times higher than expected for African American students, and six times higher than expected for Hispanic students.

Under-representation was generally more common for White students than it was for both African American and Hispanic students, occurring in 69% of the sample using the duplicated count and 82% of the sample using the unduplicated count. However, Hispanic students had a similar number of schools reflecting under-representation when the duplicated count was used, with 70% of the schools reflecting this finding.

The duplicated and unduplicated referral counts for White students did not produce a clear pattern where one type of count resulted in a higher value in a large majority of schools. When the results from these three groups are considered in combination, it seems that Hispanic students are least likely to be represented among those who receive referrals as compared to African American or White students.

Changing from the duplicated to unduplicated count impacted school-level interpretations of disproportionality for Hispanic students as well as White students, but in opposite directions. For Hispanic students, seven of the nine instances in which the finding of disproportionality differed according to the type of count used were instances where the duplicated count reflected under-representation. In other words, for Hispanic students, removing the impact of "frequent flyer" students (i.e., switching to the unduplicated count) resulted in a finding of over-representation. By contrast, when there was a discrepancy in interpretation for White students, the duplicated count usually resulted in a finding of over-representation, suggesting that in these schools White students may have been relatively more likely to receive multiple referrals.

Overall, the findings suggest that while over-representation in office referrals can be an issue for students who are identified as African American, Hispanic, *and* White, it is more common and dramatic for African American students than it is for Hispanic or White students. Over-representation for African American students was found in almost every school in the sample, and generally persisted regardless of the type of count used in the calculations. In addition, the findings suggest that Hispanic students may be less likely to receive multiple referrals than either African American or White students. Table 10 summarizes the differences in findings for each racial group.

Table 10

Summary of Differences in Referrals by Racial Group

	African American Students		Hispanic Students		White Students	
	Min.	Max.	Min.	Max.	Min.	Max.
Percent of Student Body	0.7%	70.8%	2.2%	43.3%	24%	96%
Number of Schools with Over-Representation based on the Duplicated Count	30 / 37		10 / 37		12 / 39	
Number of Schools with Over-Representation based on the Unduplicated Count	31 / 37		16 / 37		7 / 39	
Degree of over-representation in referrals *(Difference between percent of referrals accounted for and percent of student body)*	7%	1,114.3%	8%	608.8%	1%	30%
Degree of over-representation in students who received a referral *(Difference between racial group composition in students who received a referral and percent of student body)*	8.4%	278.8%	0.6%	73%	< 1%	23.4%
Degree of under-representation in referrals	-3.8%	-64.3%	-2.1%	-78.5%	-0.8%	-54.3%
Degree of under-representation in students who received a referral	-9.4%	-72.2%	-0.4%	-73.1%	-0.1%	-38.9%
Representation in duplicated count compared to representation in unduplicated count	Mostly higher		Mostly lower		Mostly lower	
Impact of changing from duplicated count to unduplicated count on school-level interpretations of disproportionality	Mostly none *(3 schools)*		Little *(9 schools)*		Little *(7 schools)*	

Risk ratios: Students identified as African American. While a comparison analysis may in some ways be a straightforward way to examine over- and under-representation, the method suffers from notable shortcomings. For one, a simple comparison will only provide information about a single group. Understanding that one group has over- or under-representation is important, but that information by itself may not be sufficient to justify a re-distribution of resources to address the problem. Over- and under-representation must be

compared across all groups of students in order to prioritize the level of disparity and target resources accordingly. Comparison analyses, because they require two separate pieces of information (percent in a target category and percent of the population), are difficult to use in this way, especially when large numbers of schools and multiple categories of students are involved.

Fortunately, the risk ratio provides a single unit that encapsulates disparities across all groups of students. A measure of one group's risk compared to the risk of all other students, the risk ratio simultaneously summarizes information about a target group's demographic distribution (such as percent of student population identified as African American) and composition within a category (such as percent of students who received an office referral who were identified as African American) and provides information about how this relates to other groups within the population. In the current study, the risk ratio was calculated as follows:

$$\text{Risk Ratio} = \frac{\text{Risk for racial group}}{\text{Risk for comparison group}} =$$

$$\frac{(\text{\# African American students who received a referral}) \div (\text{\# African American students in the school})}{(\text{\# of all other students who received a referral}) \div (\text{\# of all other students in the school})}$$

Risk ratios of 1.0 indicate that there is no difference between the numerator and denominator – that there is no difference between the target group and other students (e.g., between African American students who received referrals and all other students who received referrals). Risk ratios higher than 1.0 indicate over-representation (e.g., a risk ratio of 1.5 indicates a 50% higher risk of receiving a referral, or a risk ratio of 2.0 indicates twice the risk of receiving a referral), and risk ratios lower than 1.0 indicate under-representation (e.g., a risk ratio of .75 indicates a target group would be 25% *less* likely to receive a referral). The risk ratios for office discipline referrals for African American students are displayed in Figure 15. All 39 schools

(those which had racial group information on at least 90% of their office discipline referrals) were included in this analysis, regardless of whether they reported any referrals for African American students.

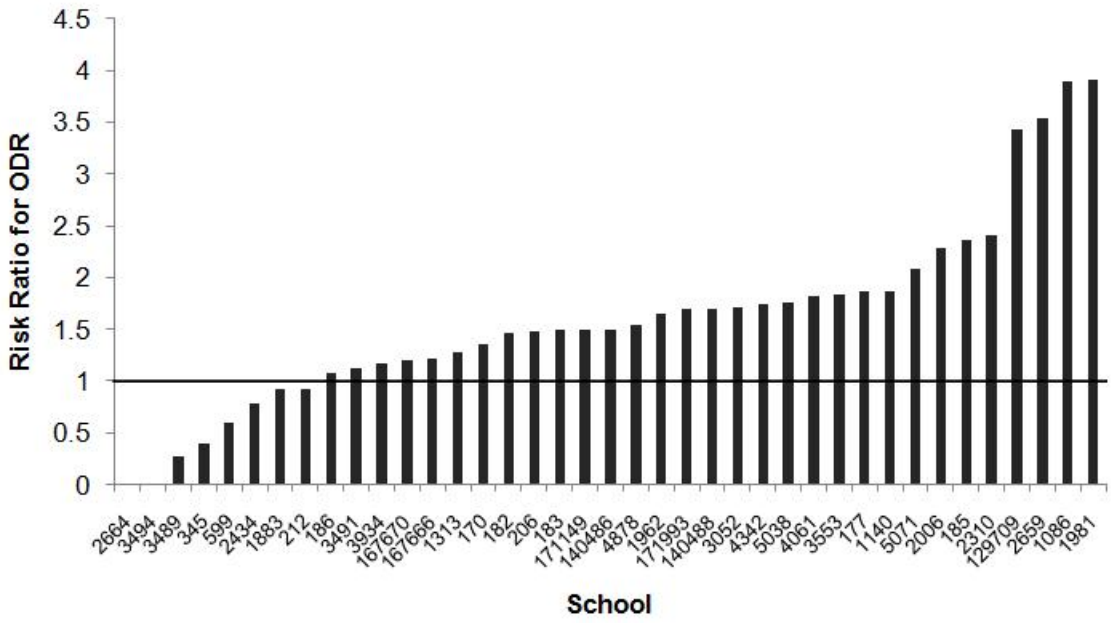

Figure 15. Risk ratios for African American students' office discipline referrals.

The results for African American students' risk ratios for referrals found over-representation in 31 of the 39 schools – which was the same finding produced by the unduplicated comparison analysis. This would be expected, as the unduplicated count is based on the same unit of analysis: the number of students who received at least one referral, as opposed to the number of referrals accounted for by students (i.e., the duplicated count). However, the risk ratio analysis provides additional information: namely, that in 20 of the 39 schools (51% of the sample) African American students were at least 50% more likely than students from all other racial groups to receive an office discipline referral. In eight schools (21% of the sample), African American students were more than *twice* as likely as students from all other racial groups to receive an office discipline referral. In the most extreme case, African

American students were almost four times (3.91) as likely as all other students to receive an office discipline referral.

The school with the most extreme value for African American risk ratio (3.91) was also one of the schools in which the comparison analysis indicated one of the highest levels of disproportionality existed. In the comparison analysis for unduplicated referrals, school #1981 showed that 34.7% of the students who received office referrals at this school were identified as African American. African American students made up 12% of the school population, providing an example where African American students received almost twice (1.9 times) as many referrals as would be expected given their proportion of the student population. Even though this is a large difference, this was not the largest difference seen in the comparison analysis across all of the schools: three schools had differences for African American students that were even larger (2.2, 2.6, and 2.8). However, once the referral patterns for other groups of students within these schools are taken into account (as is the case when the risk ratio is calculated), the risk for African American students within these three schools was comparatively smaller than was is at school #1981. In other words, what appears to be a large problem in a comparison analysis may not be as large of a problem in more relative terms.

The importance of placing findings of representation in relation to other groups also works in the opposite direction. Table 11 provides an example of three schools for which the risk ratio is approximately the same. In each of the three schools, African American students were found to have about a 50% higher risk of receiving a referral compared to all other students in the building. In two of those schools (#171149 and #4878), there were about 40% more African American students who received referrals than would be expected given their proportion of the student population. However, in the third school, there were only about 20% more African American students who received referrals than would be expected. So even though African American students at all three schools had about the same magnitude of risk compared to all other students in their building, the extent of the over-representation within their building was noticeably different – in two of the schools, noticeably more African American students were involved. Further, the two schools with similar degrees of over-representation in the unduplicated

count had vastly different degrees of over-representation in the duplicated count, indicating that at one school, it was much more common for African American students to receive multiple referrals. While these differences don't change the overall finding of "over-representation" for any one of the schools, they would likely impact how each of the schools needs to intervene. Consideration of the risk ratio and composition analyses together results a more accurate picture than any of the measures alone. A table listing African American students' risk ratios and magnitudes of disproportionality from the comparison analysis for can be found in Appendix D.

Table 11

Relationship of Risk Ratio and Comparison Analysis Findings for Referrals, African American students

School ID	Risk Ratio	Unduplicated Comparison	Duplicated Comparison
171149	1.49	46.5%	148.6%
140486	1.50	22.8%	32.7%
4878	1.54	43.5%	50%

Note: Unduplicated Comparison and Duplicated Comparison reflect the percent difference of each of these measures compared to African American students' proportion of the student population (i.e., "magnitude of disproportionality").

Risk ratios: Students identified as Hispanic*.* Figure 16 displays risk ratios for students who were identified as Hispanic. Information from all 39 schools in the final sample was included, regardless of whether a school reported any referrals for Hispanic students.

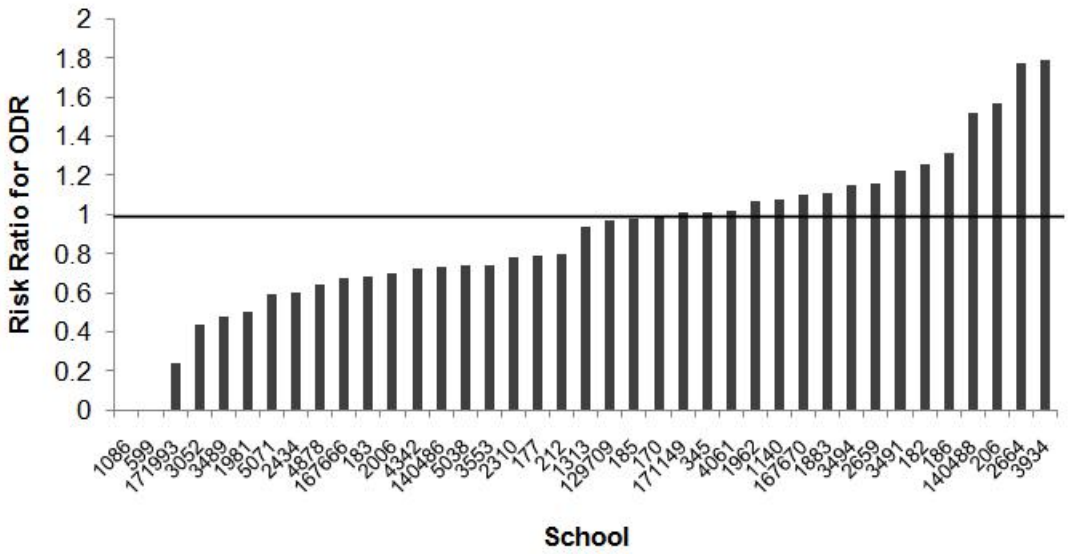

Figure 16. Risk ratios for Hispanic students' office discipline referrals.

The results for Hispanic students' risk ratios for referrals found over-representation in 16 of the 39 schools (41% of the sample). The relative degree of over-representation was much less severe than it was for African American students: in the most extreme case, Hispanic students were 79% more likely than all other students to receive a referral (risk ratio = 1.79). Overall, there were only four schools where the risk ratio for Hispanic students was higher than 1.5, indicating a 50% or higher risk for these students to receive a referral compared to students from all other groups.

The risk ratio results supported the findings of the comparison analysis for Hispanic students' unduplicated referrals, but there was not an example where schools with similar risk ratios had dissimilar results in the unduplicated comparison analysis, as was found for African American students. However, this may be due to the finding that overall, there was less variation in the range of values for the magnitudes of disproportionality in the unduplicated count for Hispanic students (i.e., the maximum value was 73.0%, compared to 278.8% for African American students). A table listing Hispanic students' risk ratios and magnitudes of disproportionality from the comparison analysis for can be found in Appendix E.

Comparison of African American students' and Hispanic students' risk ratios.

When each school's risk ratio for African American students and Hispanic students was plotted together, a general pattern emerged: as the risk ratios for African American students grew larger, the risk ratios for Hispanic students remained low (frequently under 1.0). Further, whenever a risk ratio for either racial group exceeded 1.5, the risk ratio for the other racial group fell at or below that value, suggesting that when disproportionality reaches higher magnitudes, there tends to be only one racial group which experiences a relatively intense problem. Figure 17 shows each school's risk ratios for African American and Hispanic students.

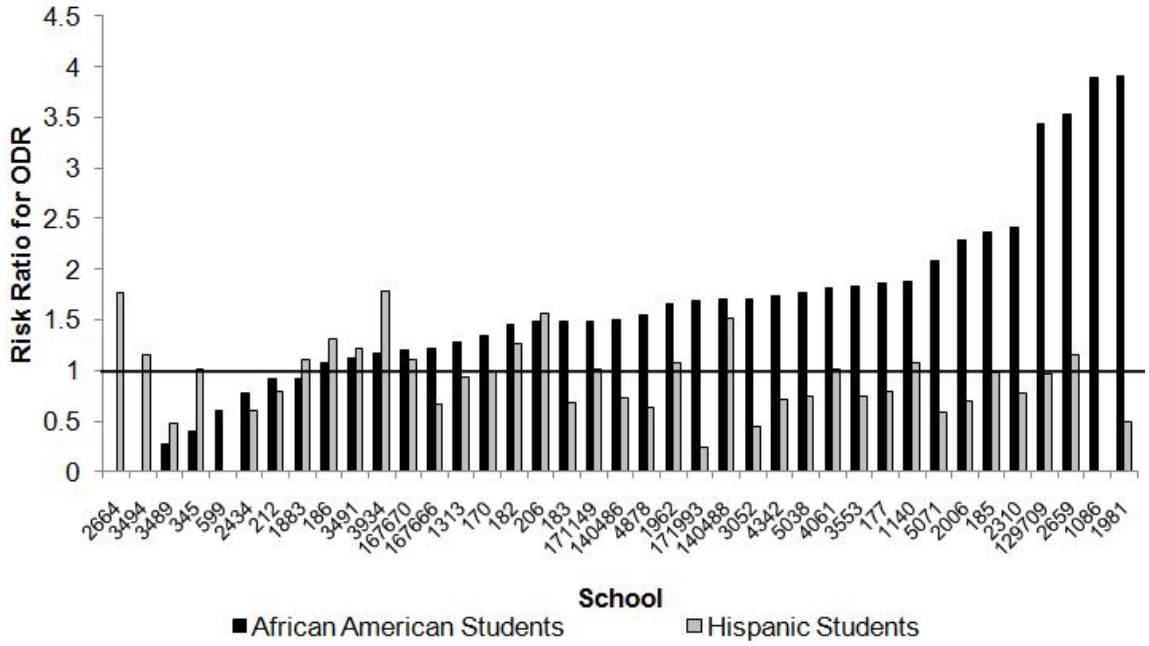

Figure 17. School-by-School distribution of African American and Hispanic students' risk ratios for office discipline referrals.

School-Level Distributions of Out-of-School Suspensions

Overall suspension rates (duplicated count). Turning to out-of-school suspensions (OSS), the overall suspension rates were calculated for the original sample (n=83 schools). The degree to which out-of-school suspensions (OSS) were utilized within a school building was measured by the suspension rate, which was calculated by dividing the number of OSS events by

the number of students enrolled in the school. Just as with the referral rate, the suspension rate gives a rough measure of how many suspensions could have been given per student at a school. For instance, if a school had a suspension rate of 1.0, an observer might interpret that to mean that enough suspensions were issued so that every student in the school would have an opportunity to receive one. It should be noted that the suspension rate for this study did *not* include a measure of how long each suspension might last; each suspension event could have lasted a half day, a whole day, or multiple days. Figure 18 shows the distribution of suspension rates by school based on a duplicated count of suspension events (i.e., if a student received multiple suspensions, all events were counted in the total).

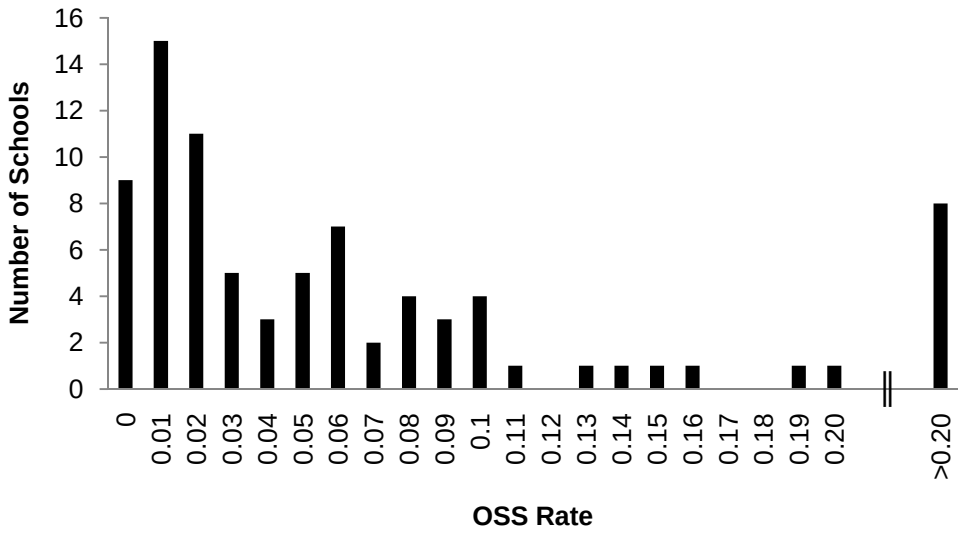

Figure 18. Overall suspension rates by school (duplicated count).

The distribution of duplicated suspension rates ranged from zero to 1.18 suspensions per student (*M* = .09, *SD* = .17). Nine schools had an OSS rate of zero: of these, six schools reported only one or two OSS events for the year, which resulted in suspension rates smaller than 0.01. The other three schools reported zero suspensions for the year. The distribution of OSS rates was significantly positively skewed and centered around the mean (skew = 4.28, *SE* of skew = .264; kurtosis = 22.65, *SE* of kurtosis = .523). In fact, rates at the upper end of the

distribution were so much more spread out than the rates at the lower end of the distribution, that only 10% of the sample had OSS rates larger than 0.20 suspensions per student (eight schools had rates which spanned from .27 to 1.18). Considering the shape of the distribution, it is likely that the few schools at the higher end of the distribution had a large impact on the distribution's mean value. Overall, it appears that the schools in the original sample had such a low rate of suspensions that for all but eight schools a suspension could have been given for fewer than one out of every five students.

Overall suspension rates (unduplicated count). To get a sense of the degree to which students who received more than one suspension impacted the schools' overall suspension rates, a second analysis was performed with an unduplicated count of suspensions. The resulting distribution is shown in Figure 19.

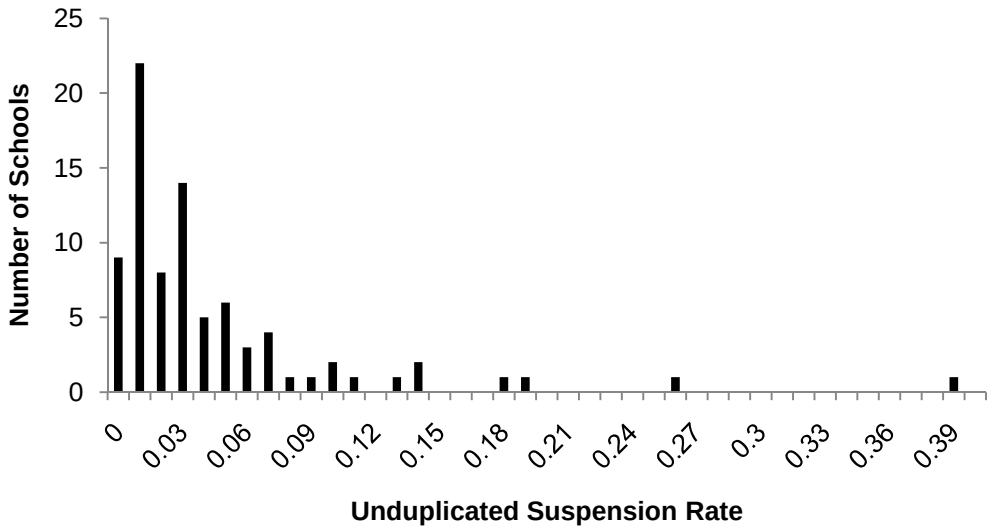

Figure 19. Suspension rates (unduplicated count) by school.

Changing from the duplicated count of OSS to the unduplicated count of OSS did *not* result in the OSS rates bearing more of a resemblance to a normal distribution, as it did with the referral rate. The distribution of unduplicated suspension rates ranged from zero to .39 OSS per student (M = .04, SD = .06). Loosely translated, the school with the highest suspension rate

based on the unduplicated count issued a suspension for approximately every 2.5 students. The distribution of OSS rate based on the unduplicated count remained significantly positively skewed and centered around the mean (skew = 3.27, *SE* of skew = .264; kurtosis = 13.82, *SE* of kurtosis = .523).

Differences in the duplicated and unduplicated counts (overall suspension rate). Changing from the duplicated to unduplicated count of suspensions did not have as dramatic an impact on the distributions of suspension rates as it did for referrals. Table 12 outlines the differences in distributions of OSS rates based on the different types of counts. The overall range of suspension rates was small for both counts, reaching no more than just over one OSS per student under the duplicated count of suspension. Compared to the range of referral rates, which under the duplicated count had a maximum value of 7.76 referrals per student, the lack of variation between duplicated and unduplicated counts of suspension does not seem surprising: with such small values, there was little room for change. Still, it was not uncommon for the unduplicated count of suspension resulted in smaller values than the duplicated count, suggesting that some students in this sample of schools received multiple suspensions.

Table 12

Characteristics of OSS Rates (OSS per student) Based on Duplicated and Unduplicated Counts of Suspensions

	Duplicated Count	Unduplicated Count
Minimum	0	0
Maximum	1.18	.39
Mean	.09	.04
Standard Deviation	.17	.06

Suspensions by gender (duplicated count). The percentage of each school's suspensions accounted for by both genders is reflected in Figure 20. It should be noted that three schools did not report any suspensions, so were excluded from the analysis. Across the sample (n = 80), the percentage of suspension events accounted for by boys ranged from zero to 100%

(*M* = 84.6, *SD* = 20.5), while the percentage of suspension events accounted for by girls also ranged from zero to 100% (*M* = 15.4, *SD* = 20.5). As can be surmised by the two ranges, there were occasions for both genders where all suspensions were assigned to the other, but this happened for girls in only two schools (compared to 25 schools where boys accounted for 100% of the suspensions). Across the sample, boys received on average over five times as many suspension events as girls.

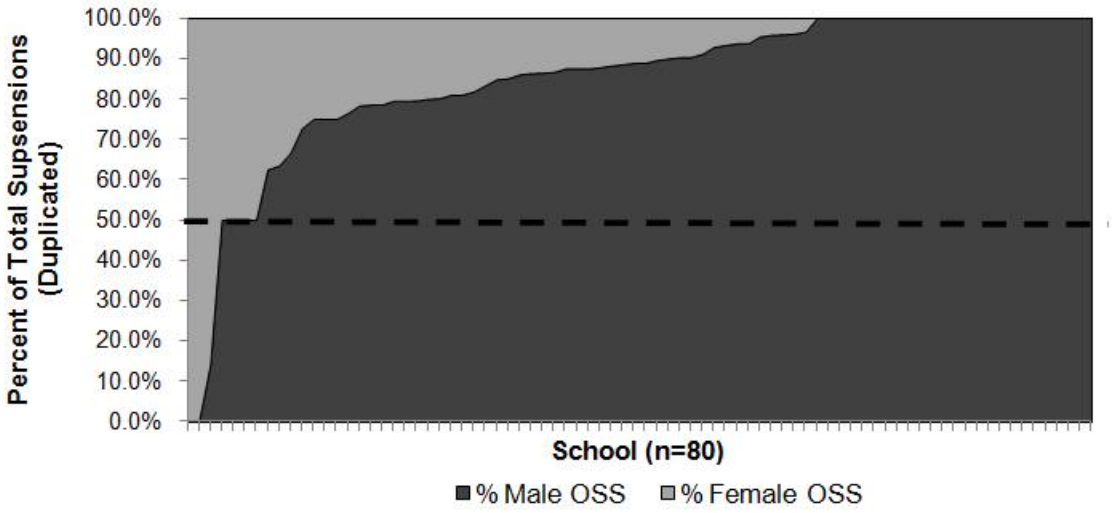

Figure 20. Percentage of referrals accounted for by gender (duplicated count).

Typically (but not exclusively), schools with a small number of total suspension events had more extreme gender disproportionality than schools with a higher number of events. For example, 18 of the 27 schools that showed either boys or girls accounting for 100% of the suspensions had seven or fewer suspension events. Of the schools that had more variation on the percent of suspensions accounted for by each gender, there were only three examples where the total number of suspension events was that small

Unlike the earlier analysis of referrals by gender, the analysis of suspensions by gender revealed a few instances where girls accounted for more of the consequence than boys. In total, there were three schools in which girls received more OSS than boys, two of which reflected

outcomes where girls accounted for 100% of the OSS events. In the third school, girls accounted for just over 85% of the OSS events. In all three of these instances, the total number of OSS events was small: one school reported seven events, another school reported four events, and the third school reported only one event (this was one of the schools with 100% of events accounted for by girls).

In addition, there were four schools in which each gender accounted for an equal number of suspensions. In these instances, each school had only two suspensions – one that was assigned to a boy, and one that was assigned to a girl. Across the sample, only seven schools (8% of the sample) had an equal or greater number of suspensions assigned to girls than they did to boys.

Students receiving suspensions by gender (unduplicated count). The percentage of students who received a suspension who were male and female is reflected in Figure 21. Just as with the duplicated count of suspensions, school totals reflected the same range for boys and girls (ranging from 0% to 100%), but with dramatically different mean values. On average, the percent of students who received suspension who were male was 82.6% (*SD* = 20.2%), while the average percent who were female was 17.4%. These values were slightly lower than the values found under the duplicated count, and can be taken to mean that on average boys were represented among the students who received suspensions at a rate that was about 65% higher than expected (assuming approximately equal proportions of each gender in the school populations). Based on these findings, as with referrals, many more boys receive suspension than girls, but the number of schools reflecting over-representation is not quite as consistent as it was in the referral analysis.

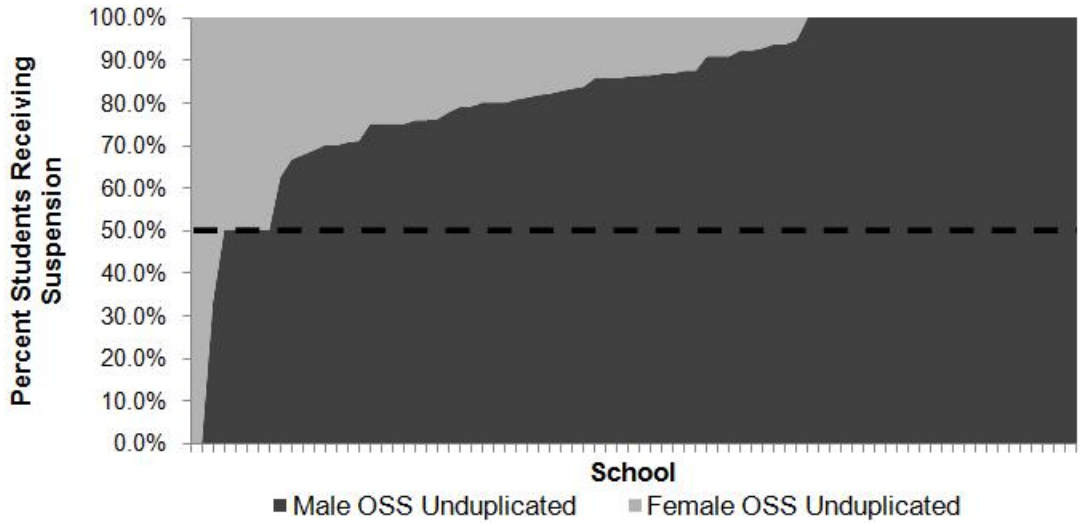

Figure 21. Percentage of students who received suspension by gender (unduplicated count).

Differences between the duplicated and unduplicated suspension counts (gender). Table 13 summarizes the differences in the gender analysis findings based on the duplicated and unduplicated counts of suspensions. In contrast to earlier analyses of values between the duplicated and unduplicated counts, the analysis of suspensions by gender reflected a noticeable number of schools in which both counts produced the same value. Typically, the total number of suspensions in these schools was small (see earlier discussion). In total, there were 34 schools (43% of the sample) in which the duplicated and the unduplicated counts were the same. In all but three instances, the percentage of suspensions accounted for was either 100% or 50%. In addition to this finding, the results showed that the duplicated count produced a higher rate of representation for boys in 45% of the sample (37 schools), but only in 11% of the sample for girls (9 schools). This suggests that in general girls are less likely to be represented among students who receive multiple suspensions than boys, which is consistent with the findings of the analysis of referrals by gender. However, there was more variation for suspension with regards to the extent to which boys received multiple suspensions. Given that 58% of the schools in the sample did *not* experience higher magnitudes of representation for boys when examining the duplicated count, this suggests that boys were not as likely to receive multiple suspensions as they were to

receive multiple referrals, where the duplicated count produced higher representation in all but four schools.

Table 13

Characteristics of Suspension Events by Gender Based on Duplicated and Unduplicated Counts of Suspensions

	Duplicated Count		Unduplicated Count	
	Boys	Girls	Boys	Girls
Minimum	0	0	0	0
Maximum	100	100	100	100
Mean	84.6	15.4	82.6	17.4
Standard Deviation	20.5	20.5	20.2	20.2

Impact of suspension counts on school-level interpretations (gender). The change from the duplicated to unduplicated count of suspensions did not impact school-level interpretations of disproportionality in suspensions as it did with disproportionality in referrals. Schools with the most proportional distributions of suspensions assigned 50% of events to each gender, which did not change based on the duplicated or unduplicated counts. The distribution of suspension events likewise didn't vary in schools with the most extreme levels of disproportionality (where one gender received 100% of the suspensions). The impact of students who receive multiple suspensions was difficult to determine for schools that fell in between the most proportional distribution and the most extreme distribution through visual analysis alone. Given the lack of a clear pattern, the influence of "frequent flyer" students on school-level interpretations of gender disproportionality in suspensions remains to be determined. A list of schools rank-ordered according to the percent of suspension events accounted for by boys can be found in Appendix F.

Suspensions by Race/Ethnicity

Suspensions by race/ethnicity: Students identified as African American (duplicated count). Returning to an examination of referrals by race and ethnicity, comparison analyses were

performed for suspension events for the 39 schools which kept track of race/ethnicity information on at least 90% of their referrals. The comparison of the percent of suspensions accounted for by African American students (based on the duplicated count) to their percent of the school's population is presented in Figure 22. It should be noted that the total number of schools displayed in this part of the analysis was only 25, as fourteen of the schools with mostly complete data did not report any suspensions for African American students.

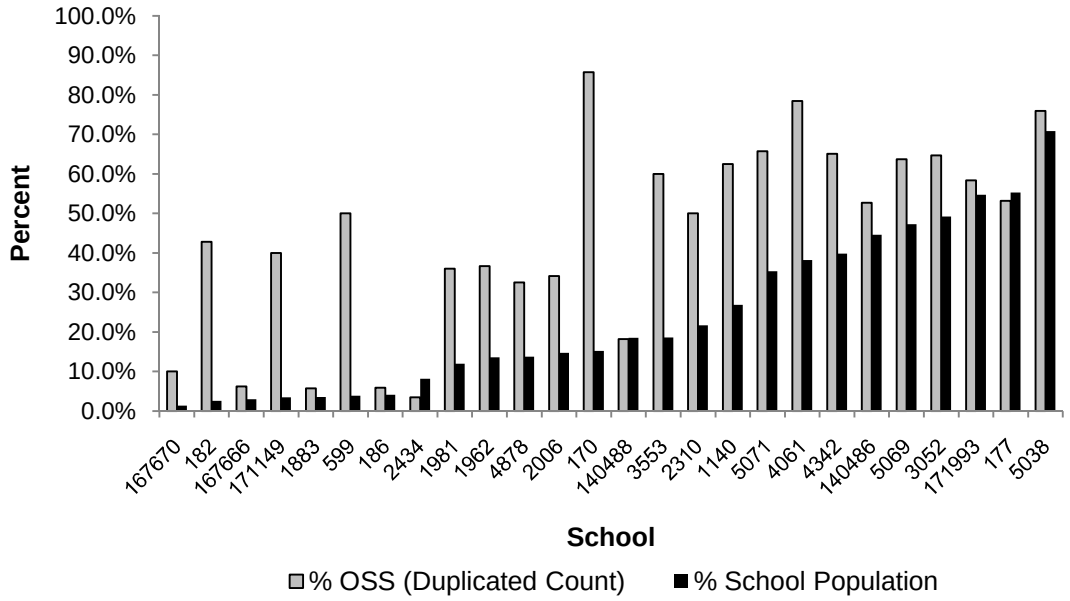

Figure 22. Percent of suspension events accounted for by African American students compared to their proportion of the school's population.

With the duplicated count of suspension events, African American students were over-represented in all but three schools (88% of this sample). The percentage of suspension events accounted for by African American students ranged from 3.4% to 85.7%, while African American students' percent of their school population ranged from 1.4% to 70.8% in this set of schools. The school-by-school differences between percent of referrals accounted for by African American students and percent of population ranged from -57.9% to 1,548.4%, meaning that at the lowest end of the distribution African American students accounted for almost 58% fewer suspension events than would be expected given their proportion of the school population, while at the

highest end of the distribution African American students accounted for more than 15 times more suspension events as would be expected.

Students receiving suspensions by race/ethnicity: Students identified as African American (unduplicated count). Figure 23 compares the percent of students who received suspension who were identified as African American (unduplicated count) to their percent of the school's population. Using the unduplicated count, African American students were over-represented in all but one of the 25 schools in this sample (96%). The percent of students who received suspension who were identified as African American ranged from 6.7% to 79.3%, while African American students' percent of their school population ranged from 1.4% to 70.8% in this set of schools. The school-by-school differences between percent of students who received suspension and percent of population ranged from -18.7% to 920.4%, meaning that at the lowest end of the distribution the percent of students who received suspension who were African American was about 18% less than expected based on their proportion of the school population, while at the highest end of the distribution the percent of students who received suspension who were African American was over nine times higher than expected. Overall, in the current sample African American students were over-represented in suspension events in almost every school, regardless of the type of count used in the population.

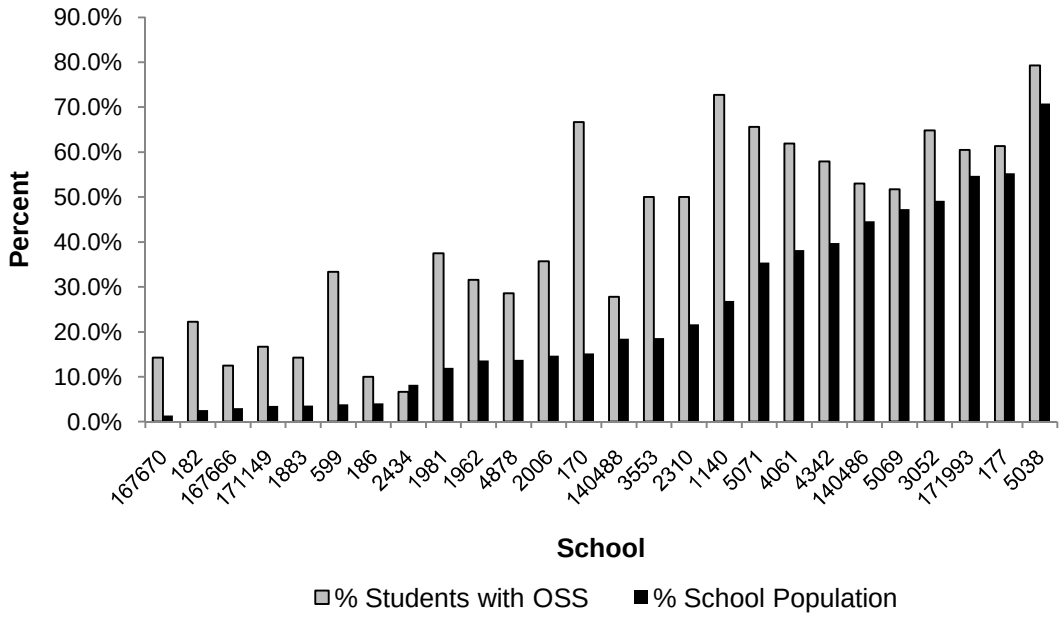

Figure 23. Percent of students who received suspension who were identified as African American compared to their proportion of the school's population.

Magnitude of over-representation of African American students. In schools where there was over-representation, the average difference between the percent of referrals accounted for by African American students and their percent of the student population was just over 295% (*SD* = 424.4), meaning that on average, in the 22 schools with over-representation in their duplicated count, African American students accounted for almost three times as many suspension events as would be expected given their proportion of the student population. The range of over-representation in the duplicated count spanned from 6.6% more suspensions than expected to more than 15 times as many suspensions than would be expected. The range of over-representation in the 25 schools using the unduplicated count began at just over 10% more African American students than expected, and ended with more than nine times the number of African American students as would be expected (*M* = 220.6%, *SD* = 252.9).

Magnitude of under-representation of African American students. There were only three instances using the duplicated count of suspensions that resulted in under-representation of African American students, and only one instance using the unduplicated count. With the

duplicated count, the differences between the percent of suspensions accounted for and the percent of school population ranged from -1.7% to -57.9% (X = 21.2%, SD = 31.9). With the unduplicated count there was only *one* school with under-representation, such that the difference between the percent of students who received a suspension who were identified as African American and their percent of the school population was 18.7% less than expected. The three schools with under-representation had varying sizes of African American student populations (8%, 18% and 55%), so there did not appear to be any relationship between under-representation in suspension events and small African American enrollment – or between over-representation and enrollment, for that matter.

Differences between the duplicated and unduplicated counts (African American students). Table 14 lists the school-by-school breakdowns of African American students' composition within a school building, their percent of suspension events, and their composition within the body of students who received suspension. African American students received a smaller percentage of suspensions using the duplicated count in 60% of the sample (15 of the 25 schools), suggesting that fewer African American students received multiple suspensions. Looking back at the referral analysis, African American students accounted for a lower percentage of referrals using the duplicated count in about a third of the sample, suggesting that in general it is more common for African American students to receive multiple referrals than it is for this group to receive multiple suspensions

Table 14

Differences in Composition of School Enrollment, Suspensions, and Students who Received Suspensions for African American Students

School ID	Percent School Population	Percent of Suspensions, Duplicated Count	Difference from Population *(Duplicated)*	Percent of Students Receiving Suspension, Unduplicated Count	Difference from Population *(Undupl.)*
167670	1.4%	10.0%	614.3%	14.3%	920.4%
182	2.6%	42.9%	1548.4%	22.2%	754.7%
167666	3.0%	6.3%	108.3%	12.5%	316.7%
171149	3.5%	40.0%	1042.9%	16.7%	376.2%
1883	3.6%	5.7%	58.7%	14.3%	296.8%
599	3.9%	50.0%	1182.1%	33.3%	754.7%
186	4.1%	5.9%	43.5%	10.0%	143.9%
2434	8.2%	3.4%	-57.9%	6.7%	-18.7%
1981	12.0%	36.0%	200.0%	37.5%	212.5%
1962	13.6%	36.7%	169.6%	31.6%	132.2%
4878	13.8%	32.6%	135.9%	28.6%	107.0%
2006	14.7%	34.1%	132.3%	35.7%	143.0%
170	15.2%	85.7%	463.9%	66.7%	338.6%
140488	18.5%	18.2%	-1.7%	27.8%	50.2%
3553	18.6%	60.0%	222.6%	50.0%	168.8%
2310	21.7%	50.0%	130.4%	50.0%	130.4%
1140	26.9%	62.5%	132.3%	72.7%	170.4%
5071	35.4%	65.7%	85.6%	65.6%	85.4%
4061	38.2%	78.4%	105.4%	61.9%	62.1%
4342	39.8%	65.1%	63.5%	57.9%	45.5%
140486	44.6%	52.7%	18.2%	53.0%	18.8%
3052	49.2%	64.7%	31.5%	64.9%	31.8%
171993	54.7%	58.3%	6.6%	60.5%	10.7%
177	55.3%	53.2%	-3.8%	61.4%	11.0%
5038	70.8%	76.0%	7.3%	79.3%	11.9%

Impact of suspension counts on school-level interpretations (African American students). The change from duplicated to unduplicated count resulted in a different interpretation of disproportionality in only two cases. Both times, the duplicated count reflected a small degree of under-representation (percent differences from the student population were -1.7% and -3.8%), while the unduplicated count resulted in a finding of over-representation (with percent differences of 50.2% and 11%). As with referrals, while the finding of over-representation

in suspensions for African American students is overwhelmingly consistent, the type of count used in the calculation has the potential to influence the finding.

Suspensions by race/ethnicity: Students identified as Hispanic (duplicated count). The comparison of the percent of suspensions accounted for by Hispanic students (the duplicated count) to their percent of the school's population is presented in Figure 24. A total of 21 schools reported suspension events for Hispanic students, seven of which did not report any suspension events for African American students. In addition, eleven schools reported suspension events for African American students, but none for students who were identified as Hispanic. In all, fourteen of the same schools were represented between the suspension analysis for African American students and Hispanic students.

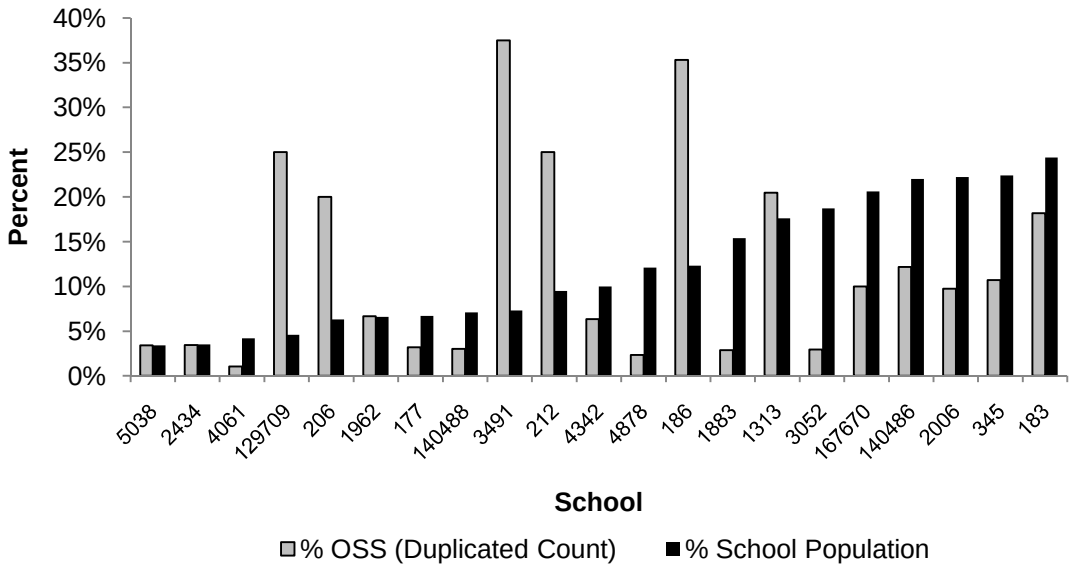

Figure 24. Percent of suspensions accounted for by Hispanic students compared to their proportion of the school's population.

Using the duplicated count, Hispanic students were over-represented in seven of the 21 schools (33% of this sample). The percentage of suspensions accounted for by Hispanic students ranged from 1.1% to 37.5%, while their percent of the student population ranged from 3.4% to 24.4%. The school-by-school differences between percent of suspensions and percent of

population ranged from -84.3% to 443.5%, meaning that at the lowest end of the distribution Hispanic students accounted for close to 85% fewer suspensions than would be expected given their proportion of the student population, while at the highest end of the distribution Hispanic students accounted for close to 4.5 times as many suspensions as would be expected.

Students receiving suspensions by race/ethnicity: Students identified as Hispanic (unduplicated count). Figure 25 compares the percent of students who received suspensions who were identified as Hispanic (unduplicated count) to their percent of the school's population. Using the unduplicated count, Hispanic students were over-represented in nine of the 21 schools (43% of the sample). The percent of students who received suspension who were identified as Hispanic ranged from 3.6% to 40%, while Hispanic students' percent of their school population ranged from 3.4% to 24.4%. The school-by-school differences between the percent of students who received suspension who were identified as Hispanic and the percent of their school population ranged from -71.1% to 624.6%, meaning that at the lowest end of the distribution the percent of students who received suspensions who were Hispanic was just over 70% less than expected, while at the highest end of the distribution the percent of students who received suspension who were Hispanic was over six times higher than expected. Overall, findings of over-representation for Hispanic students in suspension was similar to what was found when examining referrals, occurring in less than 50% of the schools in this sample, with the small differences in the number of schools with over-representation based on the type of count used.

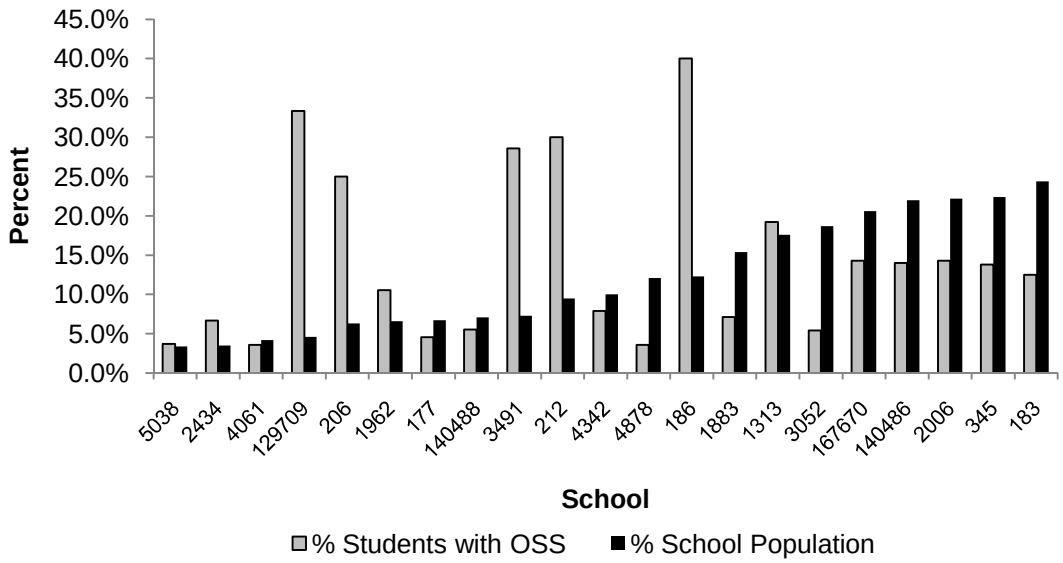

Figure 25. Percent of students who received a suspension who were identified as Hispanic compared to their proportion of the school's population.

Magnitude of over-representation of Hispanic students. In the seven schools where Hispanic students received more suspensions than expected (using the duplicated count), the average difference between the percent of suspensions accounted for and the percent of the school population was 206% (*SD* = 173.2), meaning when there was over-representation for Hispanic students in suspension events, on average they accounted for about two times as many suspensions as would be expected given their proportion of the student population. The range of over-representation in the duplicated count spanned from suspensions that were just 1% higher than expected to almost four and a half times higher (443.5%) than expected. There did not appear to be any schools with values that set them apart from the rest of the sample, as was found in the referral analysis (i.e., there were no outliers among the schools).

The range of over-representation with the unduplicated count (nine schools) spanned from 8.9% to 624.6% (*X* = 202.4%, *SD* = 194.8). A closer examination of the range of over-representation revealed that the school with the maximum difference (624.6%) appeared to be something of an outlier, as the next highest magnitude of over-representation was only 296.8%, or almost three times higher than expected. If one were to remove this school, the range for the

unduplicated count became smaller than the range for the duplicated count (the range for these eight schools would span from 8.9% to 296.8%, with an average of 149.7% and *SD* of 121.3).

Magnitude of under-representation of Hispanic students. In the 14 schools where Hispanic students received fewer suspensions than expected (under the duplicated count), the magnitude of under-representation ranged from -0.1% to -84.3% (*M* = -49.9%, *SD* = 27). Using the unduplicated count (12 schools), the magnitude of under-representation ranged from -15% to -71% (*M* = -39.6%, *SD* = 18.2). Overall, the unduplicated count resulted in a slightly smaller number of schools with under-representation, a smaller range of values for the magnitude of under-representation, and a slightly smaller average. In spite of these differences, the overall outcomes remained similar between the two types of counts.

Differences between the duplicated and unduplicated counts (Hispanic students). Table 15 lists the school-by-school percentages for Hispanic students' composition within a school building, their percent of suspensions, and their composition within the body of students at that school who received suspension. The duplicated count produced a smaller value than the unduplicated count in all but three schools in the sample, suggesting that overall there were relatively fewer Hispanic students who received multiple suspensions. This was similar to the finding for referrals, where the duplicated count was lower than the unduplicated count in 73% of the sample.

Table 15

Differences in Composition of School Enrollment, Suspensions, and Students who Received Suspensions for Hispanic Students

School ID	Percent School Population	Percent of Suspensions, Duplicated Count	Difference from Population *(Duplicated)*	Percent of Students Receiving Referrals, Unduplicated Count	Difference from Population *(Undupl.)*
5038	3.4%	3.4%	-0.1%	3.7%	8.9%
2434	3.5%	3.4%	-1.5%	6.7%	90.5%
4061	4.2%	1.1%	-74.8%	3.6%	-15.0%
129709	4.6%	25.0%	443.5%	33.3%	624.6%
206	6.3%	20.0%	217.5%	25.0%	296.8%
1962	6.6%	6.7%	1.0%	10.5%	59.5%
177	6.7%	3.2%	-52.4%	4.5%	-32.2%
140488	7.1%	3.0%	-57.3%	5.6%	-21.8%
3491	7.3%	37.5%	413.7%	28.6%	291.4%
212	9.5%	25.0%	163.2%	30.0%	215.8%
4342	10.0%	6.3%	-36.5%	7.9%	-21.1%
4878	12.1%	2.3%	-80.8%	3.6%	-70.5%
186	12.3%	35.3%	186.9%	40.0%	225.2%
1883	15.4%	2.9%	-81.4%	7.1%	-53.6%
1313	17.6%	20.5%	16.2%	19.2%	9.3%
3052	18.7%	2.9%	-84.3%	5.4%	-71.1%
167670	20.6%	10.0%	-51.5%	14.3%	-30.7%
140486	22.0%	12.2%	-44.7%	14.0%	-36.4%
2006	22.2%	9.8%	-56.1%	14.3%	-35.6%
345	22.4%	10.7%	-52.2%	13.8%	-38.4%
183	24.4%	18.2%	-25.5%	12.5%	-48.8%

Impact of suspension counts on school-level interpretations (Hispanic students).

As with the analysis of suspensions for African American students, the finding of over- or under-representation was overwhelmingly consistent regardless of the type of count used. For Hispanic students, there were only two cases where the conclusions based on the different counts did not agree. Both times, the duplicated count resulted in a finding of slight under-representation (differences of -1.5% and -0.1%), while the unduplicated count resulted in a finding of over-representation (differences of 90.5% and 8.9%). So it seems that even though the finding of over-

representation will likely be the same regardless of the type of count used in the calculations, there is the potential for the type of count to influence the finding.

Differences in suspension patterns between African American and Hispanic students. Over-representation in suspensions was a reliable finding for African American students, occurring in almost all schools. Over-representation for Hispanic students occurred much less frequently in just 33%-43% of the schools, depending on the type of count. However, when over-representation in suspensions was found for Hispanic students, the intensity of the finding was relatively similar to the intensity of the finding for African American students. While there was variation in the maximum values for the range of differences between suspension events and proportion of school population (duplicated count), the average value for this difference was similarly intense: African American students on average experienced almost three times as many suspension events as would be expected, while Hispanic students on average experienced a little more than twice as many suspension events as would be expected. With the unduplicated count, both African American and Hispanic students were represented among those who received suspension at a rate that was more than twice what would be expected, on average.

As with the findings in the referral analysis, Hispanic students were found to be under-represented much more frequently than African American students, and often to a much greater extent. Over half of the sample (14 schools, or 54%) reflected under-representation in suspensions for Hispanic students using the duplicated count, while 46% of the sample (12 schools) had the same finding with the unduplicated count. On average, Hispanic students were found to have approximately 50% fewer suspensions than expected, and were represented among students who received suspension at a rate that was approximately 40% lower than expected. Compare this to findings of under-representation for African American students, which occurred in only three schools and had an average difference of 21% fewer suspensions than expected, and who were under-represented in the percent of students receiving suspension in only one school at a rate that was 18.7% lower than expected. Taken together, it becomes clear

that over-representation in suspensions is a much more pervasive problem for African American students than Hispanic students.

In general, the duplicated count produced a smaller value than the unduplicated count for both African American students (60% of schools) and Hispanic students (86% of schools). Compared to the finding for referrals, it was relatively less common for African American students to receive multiple suspensions as it was for this group to receive multiple referrals (the duplicated count for referrals was lower in just 32% of schools). While it was also true that Hispanic students were less likely to receive multiple suspensions than multiple referrals, the difference was much less: the duplicated count was lower for Hispanic students for suspension in 86% of schools, compared to 73% of schools for referrals. Between the two groups of students, results for Hispanic students were more consistent, with the duplicated count producing a smaller value than the unduplicated count in a clear majority of schools for both referrals and suspensions.

Changing from the duplicated to unduplicated count had little impact on school-level interpretations of disproportionality for both groups of students. For both African American and Hispanic students, there were only two cases in which the duplicated count produced a finding of under-representation and the unduplicated count produced a finding of over-representation. For better or worse, the unit of analysis for suspensions has less of an impact on interpretations of disproportionality than it does with referrals.

Suspensions by race/ethnicity: Students identified as White (duplicated count). The comparison of the percent of suspension events accounted for by White students (the duplicated count) to their percent of the school's population is presented in Figure 26. All 39 schools in the final sample reported suspensions for students who were White.

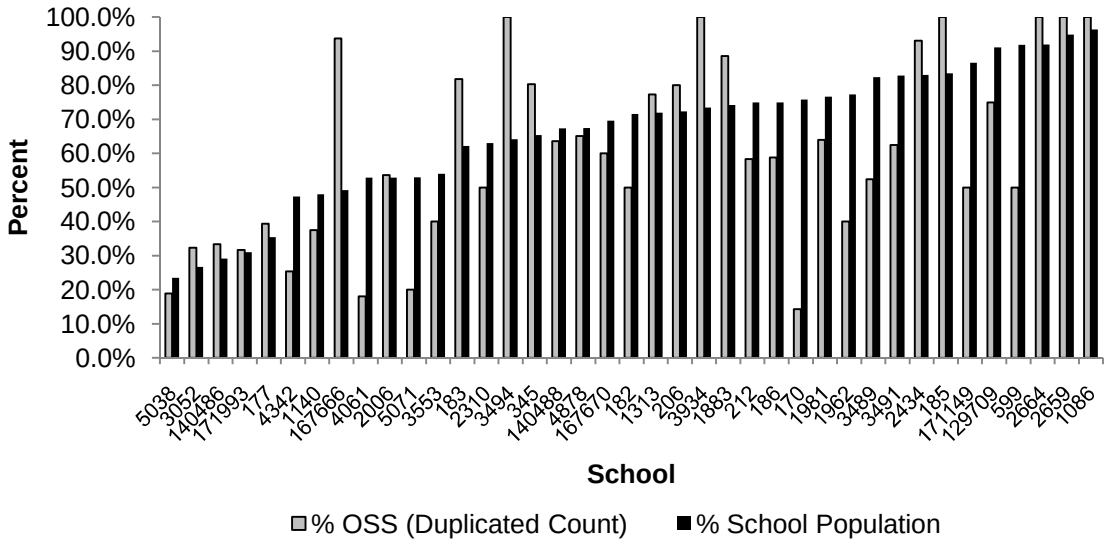

Figure 26. Percent of suspensions accounted for by White students compared to their proportion of the school's population.

Using the duplicated count, White students were over-represented in 18 of the 39 schools (46% of the final sample). The percentage of suspension events accounted for by White students ranged from 14.3% to 100%, while White students' percent of their school population ranged from 23.5% to 96.4%. The school-by-school differences between percent of suspensions accounted for by White students and their percent of the school population ranged from -81.2% to 90.5%, meaning that at the lowest end of the distribution White students accounted for approximately 80% fewer suspensions than would be expected, while at the highest end of the distribution White students accounted for just over 90% more suspensions than would be expected.

Students receiving referrals by race/ethnicity: Students identified as White (unduplicated count). Figure 27 compares the percent of students who received suspension who were identified as White (unduplicated count) to their percent of the school's population. With the unduplicated count, White students were over-represented in 16 of the 39 schools (41% of the sample). The percent of students who received suspension who were identified as White ranged from 14.1% to 100%, while White students percent of their school population ranged from 23.5% to 96.4%. The school-by-school differences between percent of students who received

suspension who were White and percent of school population ranged from -56% to 77.8%, meaning that at the lowest end of the distribution the percent of students who received referrals who were White was about 56% less than expected, while at the highest end of the distribution the percent of students who received referrals who were White was close to 80% higher than expected. Overall, with the duplicated count White students were a little more likely than Hispanic students to receive suspension, and with the unduplicated count White students were about as likely as Hispanic students to be represented among students who receive suspension. With both types of counts, African American students were over-represented in suspension more often than either White or Hispanic students.

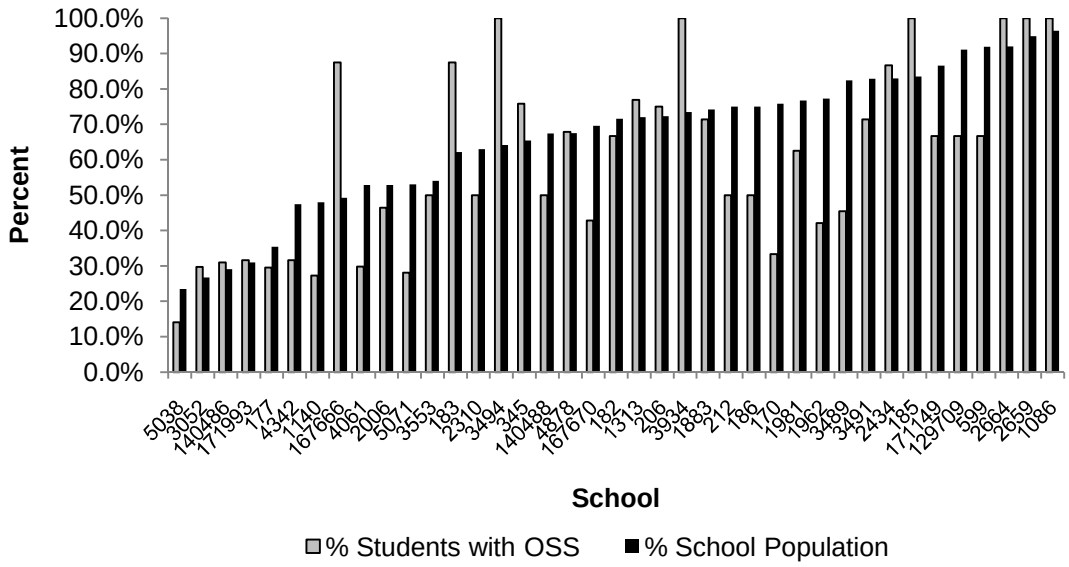

Figure 27. Percent of students who received suspension who were identified as White compared to their proportion of the school's population.

Magnitude of over-representation of White students. In the 18 schools where White students received more suspensions than would be expected (using the duplicated count), the average difference between the percent of suspensions accounted for and the percent of the student population was 20.8% (*SD* = 22.1), meaning that on average, White students accounted for about 20% more suspension events than would be expected given their proportion of the

student body. The range of over-representation in the duplicated count spanned from 1.4% to 90.5%.

In the 16 schools that revealed over-representation in the unduplicated count, White students were represented at a rate that was on average 18.7% higher than expected. The range of over-representation in the unduplicated count spanned from 0.5% to 77.8%.

Magnitude of under-representation of White students. There were 21 schools (54% of the sample) with under-representation in their duplicated count for White students, with differences between the percent of suspensions accounted for by White students and their percent of the school population that ranged from -3.5% to -81.2% (X = -32%, SD = 20.3). For the 23 schools (59% of the sample) with under-representation of White students in their unduplicated count, differences ranged from -3.7% to -56% (X = -28.8%, SD = 14.7). Overall, regardless of the type of count, under-representation in suspension for White students was found in a higher number of schools than was over-representation. In addition, the magnitude of over-representation for White students was considerably less intense than it was for both African American or Hispanic students, and the magnitude of under-representation fell in-between those found for African American and Hispanic students.

Differences between duplicated and unduplicated counts (White students). Table 16 lists the school-by-school percentages for White students' composition within a school building, their percent of suspensions, and their composition within the body of students at that school who received suspension. The duplicated count produced a lower value than the unduplicated count in 31% of the sample (12 schools), and there were seven schools in which the value for the duplicated and unduplicated counts were the same. With 51% of the schools in the sample reflecting higher values in their duplicated counts and 49% of the sample reflecting lower or equal values in their duplicated counts, there is no clear pattern for the likelihood of representation among "frequent flyer students" for White students.

Table 16

Differences in Composition of School Enrollment, Suspensions, and Students who Received Suspension for White students

School ID	Percent School Population	Percent of Suspensions, Duplicated Count	Difference from Population *(Duplicated)*	Percent of Students Receiving Suspension, Unduplicated Count	Difference from Population *(Undupl.)*
5038	23.5%	18.9%	-19.4%	14.1%	-40.1%
3052	26.7%	32.4%	21.2%	29.7%	11.3%
140486	29.1%	33.3%	14.5%	31.0%	6.5%
171993	31.0%	31.7%	2.2%	31.6%	1.9%
177	35.4%	39.4%	11.2%	29.5%	-16.5%
4342	47.4%	25.4%	-46.4%	31.6%	-33.4%
1140	48.0%	37.5%	-21.9%	27.3%	-43.2%
167666	49.2%	93.8%	90.5%	87.5%	77.8%
4061	52.9%	18.0%	-65.9%	29.8%	-43.7%
2006	52.9%	53.7%	1.4%	46.4%	-12.2%
5071	53.0%	20.0%	-62.3%	28.1%	-46.9%
3553	54.0%	40.0%	-25.9%	50.0%	-7.4%
183	62.2%	81.8%	31.5%	87.5%	40.7%
2310	63.0%	50.0%	-20.6%	50.0%	-20.6%
3494	64.2%	100.0%	55.8%	100.0%	55.8%
345	65.4%	80.4%	22.9%	75.9%	16.0%
140488	67.4%	63.6%	-5.6%	50.0%	-25.8%
4878	67.5%	65.1%	-3.5%	67.9%	0.5%
167670	69.6%	60.0%	-13.8%	42.9%	-38.4%
182	71.6%	50.0%	-30.2%	66.7%	-6.9%
1313	72.0%	77.3%	7.3%	76.9%	6.8%
206	72.3%	80.0%	10.7%	75.0%	3.7%
3934	73.5%	100.0%	36.1%	100.0%	36.1%
1883	74.2%	88.6%	19.4%	71.4%	-3.7%
212	75.0%	58.3%	-22.2%	50.0%	-33.3%
186	75.0%	58.8%	-21.6%	50.0%	-33.3%
170	75.8%	14.3%	-81.2%	33.3%	-56.0%
1981	76.7%	64.0%	-16.6%	62.5%	-18.5%
1962	77.3%	40.0%	-48.3%	42.1%	-45.5%
3489	82.4%	52.4%	-36.4%	45.5%	-44.8%
3491	82.9%	62.5%	-24.6%	71.4%	-13.8%
2434	83.0%	93.1%	12.2%	86.7%	4.4%
185	83.5%	100.0%	19.8%	100.0%	19.8%
171149	86.6%	50.0%	-42.3%	66.7%	-23.0%
129709	91.1%	75.0%	-17.7%	66.7%	-26.8%
599	91.9%	50.0%	-45.6%	66.7%	-27.5%
2664	92.0%	100.0%	8.7%	100.0%	8.7%
2659	94.9%	100.0%	5.4%	100.0%	5.4%
1086	96.4%	100.0%	3.7%	100.0%	3.7%

Impact of suspension counts on school-level interpretations (White students). In spite of the lack of a general pattern as to whether the duplicated or unduplicated count generated higher representation of White students in suspension, there were four instances where switching from the duplicated to unduplicated count resulted in a different school-level interpretation of disproportionality. In three of those instances, the duplicated count produced a finding of over-representation – which is the opposite of the findings for African American and Hispanic students, where for conflicting instances, the duplicated count produced a finding of under-representation.

Differences in suspension patterns between African American, Hispanic and White students. Over-representation in suspensions was generally found less often for White students than it was found for either African American or Hispanic students. However, if a duplicated count was used, Hispanic students were found to be over-represented in the smallest number (and percentage) of schools. When over-representation was found for White students, it was much less intense than it was for either of the other groups: White students' average percent difference from population (regardless of the type of count) was approximately 20%, while the average percent differences for African American students were well over 200%, and the average percent differences for Hispanic students were close to 200%.

Under-representation in suspension was generally more common for White students than it was for both African American and Hispanic students, occurring in 54% of the sample using the duplicated count and occurring in 59% of the sample using the unduplicated count. However, the same percentage of schools reported under-representation in suspension for Hispanic students (54% of the schools reporting suspension for Hispanic students) using the duplicated count.

The clearest pattern to emerge with respect to whether the type of suspension count produced higher or lower values for groups of students was found for Hispanic students. With this group of students, the duplicated count tended to produce lower values in the majority of schools in the sample, suggesting that Hispanic students were relatively less likely to receive multiple suspension events. African American students also had a higher number of schools (60%) where

the duplicated count was lower than the unduplicated count, but there was no clear majority of schools in which this was true for White students.

Changing from the duplicated to unduplicated count impacted school-level interpretations of disproportionality for White students more often than it did for either African American or Hispanic students, but this occurred on a very infrequent basis (four times across 29 schools). This suggests that much of the time the type of count used in the calculation may not be functionally important for determining disproportionality in suspension.

Overall, the findings suggest that over-representation in suspension can be an issue for African American, Hispanic, and White students, but is more common and more intense for African American students. When over-representation occurs for Hispanic students, it occurs at magnitudes that are similar to those found for African American students, but when over-representation occurs for White students, it tends to be relatively mild. Hispanic students tended to have lower representation in the number of suspension events and higher representation among students who receive suspension, suggesting that students of other races/ethnicities may be relatively more likely to experience multiple suspensions. Table 17 summarizes the differences in findings for each racial group.

Table 17

Summary of Differences in Suspension by Racial Group

	African American Students		Hispanic Students		White Students	
	Min.	Max.	Min.	Max.	Min.	Max.
Percent of Student Body	1.4%	70.8%	3.4%	24.4%	23.5%	96.4%
Number of Schools with Over-Representation based on the Duplicated Count	22 / 25		7 / 21		18 / 39	
Number of Schools with Over-Representation based on the Unduplicated Count	24 / 25		9 / 21		16 / 39	
Degree of over-representation in suspensions *(Difference between percent of suspensions accounted for and percent of student body)*	6.6%	1,548.4%	1%	443.5%	1.4%	90.5%
Degree of over-representation in students who received a suspension *(Difference between racial group composition in students who received a suspension and percent of student body)*	9%	920.4%	8.9%	624.6%	< 1%	77.8%
Degree of under-representation in suspensions	-1.7%	-57.9%	-0.1%	-84.3%	-3.5%	-81.2%
Degree of under-representation in students who received a suspension	-18.7%		-15%	-71%	-3.7%	-56%
Representation in duplicated count compared to representation in unduplicated count	Mixed; duplicated count resulted in lower values in 60% of the sample		Mostly lower		No pattern	
Impact of changing from duplicated count to unduplicated count on school-level interpretations of disproportionality	Mostly none *(2 schools)*		Mostly none *(2 schools)*		Little *(4 schools)*	

Suspension risk ratios: Students identified as African American. Turning now to the risk ratio analysis of suspensions, Figure 28 displays the risk ratios for African American students for suspension. All 39 schools (those which had racial group information on at least 90% of their office referrals) were included in the analysis, regardless of whether they reported any suspensions for African American students.

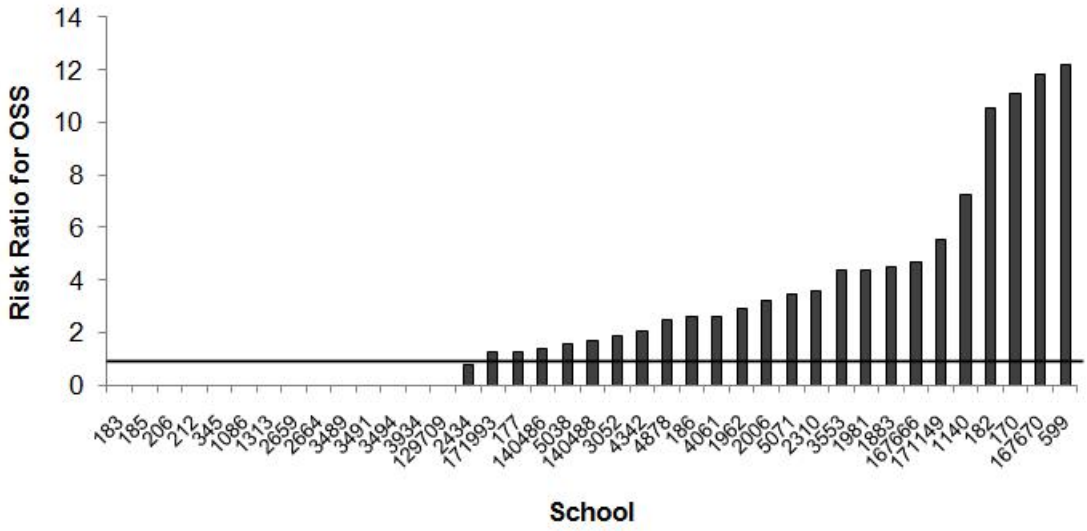

Figure 28. Risk ratios for African American student suspension.

The results for African American students' risk ratios for suspension found over-representation in 24 of the 39 schools (62% of the sample). At first glance, this seems to be a very different result from the comparison analysis for suspensions (unduplicated count), which found over-representation in 96% of the sample. However, 13 schools were included in the risk ratio analysis that weren't in the comparison analysis – these were schools excluded because they did not report any suspensions for African American students. With the comparison analysis, the value that would be graphed for schools which reported zero suspensions for African American students would only reflect the group's percent of the school population. While this is good information to offer, it could also be argued that the information visually clutters the graph. With the risk ratio, schools which report zero suspensions for African American students are represented as a zero on the X axis, and the amount of risk for all of the schools in the sample can be quickly assessed. Thus, another benefit of the risk ratio is that it may be easily calculated and compared across a large number of schools. Regardless, once the schools that did not report suspension for African American students are taken into account, the results of the risk ratio analysis for African American students' suspension were consistent with what was found in the comparison analysis (unduplicated count).

For the 24 schools that evidenced over-representation through the risk ratio, 21 schools (54% of the sample) showed African American students to have at least a 50% higher risk of receiving a suspension compared to all other students. In 18 schools (46% of the sample), African American students were more than *twice* as likely as students from all other racial groups to receive a suspension. There were four schools in which the risk ratio showed African American students to have more than 10 times the risk for suspension than any other racial group.

Taking a closer look at the schools with risk ratios of 10 or higher, it was discovered that each of these four schools had fewer than 10 students (across all racial/ethnic groups) experience suspension. Table 18 lists the breakdown for these schools. Having a low number of total students who received suspension, however, did not always lead to extreme risk ratios – 10 of the 18 schools that had fewer than 10 students experience suspension showed risk ratios of zero for African American students. In addition, it was discovered that schools with a total of 100 or more students who experienced suspension reflected risk ratios close to 1.5 (for one school with this criterion, the risk ratio was 1.4; for the other school with this criterion, the risk ratio was 1.58). So while it appears that the total number of students receiving a consequence can impact a school's risk ratio, the relationship is not absolute. A list of schools' risk ratios, number of African American students receiving suspension, and the magnitude of disproportionality found in the unduplicated and duplicated analyses can be found in Appendix G.

Table 18

Number of Students with Suspensions in Schools with Extreme Risk Ratios

Risk Ratio	Number of African American Students with OSS	African American Student Enrollment	Number of "All Other" Students with OSS	"All Other" Student Enrollment
12.20	1	15	2	366
11.83	1	7	6	497
11.12	2	34	1	189
10.54	2	9	7	332

Overall, the of risk for African American students to be over-represented in suspensions was less than their risk for being over-represented in referrals - only 62% of the sample reflected risk ratios of greater than 1.0 for suspension, compared to 79.5% of the sample for office referrals. However, risk ratios for suspensions were larger than those for referrals, with 46% of the sample (18 schools) reflecting risk ratios higher than 2.0 (including the schools with risk ratios higher than 10.0).

Suspension risk ratios: Students identified as Hispanic. Figure 29 displays risk ratios for students were identified as Hispanic. Information from all 39 schools in the final sample was included, regardless of whether a school reported any referrals for Hispanic students.

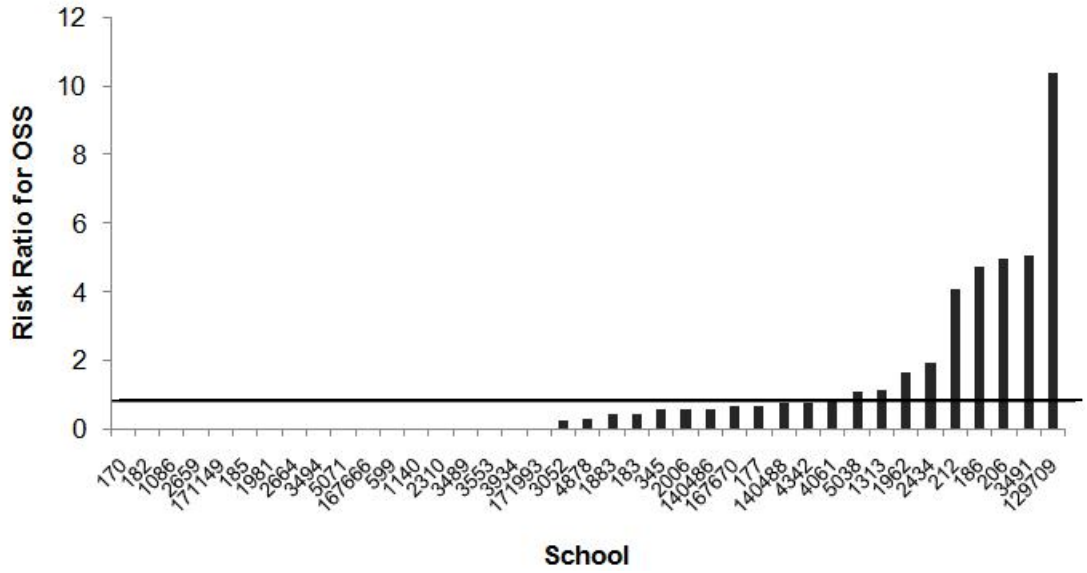

Figure 29. Risk ratios for Hispanic students' suspension.

The results for Hispanic students' risk ratios for suspension found over-representation in nine of the 39 schools (23% of the sample). In seven of the schools (18% of the sample), Hispanic students had at least a 50% higher risk of receiving a suspension than all other students, and in five schools (13% of the sample), the risk was more than twice as high as all other students. In general, the risk ratio analysis supported the results of the comparison analysis, except for occasional shifts in the rank-order of individual schools (e.g., school #3491).

A list of schools' risk ratios, number of Hispanic students receiving suspension, and the magnitude of disproportionality found in the unduplicated and duplicated analyses can be found in Appendix H.

Overall, the of risk for Hispanic students to be over-represented in suspensions was less than their risk for being over-represented in referrals (about half as many schools reflected risk ratios of greater than 1.0), but when over-representation was indicated, the degree was frequently more intense.

Comparison of African American students' and Hispanic students' risk ratios for suspension. When each school's risk ratios for suspension of African American and Hispanic students were plotted together, it was clear that African American students received suspension in more schools and experienced higher magnitudes of risk than Hispanic students. Figure 30 shows each school's risk ratios for African American and Hispanic students' suspension. There were half as many schools reflecting under-representation (including zero risk) for suspension of African American students as for Hispanic students (15 schools with risk ratios below 1.0, compared to 30 schools for Hispanic students). Conversely, there were more than three times as many schools reflecting high magnitudes of *over*-representation (risk ratios of 2.0 or greater) for suspension of African American students as for Hispanic students (18 schools, versus five schools for Hispanic students). Overall, the findings in the current study indicate that African American students frequently experience more widespread and intense problems with disproportionality in suspension than do Hispanic students.

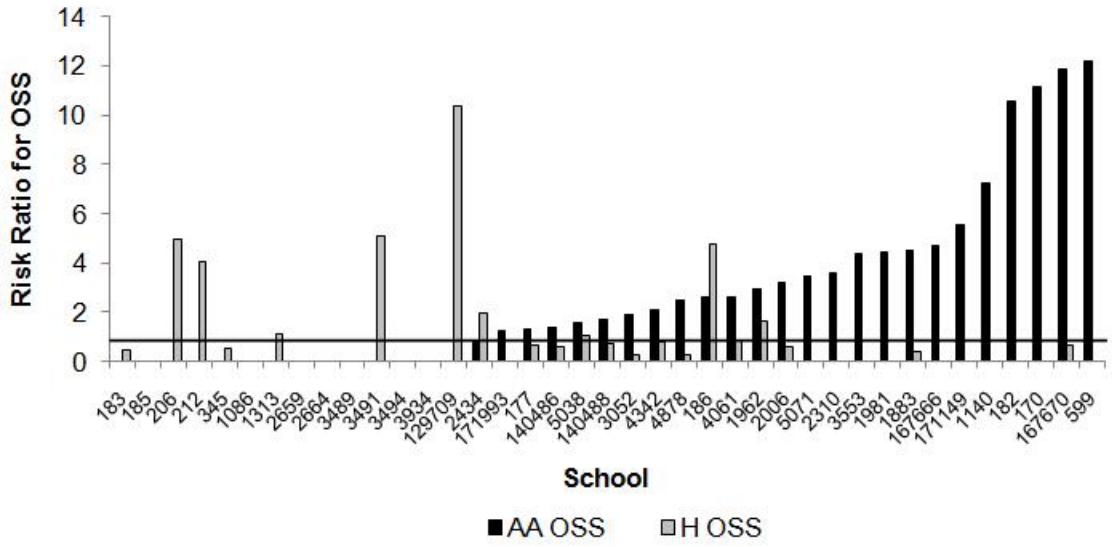

Figure 30. African American and Hispanic students' risk ratios for suspension.

Comparison of African American and Hispanic students' risk across referrals and suspensions. Table 19 provides a summary of the risk ratio findings for African American and Hispanic students for both office referrals and suspensions. The patterns of risk ratios for office referrals and suspensions were similar within each group of students, but the patterns differed for each racial group. For African American students, there were about an equal number of schools reflecting under-representation as there were reflecting higher magnitudes of *over*-representation for both office referrals and suspensions. For Hispanic students, there were many more schools which reflected under-representation than reflected higher magnitudes of over-representation for both office referrals and suspensions. Between the two groups of students, African American students experienced higher magnitudes of over-representation in more schools than Hispanic students, regardless of whether office referrals or suspensions were being examined. Overall, these findings suggest that over-representation in office referrals and suspension is more likely for African American students than Hispanic students, and more common for suspension than for office referrals.

Table 19

Summary of Risk Ratio Findings for Referrals and Suspensions

	African American ODR	Hispanic ODR	African American OSS	Hispanic OSS
Schools with Risk Ratios of <1	8	23	15	30
Schools with Risk Ratios of 2.0+	8	0	18	5

Identification of High- and Low-Implementing Schools

Level of school-wide positive behavior support (SW-PBS) implementation. The Benchmarks of Quality (BoQ) score was used to indicate the degree to which a school implemented SW-PBS with fidelity. All of the schools in the sample completed the BoQ during the spring or summer semesters of 2008, giving school PBS teams at least half of a school year to implement the critical elements that are measured by the tool. A copy of the BoQ Scoring Form and Scoring Guide can be retrieved from the national Technical Assistance Center on Positive Behavioral Interventions and Supports website: www.pbis.org. The critical elements measured by the BoQ include the PBS team, faculty commitment, discipline procedures, data entry and analysis, the school-wide expectations and rules, the lesson plans that are used to teach the expectations and rules, the reward/recognition program, the implementation and crisis plans, and evaluation (Cohen et al., 2007). The BoQ total score that was used in this analysis represents a cumulative total of the scores from all of the critical elements. For the 83 schools in the original sample, total BoQ scores ranged from 48 to 99, with a mean of 81 (*SD* = 11.2). Figure 31 provides a visual representation of the distribution of scores for the original sample.

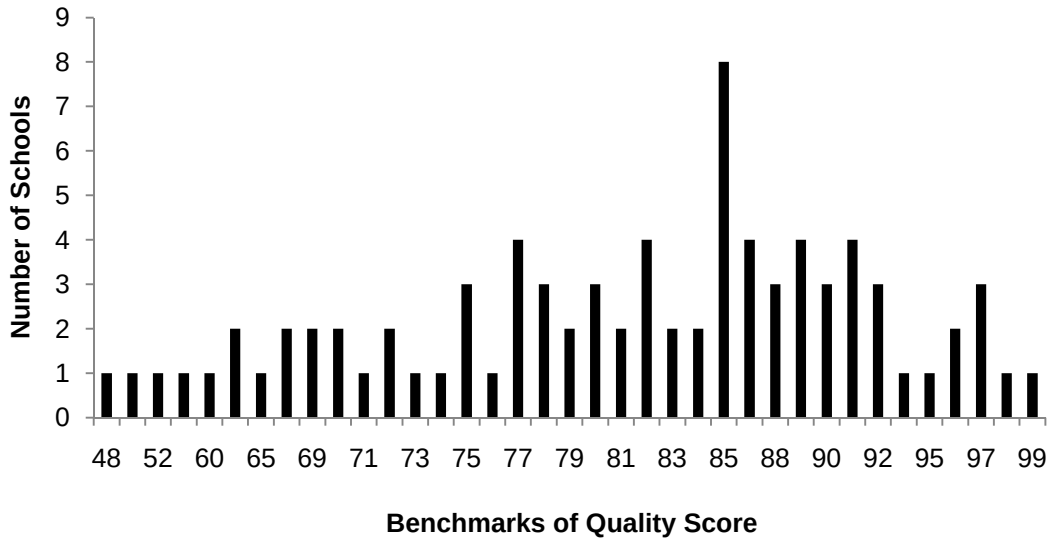

Figure 31. Distribution of schools' Benchmarks of Quality scores (*n* = 83).

The amount of negative skew in the distribution was significant (skewness = -.86, *SE* of skewness = .264), suggesting relatively few schools were implementing with low levels of fidelity. Unlike other school characteristics described thus far, the distribution of BoQ scores also had a prominent peak: eight schools shared modal scores (85), which was twice the number of schools that shared any other score in the Benchmarks' distribution (e.g., four schools shared a score of 77, four schools shared a score of 82, etc.). Just over 40% of the sample (35 schools) scored from 80 to 90. In spite of this peak, the concentration of scores around the mean was found to be relatively normal (kurtosis = .65, *SE* of kurtosis = .523). Taken together, although the total range of scores was spread out, many schools' scores were similar to each other and reflected higher levels of implementation, whereas relatively fewer schools had scores that reflected lower levels of implementation.

Implementation scores and risk ratios. To address the first research question, "What are the risk ratios for office discipline referrals and incidents of out-of-school suspensions for African American and Hispanic students in schools that implement SW-PBS," each school's Benchmarks of Quality score was plotted against their risk ratio as a way to facilitate visual analysis of a possible relationship between implementation and disproportionality. To help ensure

their risk ratios were relatively accurate representations of the disciplinary actions within a school building, schools were filtered to exclude those that were missing more than 10% of their referral-level race/ethnicity data, resulting in a sample of 39 schools. Figure 32 reflects schools' BoQ scores compared to their risk ratio for referrals for African American students. It should be noted that the graph truncates the horizontal axis to allow for a closer examination of the variables.

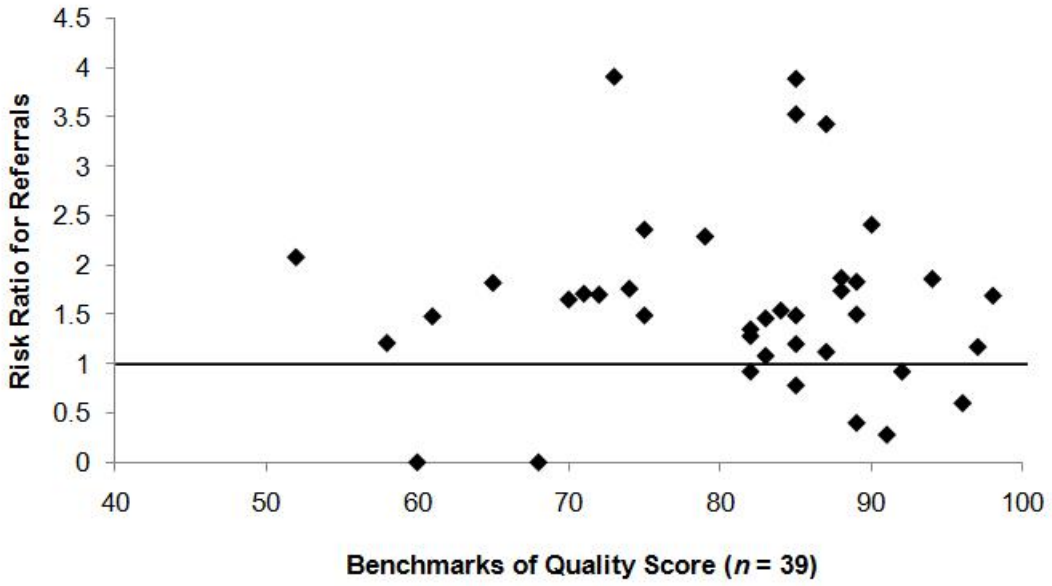

Figure 32. Distribution of BoQ scores by risk ratios for referrals for African American students.

Schools with risk ratios over 2.0 had BoQ scores which ranged across almost the entire distribution (from 52 to 90). Looking at under-representation, there were two schools with risk ratios of zero who's BoQ scores were in the low 60's, but six schools with risk ratios of less than 1.0 and BoQ scores in the high 80's-90's. Overall, while there was a slight pattern for schools with low risk ratios to have a higher BoQ score, there was no obvious pattern for schools with high risk ratios.

Figure 33 presents the BoQ scores and risk ratios for referrals for Hispanic students. There appeared to be a slight inverse relationship between BoQ score and risk ratio for referrals for Hispanic students. With the exception of one school which had a BoQ score of 97 and a risk

ratio of 1.79, risk ratios fell under 1.4 for every school with a Benchmarks score higher than 80. In addition, there were more schools with risk ratios of less than 1.0 which had BoQ scores higher than 70 than there were at the lower end of the Benchmarks distribution.

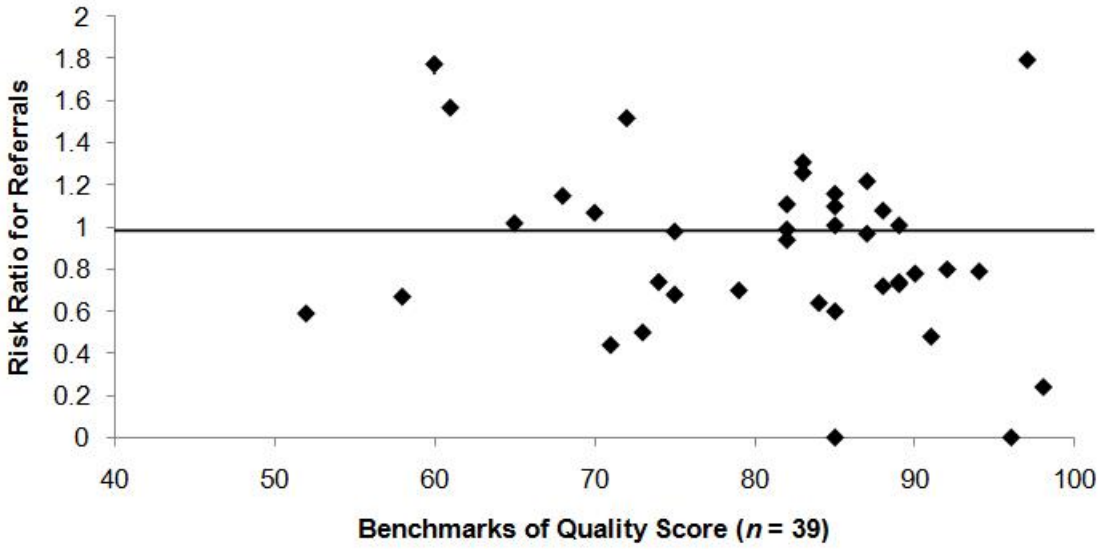

Figure 33. Distribution of BoQ scores by risk ratios for referrals Hispanic students.

Figure 34 depicts BoQ scores and risk ratios for suspension for African American students. While there were more schools with risk ratios of zero when Benchmarks scores exceeded 80, there were also more schools with higher risk ratios in this range (risk ratios of 3.0 and greater) – even after excluding the schools that had extreme risk ratios (values greater than 10). Overall, the distribution was flat, and did not support a strong relationship.

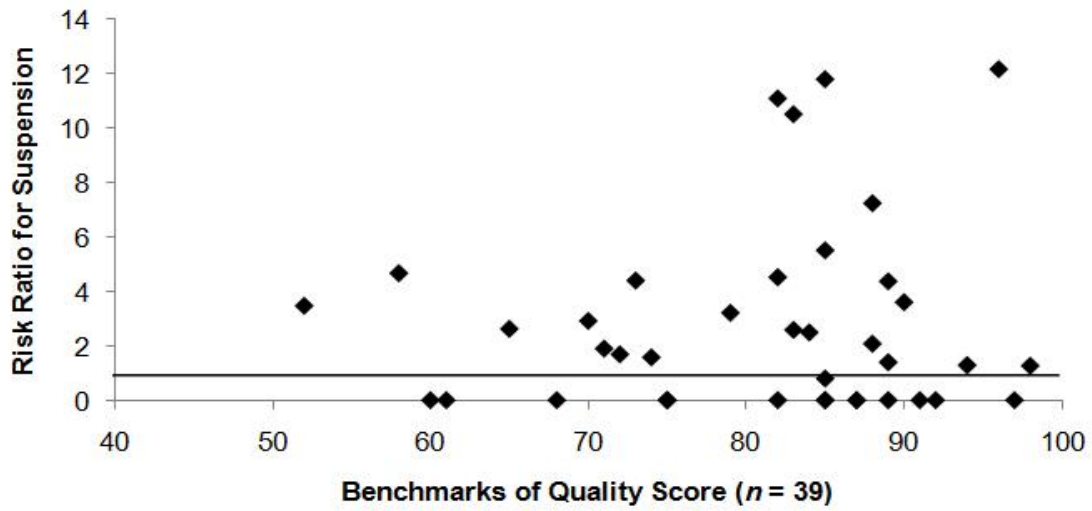

Figure 34. Distribution of BoQ scores by risk ratios for suspension for African American students.

The distribution of risk ratios for suspension for Hispanic students and Benchmarks scores is presented in Figure 35. Compared to the other scatter plots, the distribution for Hispanic students' suspension was extremely flat – only five schools had risk ratios higher than 2.0. Given the extreme lack of variation, the visual analysis did not support a relationship between implementation and disproportionality in suspension for Hispanic students.

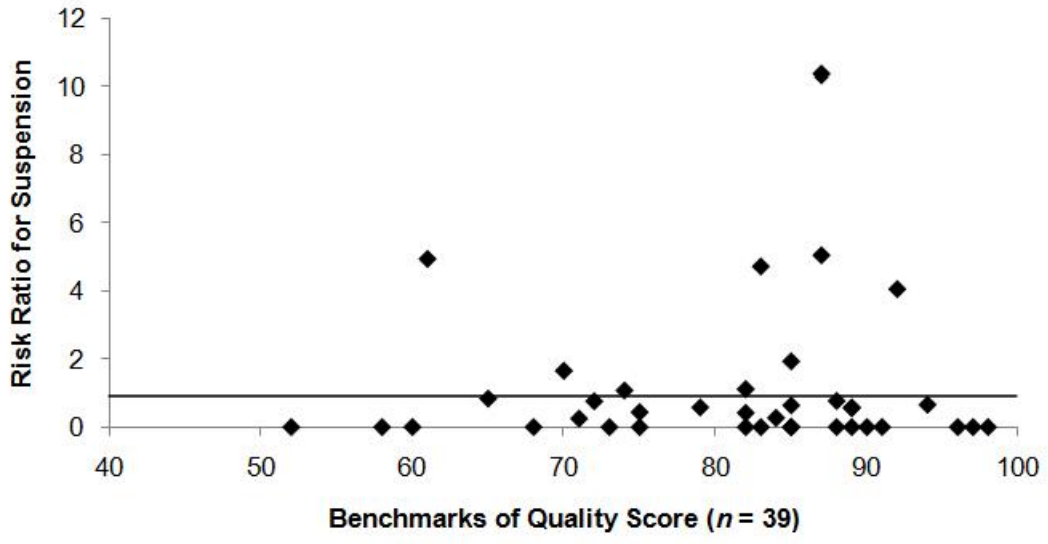

Figure 35. Distribution of BoQ scores by risk ratios for suspension for Hispanic students.

Cut scores for high- and low-implementing schools. The second research question, "Is there a relationship between the level of implementation of SW-PBS and levels of disproportionality in office discipline referrals and suspensions," requires the sample to be broken down into categories of high- and low-implementing schools. In an attempt to identify a socially valid way of differentiating schools that were implementing with higher and lower levels of fidelity, BoQ scores for the *original* sample of 83 schools were examined for potential differences across the distribution. With 42% of the elementary school sample scoring from 80 to 90, it seemed that these scores might have potential for identifying practical differences in implementation levels. Given that so many schools obtained scores within the same range, these schools likely reflected similar levels of fidelity. Using a score of 91 as the cut score for high implementing schools and 79 as the cut score for low implementing schools would offer a degree of separation that might reflect practical differences at the school level. Unfortunately, using those values would result in a total of 32 low-implementing schools and 16 high-implementing schools, and with such a small overall sample size, could increase the chance of a Type 1 error (finding a difference when none actually exists).

However, including schools that scored *higher than* the mean (81), median (83), and mode (85) in the high-implementing category would increase the balance in the number of schools categorized as high- and low-implementing, while still excluding two-thirds of the schools with scores in the common range. This resulted in 30 high-implementing schools with scores from 87-99, and 32 low-implementing schools with scores from 48-79.

Differences in outcomes for low- and high-implementing schools. To examine whether this categorization reflected actual differences in school outcomes, office discipline referral (ODR) rates for high- and low-implementing schools were examined. In order to maximize the likelihood of finding differences between the two groups of schools, the duplicated count was used as the basis for the referral rate. Figure 36 shows each of the high- and low-implementing school's BoQ score plotted against their rate of office discipline referrals per student. The shaded area on the graph reflects the "middle implementing" schools that were dropped from the analysis.

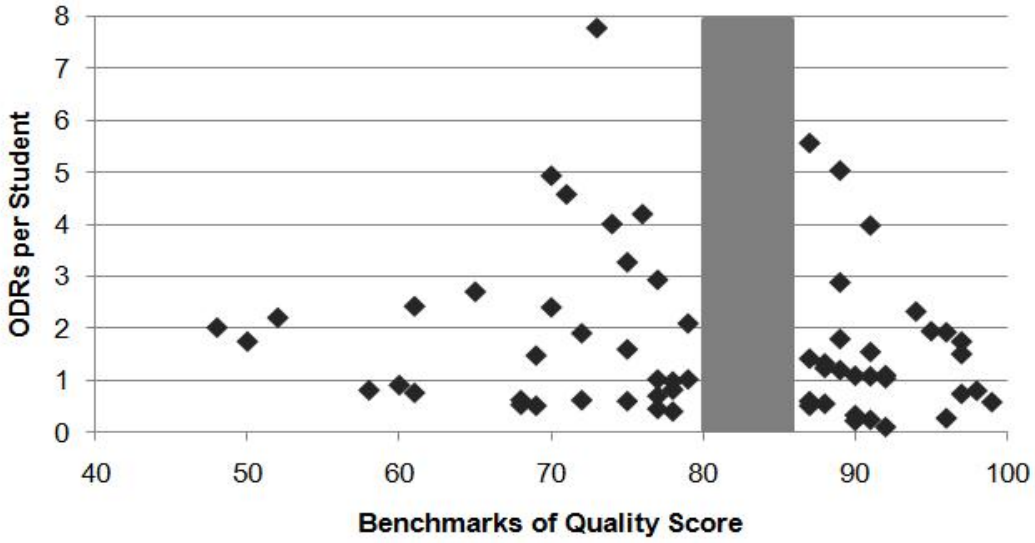

Figure 36. Office Discipline referral (ODR) rates for high- and low-implementing schools.

There were only six data points that did not overlap between the high- and low-implementing schools. There were five high-implementing schools with ODR rates that were lower than any found for the low-implementing schools, and one low-implementing school with an ODR rate that was higher than any found for the high-implementing schools. These few schools were enough to impact the average ODR rates, which for low-implementing schools was 1.96 referrals per student, and for high-implementing schools was 1.48 referrals per student (a 25% difference).

Figure 37 plots out-of-school suspension rates against schools' BoQ score to examine potential differences in this outcome measure. The shaded area on the graph reflects the "middle implementing" schools that were dropped from the analysis. The OSS rate was based on the duplicated count of suspensions in order to maximize the likelihood of finding differences between the groups of schools. As with the plot of ODR rates, there was a high degree of overlap between high- and low-implementing schools: there were no high-implementing schools with OSS rates that were lower than any found for the low-implementing schools, and only three low-implementing schools which had OSS rates that were higher than any found for the high-

implementing schools. The average OSS rate for low-implementing schools was 0.13 OSS per student, and for high-implementing schools the average rate was 0.07 OSS per student (a 46% difference). Overall, there was not a dramatic difference in the outcome measures between high- and low-implementing schools, although on average, higher implementing schools had slightly better outcomes than the lower-implementing schools.

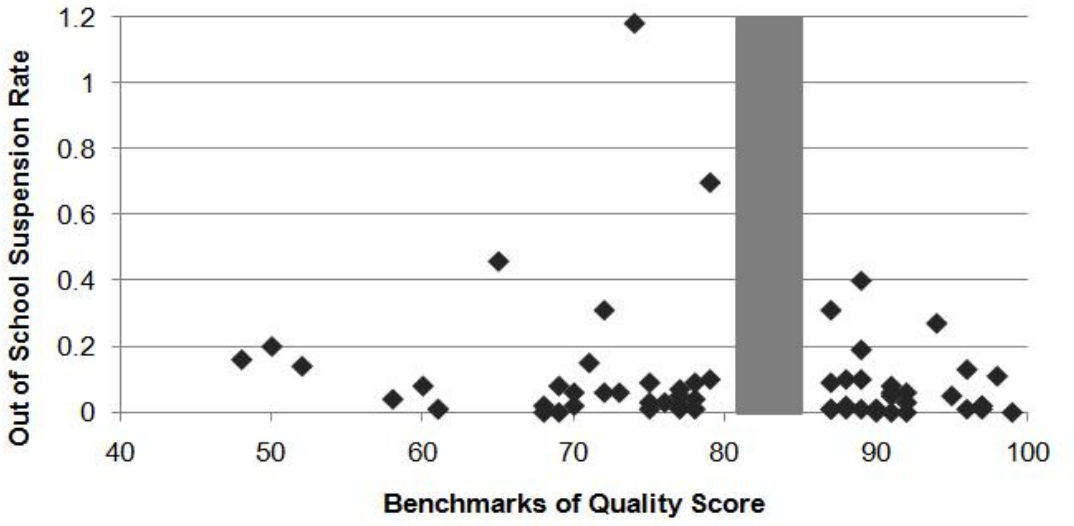

Figure 37. Out-of-School suspension rates for high- and low-implementing schools.

Implementation levels and missing data. Before moving on to explore differences in disproportionality by implementation level, the high- and low-implementing schools first had to be screened to exclude schools that were missing more than 10% of their referral-level race/ethnicity data. In all, 34 of the 62 high- and low-implementing schools had to be dropped (18 low implementing schools and 16 high implementing schools), leaving 14 high- and 14 low-implementing schools for the final analysis.

Given that so many schools had to be excluded from analysis (44 of the original 83 schools), it seemed worthwhile to examine whether the number of excluded schools varied by implementation level. Figure 38 shows the percentage of excluded schools by implementation level. Overall, there was very little difference between low- and high-implementing schools in

terms of the percent of schools that were missing more than 10% of their race/ethnicity data. Interestingly, schools in the “middle implementing” category had the lowest percent of schools that were missing data (just under half of the middle implementing schools). The lower-implementing schools had the highest percent of schools that were missing data (56%, or 18 of the 32 schools).

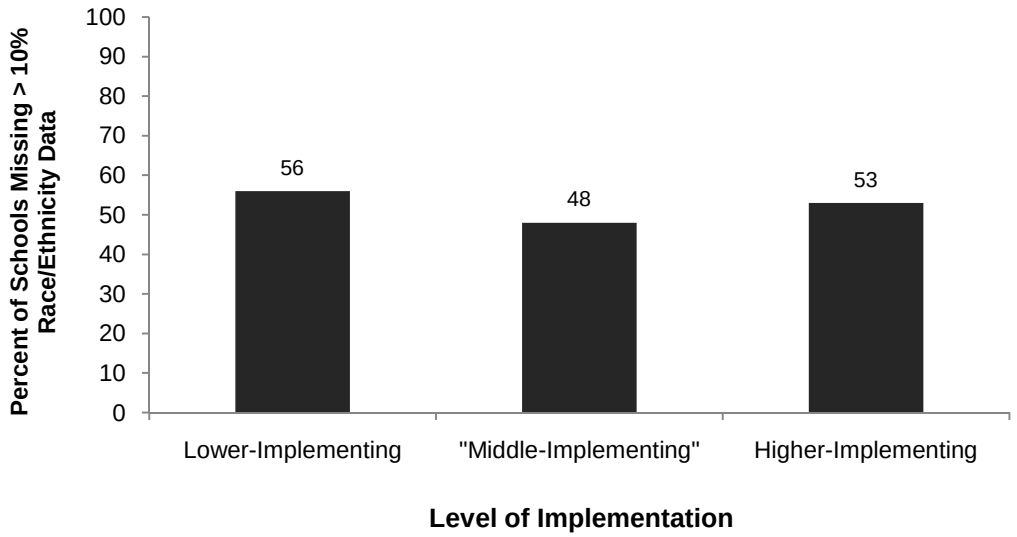

Figure 38. Percent of excluded schools by implementation level (*n*=83).

Disproportionality in referrals by implementation level: Students identified as African American. Based on the earlier discussion of the advantages of the risk ratio and it’s consistency with the comparison analysis, the exploration of disproportionality by implementation level will be limited to findings based on the risk ratio. Figure 39 lists the risk ratios for referrals for African American students by implementation level. Note that the separate columns do ***not*** represent the same school, but rather the implementation levels (high and low) of separate schools.

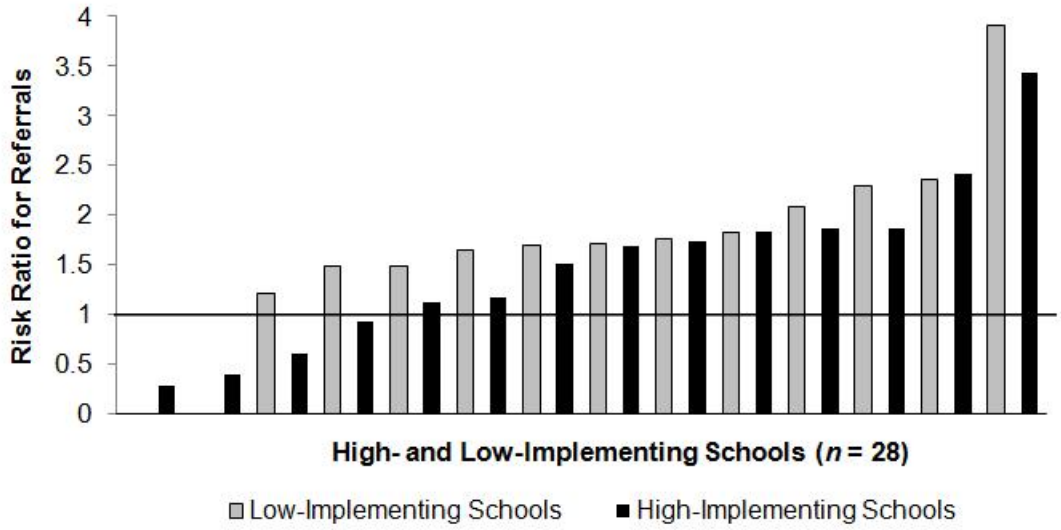

Figure 39. Risk ratios for referrals for African American students in high- and low-implementing schools.

Overall, there were small differences between high- and low-implementing schools for African American students' risk ratios for office discipline referrals, with high-implementing schools reflecting slightly better findings. Table 20 summarizes the distribution of risk ratios for referrals for African American students. The range of risk ratios for low-implementing schools spanned from zero to 3.91 (X = 1.68, SD = .96), while the range of risk ratios for high-implementing schools was slightly smaller, beginning with a risk ratio of .28 and ending with a risk ratio of 3.43 (X = 1.45, SD = .84). Higher-implementing schools had twice the number of schools with risk ratios of less than 1.0 (four high-implementing schools versus two low-implementing schools), and half as many schools with risk ratios greater than 2.0 (two high-implementing schools versus four low-implementing schools). While the small number of schools involved caution against broader interpretations, there nonetheless appears to be some tendency for high-implementing schools to have less of a problem with over-representation of African American students in office referrals.

Table 20

Distribution of risk ratios for office discipline referrals for African American students in high- and low-implementing schools

	Low-Implementing Schools (*n* = 14)	High-Implementing Schools (*n* = 14)
Min	0	.28
Max	3.91	3.43
Mean	1.68	1.45
SD	.96	.84
Median	1.71	1.60
Mode	0	---
Skewness	.22	.65
Kurtosis	2.12	.95

Note. A mode could not be calculated for high-implementing schools because every school had a different value on this measure.

Disproportionality in referrals by implementation level: Students identified as Hispanic. Figure 40 lists the risk ratios for referrals for Hispanic students by implementation level, while Table 21 provides a summary of the distribution. Few differences emerged for Hispanic students' risk ratios for referrals. At first glance, it appeared that low-implementing schools might have had slightly better outcomes than high-implementing schools. The range of risk ratios for low-implementing schools was slightly smaller than the range for high-implementing schools, beginning with a value of .44 and ending with a value of 1.77 (versus a range of 0 to 1.79 for high-implementing schools). In spite of the smaller range, however, low-implementing schools had a higher average risk ratio (X = .96) than high-implementing schools (X = .81), suggesting that risk ratios for lower-implementing schools tended to be higher overall. In addition, there were slightly more schools with risk ratios of less than 1.0 for high-implementing schools (10 high-implementing schools versus eight low-implementing schools). Neither group of schools had risk ratios greater than 2.0 for Hispanic students' office referrals. In all, there was little difference between high and low implementing schools for Hispanic students' risk ratios for referrals, but on average the high-implementing schools may have performed slightly better.

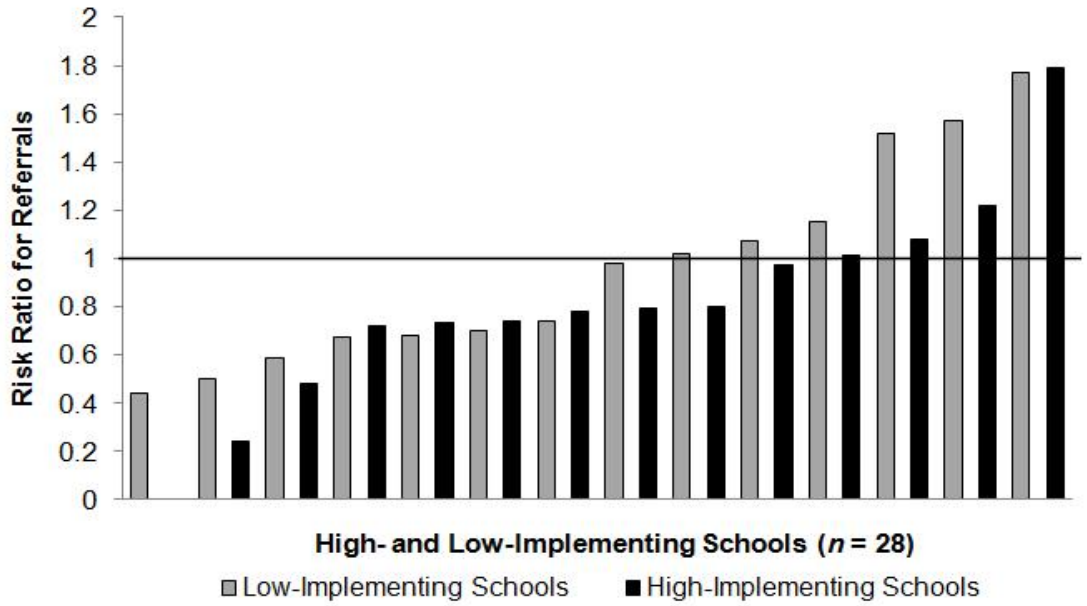

Figure 40. Risk ratios for referrals for Hispanic students in high- and low-implementing schools.

Table 21

Distribution of risk ratios for office discipline referrals for Hispanic students in high- and low-implementing schools

	Low-Implementing Schools (n = 14)	High-Implementing Schools (n = 14)
Min	.44	0
Max	1.77	1.79
Mean	.96	.81
SD	.42	.43
Median	.86	.79
Mode	---	---
Skewness	.73	.33
Kurtosis	-.57	1.68

Note. A mode could not be calculated because every school had a different value on this measure.

Disproportionality in suspensions by implementation level: Students identified as African American. Figure 41 lists the risk ratios for suspension for African American students by implementation level, while Table 22 provides a summary of the distribution. Overall, it appeared that high-implementing schools performed worse than low-implementing schools, with a larger range and higher average risk ratio than the low-implementing schools. However, the lower median value and large standard deviation for the high-implementing schools provide some insight as to why this may be the case: with a maximum risk ratio of 12.2, one of the high-implementing schools was an outlier school with a very small number of total students who had a suspension. At this particular school, only two students received suspension, one of whom was identified as African American; there were a total of 15 African American students in the school. Even though the risk ratio for this school is extremely high, it may not be an accurate depiction of disproportionality *per se*. Still, it's inclusion in the distribution of high-implementing schools has an impact on the resulting interpretations. In the end, given the similar number of high- and low-implementing schools with under-representation, and the similar number of high- and low-implementing schools with high over-representation (risk ratio greater than 2.0), it appears that level of implementation does not have a large effect on levels of disproportionality in suspensions for African American students.

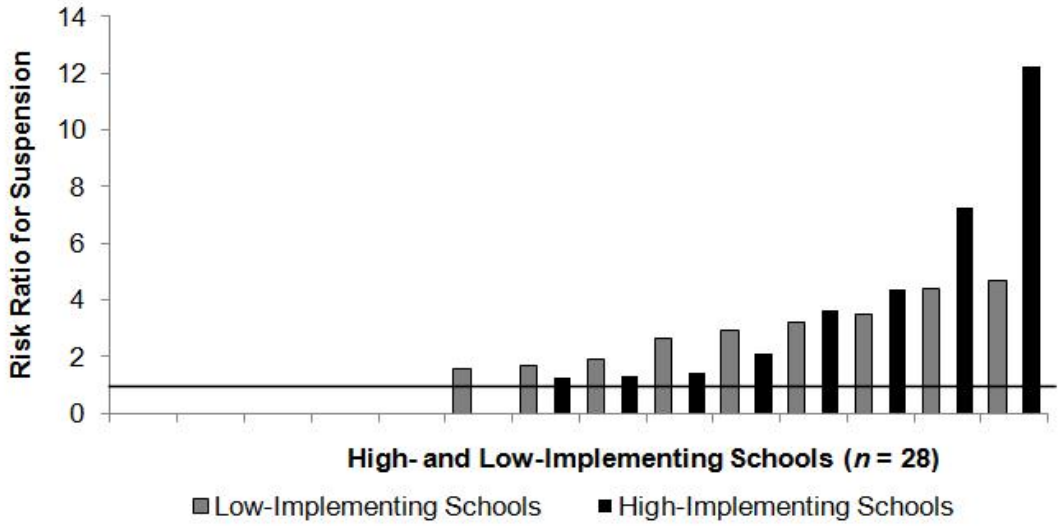

Figure 41. Risk ratios for suspension for African American students in high- and low-implementing schools.

Table 22

Distribution of risk ratios for suspension for African American students in high- and low-implementing schools

	Low-Implementing Schools (n = 14)	High-Implementing Schools (n = 14)
Min	0	0
Max	4.68	12.20
Mean	1.89	2.39
SD	1.71	3.54
Median	1.80	1.28
Mode	0	0
Skewness	.20	1.98
Kurtosis	-1.33	3.95

Disproportionality in suspensions by implementation level: Students identified as Hispanic. Figure 42 lists the risk ratios for suspension for Hispanic students by implementation level, while Table 23 provides a summary of the distribution. Similar to the findings for suspension for African American students, it appeared that high-implementing schools had worse outcomes

in suspension for Hispanic students than did low-implementing schools. High-implementing schools had a larger range of values and a larger average risk ratio, and also had more schools with high levels of over-representation (two schools with risk ratios greater than 2.0, compared to just one school in the low-implementing category). The high-implementing category also included an outlier school with a risk ratio over 10, and only two students within that school who received suspension (there were a total of 16 Hispanic students enrolled in the school). Given the presence of outliers and the small number of schools involved in the analysis, it is difficult to determine the role implementation fidelity plays in disproportional suspension outcomes.

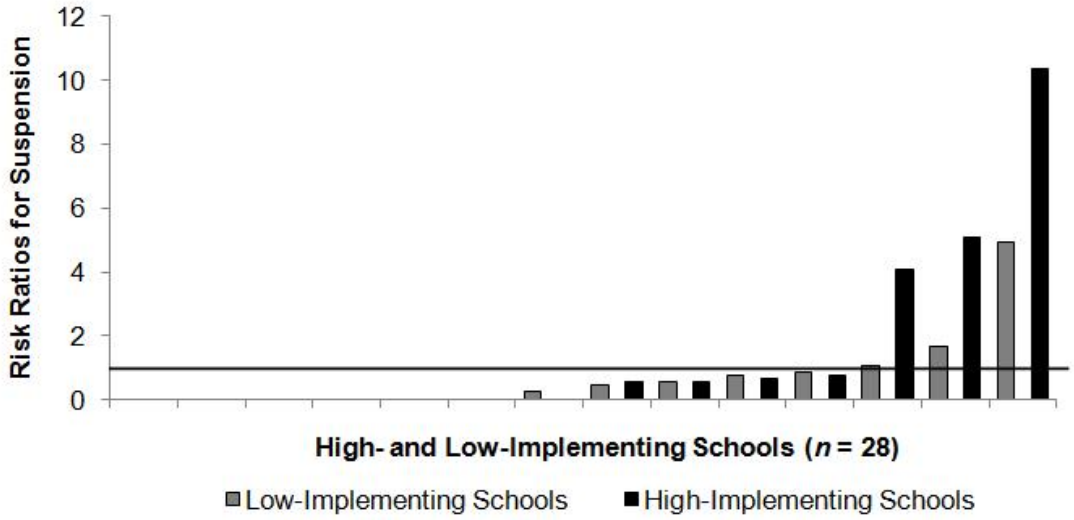

Figure 42. Risk ratios for suspension for Hispanic students in high- and low-implementing schools.

Table 23

Distribution of Risk Ratios for Suspension for Hispanic Students in High- and Low-Implementing Schools

	Low-Implementing Schools (n = 14)	High-Implementing Schools (n = 14)
Min	0	0
Max	4.96	10.38
Mean	.76	1.58
SD	1.31	3.0
Median	.35	.28
Mode	0	0
Skewness	2.87	2.36
Kurtosis	9.13	5.61

Chi-Square Analysis

Identification of disproportionality levels. Overall, it appeared that higher-implementing schools may have had slightly better outcomes with regards to risk ratios for referrals for African American and Hispanic students, but they may have had worse outcomes with regards to suspension for these groups. To help gauge the validity of these patterns, a Chi-Square test of Independence was performed to statistically evaluate the differences between the groups. The Chi-Square analysis tests the relationship between two or more categorical variables to see how the observed differences compare to chance. Chi-Square tests do not require a normal distribution, but they do require independent observations (such that each data point may only be counted only once), categorical data, and expected frequencies that are greater than five per cell (Field, 2009). In the current study, risk ratios were grouped into three categories to reflect the relative intensity of disproportionality reflected in the risk ratio measure.

The categories for the levels of disproportionality were based on a review of state definitions of "significant disproportionality" in special education placements (Burdette, 2007) and advocate guidelines. The review of state guidelines revealed a great deal of variation in how different states defined “significant” disproportionality; of the definitions that relied on a risk ratio,

values ranged from 2.0 to 3.0. Advocate guidelines (Kozleski, 2005) recommended that any risk ratio higher than 1.2 or higher should signal a school, district, or state to take action. Given that advocates' guidelines set 1.2 as the minimum risk ratio value that warrants action, the current study categorized any school with risk ratios below this value as "Low Disproportionality" schools for each of the different measures.

As state definitions of "significant disproportionality" reflected 2.0 as the minimum risk ratio value required for state intervention, the current study categorized any school with risk ratios of this value and higher as "High Disproportionality" schools. Schools with risk ratios that fell below 2.0 but above 1.19 were categorized as "Moderate Disproportionality" schools. Definitions for these categorizations are provided in Table 24.

Table 24

Definitions for Categories of Disproportionality

Category	Definition
Low Disproportionality	Schools with risk ratios below 1.20 for either African American or Hispanic students for referrals or suspension.
Moderate Disproportionality	Schools with risk ratios between 1.20 and 1.99 for either African American or Hispanic students for referrals or suspension.
High Disproportionality	Schools with risk ratios of 2.0 or higher for either African American or Hispanic students for referrals or suspension.

Tables 25, 26, 27, and 28 summarize the categorization of schools for each of the chi-square analyses. As seen in these tables, the small sample size (n = 28), would have violated a main assumption of the chi-square in each of the analyses: that each cell contain at least five observations. Fisher's Exact test provides an alternate method for calculating the chi-square that corrects for small cell values (Field, 2009). Therefore, all results in the analyses reflect the results of the Fisher's Exact test, for which only a p value is reported (i.e., Fisher's Exact does not calculate a formal test statistic or critical value).

Table 25

Classification Table for African American Students' Office Referrals

Level of Disproportionality	Level of Implementation	
	Low	*High*
Low	2	7
Moderate	8	5
High	4	2

Table 26

Classification Table for Hispanic Students' Office Referrals

Level of Disproportionality	Level of Implementation	
	Low	*High*
Low	11	12
Moderate	3	2
High	0	0

Table 27

Classification Table for African American Students' Suspension

Level of Disproportionality	Level of Implementation	
	Low	*High*
Low	5	6
Moderate	3	3
High	6	5

Table 28

Classification Table for Hispanic Students' Suspension

Level of Disproportionality	Level of Implementation	
	Low	*High*
Low	12	11
Moderate	1	0
High	1	3

Statistical results. Although the descriptive analysis and classification table suggested the potential for a relationship between levels of disproportionality in office referrals for African American students and implementation level, the relationship was not statistically significant (P = 0.1493, *FET*). Likewise, the results of the other statistical tests were not significant (disproportionality in office referrals for Hispanic students P = 1.00, *FET*; disproportionality in suspensions for African American students P = 1.00, *FET*; disproportionality in suspensions for Hispanic students P = 0.5956, *FET*). Based on these results, it seems that the level of implementation of SW-PBS is not related to schools' level of disproportionality in office referrals or suspensions for either African American or Hispanic students.

Chapter Five: Discussion

The current study set out to investigate the overall distribution of risk ratios for office referrals and suspensions for African American and Hispanic students in elementary schools which implement SW-PBS, and to determine whether schools which implement SW-PBS with higher levels of fidelity tended to have lower levels of disproportionality in these outcome measures. This chapter summarizes the noteworthy findings of the descriptive and statistical analyses, identifies major limitations, and relates these findings to considerations for future research and practice.

Noteworthy findings

Interpreting duplicated and unduplicated counts. As part of the descriptive analysis leading up to the research questions, schools' overall referral and suspension rates were calculated in order to get a sense of the range of referral-writing behaviors and suspension events within each of the buildings. Office referrals and suspensions have been shown to be useful indicators of school climate, and can provide insight into educators' general capacity for addressing problem behavior (Christle et al., 2004; Irvin, Tobin, Sprague, Sugai, & Vincent, 2004). Referral (or suspension) rates may be calculated based on the total number of referrals/suspensions written, so that if one student generates multiple incidents, all of those incidents are counted (the duplicated count). Referral/suspension rates may also be calculated based on the *number of students* who receive referrals/suspensions, so that if one student generates multiple incidents, only one is counted (the unduplicated count). Each method of calculation offers useful information, and can result in noticeably different outcomes. The duplicated count provides an indication of how frequently referrals/suspensions were written at a school, and offers a sense of the representation of particular groups of students among those who receive multiple incidents, as was done for gender and race/ethnicity in the current study.

The unduplicated count reflects the number of students who received at least one referral (or suspension), and removes the effect "frequent flyer" students may have on the school's overall rate.

Differences between the counts: Overall referral and suspension rates. In the current study, vast differences were found in overall referral rates depending on the method of calculation. Once the effect of students who receive multiple referrals was removed, schools in this study were fairly consistent in the rate in which they wrote referrals, with a standard deviation of just one referral per 6-7 students, compared to a standard deviation of 1.45 referrals per student under the duplicated count.

The overall range of suspension rates was small for both counts, reaching no more than just over one OSS per student under the duplicated count of suspension. The different counts for suspension events yielded similar findings, but with substantially fewer suspension events than referrals, differences between the counts weren't as dramatic. A number of schools in the current study (nine) had suspension *rates* of zero, and of these schools the majority reported zero suspension *events* as well.

Current referral rate vs. national referral rate. Overall, the rate at which most schools in this study issued referrals appeared to be slightly higher than other PBS schools. National averages of referrals, which were also based on the SWIS database, indicated 0.4 referrals per day per 100 students (Spaulding et al., 2010). Of course, this calculation can't be directly compared to those used in the current analysis – the standardization "per day per 100 students" is decidedly different than "referrals per student." The metric used by the SWIS database accounts for differences in the total number of school days, which varies from state to state. This consideration was not accounted for in the metrics used in the current study, which has the potential to affect the findings: a 2004 national review of the number of instructional days and hours conducted by the Education Commission of the States revealed that the minimum number of instructional days "varies widely" from state to state, and while the majority of states (30) require 180 days of instruction, limitations on the start and end dates of the school calendar and local conditions (such as snow days, days lost due to storms, etc.) may further impact the total

number of instructional hours experienced by students (D. Kincaid, personal communication, May 29, 2010; Tomlinson, 2004). However, the calculation of referrals per student used in the current study was not intended to be used as a literal measure, but only as an estimation of referral-writing and suspension characteristics of the schools in the sample.

Even though the method of calculation differed, the rates of referrals may still be compared – although generally – to national rates in the SWIS database. The website for the OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports provides national summaries of referral data from SWIS, which include the total number of schools, the total number of students, and the total numbers of referrals. Based on this information, the number of referrals per student for all K-12 schools using SWIS was calculated to be 0.90, which is lower than the average found in current study (1.65 referrals per student for the duplicated count). While it could be argued that the schools with referral rates far from the mean were outliers and increased the mean value for the sample, it is also true that the majority of schools had duplicated referral rates of 1.0 or greater. Given the impact referrals can have on academic engaged time, school climate, and student-teacher relationships, the increased rate in this sample is worth noting.

Boys' over-representation in disciplinary outcomes. The descriptive analysis was broadened to encompass another group for whom over-representation has been a highly consistent finding, and in doing so provided another example of boys receiving far more disciplinary consequences than would be expected. In the current study, boys received an average of four times more referrals than did girls, and on average accounted for a little more than 70% of the students who received referrals, even though boys typically make up only about 50% of a school's population. Caution should be used in interpreting this last finding however, as school-level enrollment by gender was not available for this sample. Without knowing for sure that boys and girls each accounted for approximately 50% of the population, judgments of over-representation should be provisional. Findings for suspension were similar, with boys accounting for an average of five times more events, and accounting for more than 80% of the students who receive the events. With suspension, the number of schools reflecting over-representation was

not as consistent as it was in the referral analysis, with some schools reflecting over-representation for girls in both the duplicated and unduplicated counts. Schools which reported over-representation in suspension for girls issued less than 10 total suspensions (regardless of gender).

These results are consistent with findings from other research. Skiba, et al. (2000) reported a number of studies that found boys to be over four times as likely as girls to be referred to the office or suspended, and Kaufman et al. (2010) reported that boys were significantly more likely to receive referrals across a number of different types of offenses. Kaufman et al. also noted that the types of offenses most likely to be sent to the office and documented in school-wide data are externalizing behaviors that boys are more likely to demonstrate (e.g., aggression, disruption, etc.). Offenses that girls are more likely to engage in – such as relational aggression and verbal teasing – are rarely captured on office referral forms. Kaufman et al. propose that gender disproportionality may be more of a result of how unacceptable behavior is defined and documented than a result of intentional or unintentional bias. While an excellent point, readers might also observe that part of the training curriculum for SW-PBS includes operationally defining and accurately recording inappropriate behavior (Florida's Positive Behavior Support Project, 2008). Given that the schools participating in the SWIS database were trained in SW-PBS and yet still experienced such a high degree of gender disproportionality in their outcomes, it may be that the current SW-PBS curriculum is not yet comprehensive enough to address their needs in this area.

Differences between counts: Referrals by gender. The current study builds on earlier findings through the analysis of referrals/suspensions (under the duplicated count) as well as the analysis of *students who receive* referrals/suspension (the unduplicated count). With respect to referrals, boys were more frequently represented among students who receive multiple referrals (i.e., when every incident was counted, there were more incidents for boys than for girls). While it can't be ruled out that a single boy in the school received a high number of referrals and so caused the overall count to be higher for boys, the consistency of the finding speaks to a broader

pattern of boys being more likely to be among a school's "frequent flyer" students, and that educators would likely benefit from bolstering their capacity for addressing boys' behavior.

For example, Ladson-Billings (2011) notes that, "While it is important to avoid gender stereotyping in the development of school curriculum we do have a responsibility to use available data to learn more about the interests and course taking patterns of boys in our schools" (p. 14). Developing activities that are more likely to engage boys in the academic curriculum is key to improving their academic achievement, and will likely lead to improvements in behavior as well. Reviewing discipline policies for minor infractions that unfairly target boys (such as wearing hats, or getting up from one's assigned seat) would also help to cut down on the negative student-teacher encounters that can lead to office referrals and suspensions.

Differences between counts: Suspensions by gender. The analysis of suspension revealed that boys were *not* represented as frequently among students who receive multiple suspensions as they were for receiving multiple referrals. There were also a number of schools (43% of the sample) where the duplicated and unduplicated count produced the same value, indicating that students (of either gender) who received suspension at those schools received the consequence only once. The relative infrequency of multiple suspension events is an encouraging finding, given the decreases in school attachment and increases in grade retention, dropout, and contact with the criminal justice system frequently experienced by students who have been suspended more than once (Christle et al., 2004; Costenbader & Markson, 1998; Gordon et al., 2000; Skiba & Rausch, 2006).

Comparison analysis: Referrals by race/ethnicity. Overall, the current study found that for office discipline referrals, African American, Hispanic, and White students were all subject to over-representation, but none to such an extent as was found for African American students. Over-representation in office referrals occurred much more frequently and to a much greater degree for African American students than it did for Hispanic or White students. For Hispanic students, there were more schools which reflected *under*-representation than over-representation, and the magnitude of the disproportionality was similar regardless of the direction of the imbalance. White students were also subject to over-representation in referrals, but the

magnitude of the disproportionality was much less than was found for Hispanic or African American students. For the most part, these findings were consistent regardless of the type of count used for the calculation. However, removing the impact of "frequent flyer" students tended to result in a greater number of schools revealing over-representation for African American and Hispanic students, whereas the opposite was true for White students (for whom over-representation was found more often using the duplicated count). It should be noted that although the number of schools indicating over-representation in referrals for African American students was higher using the unduplicated count, it was only higher by one school.

Comparison analysis: Suspensions by race/ethnicity. The over-representation of African American students in suspension was even more common than it was for referrals, and reflected larger differences between these students' percent of the student body and their representation in the duplicated and unduplicated counts. The percent of schools indicating over-representation in suspension for Hispanic students was similar to what was found with referrals, but the magnitude of the disproportionality tended to be much higher. For White students, over-representation was found *more often* with suspension than it was with referrals. This may be related to the finding that more schools exhibited over-representation for White students using the duplicated count than the unduplicated count. In other words, White students were sent to the office multiple times, and after multiple trips to the office school administrators may have been more likely to issue a suspension.

These findings are consistent with the results of earlier research showing that over-representation in school discipline begins at the point of the referral (Kaufman et al., 2010; Skiba, 2007). African American students were more likely than students of other racial backgrounds to experience over-representation in referrals, which carried over to this group of students being most likely to experience over-representation in suspension. Prior studies have found rates of over-representation for African American students in suspension that were two to more than three times higher than their White peers (Baker, Hendricks, McGowan & McKechnie, 2004; Gordon, Piana & Keleher, 2000; Skiba, Michael, Nardo & Peterson, 2000), and the current study supplements those findings by describing the number of schools where over-representation (of

any magnitude) was found, by describing the amount of referrals/suspensions attributed to students of different racial/ethnic groups, and by describing students' composition among those who receive these disciplinary consequences. The rampant over-representation of African American students in referrals and suspension was supported regardless of the type of measure used.

Under-representation of African American students. While this analysis revealed that African-American students' over-representation in referrals was widespread, it also demonstrated that *under*-representation of this group of students was possible – although the magnitude of under-representation was much less than the levels found in their over-representation. Under-representation in referrals tended to occur in schools with small African American student enrollments, but schools with African American student enrollment of less than 20% were also found to have some of the most extreme examples of *over*-representation. There did not appear to be a similar association between size of African American student enrollment and disproportionality in suspension, but there have been other examples that have found similar associations. Kozleski (2005) reported that schools with clear majority/minority populations (regardless of the racial compositions) are at high risk for having the minority population over-represented in special education. The findings in the current study provide additional support for this finding, but also caution against blanket assumptions based on the size of a school's minority enrollment.

Research question one: Relative risk for referrals. The first research question set out to investigate the overall distribution of risk ratios for office referrals and suspensions for African American and Hispanic students in elementary schools which implement SW-PBS, and posited that risk ratios for office referrals for African American students will generally indicate rates of referrals that are higher than would be expected given this group's distribution in their school's population, but that risk ratios for Hispanic students would be mixed. The results of the analysis support these hypotheses: in about half of the schools in the sample, African American students' risk for referrals was at least 50% higher than the risk of all other students, and reached levels as high as nearly four times the risk of all other students. Meanwhile, for Hispanic students only

about 10% of the schools in the sample indicated that this group's risk was at least 50% higher than the risk of all other students. However, because the comparison group used in the risk ratio was "all other students," African American students were included with students from all other race/ethnicities in the denominator of the ratio. With such widespread and high levels of over-representation for African American students, the risk ratios for Hispanic students may have been lower than if a more restricted comparison group, such as only White students, was used.

Choice of metric: "All other students." While the choice of comparison group may have lowered the apparent level of risk for Hispanic students, the metric was appropriate. Using "all other students" as the comparison group allows risk ratios to be calculated regardless of a school's demographic distribution. For example, in schools with student populations that are primarily composed of minority students, there may not be a group of White students that could be used in the calculation. When calculating risk ratios for many schools from a number of regions/states it's usually not possible to know the likely demographic makeup for all of the schools in advance; but the risk ratio will allow values to be calculated for all of the schools, regardless. The use of "all other students" as the comparison group may also be advantageous for resource-strapped school systems, as this metric clearly identifies the group in the school with the highest level of risk without having to compare calculations for other groups (e.g., the risk ratio for Hispanic students compared to White students versus the risk ratio for African American students compared to White students). The risk ratio, using "all other students" as the comparison group, provides a single summary of one group's relative risk.

Research question one: Relative risk for suspensions. The hypotheses for the first research question predicted the same pattern of results for suspensions as for referrals for the two groups of students: that risk ratios for suspensions for African American students will generally indicate rates that are higher than would be expected given this group's distribution in their school's population, but that risk ratios for Hispanic students would be mixed. These hypotheses were supported, with just over half of the schools in the sample presenting risk ratios for African American students that indicated at least a 50% higher risk of suspension than all other students. For Hispanic students, only 18% of the schools in the sample indicated that this

group's risk of suspension was at least 50% higher than the risk of all other students, but when over-representation was found, it was frequently of a magnitude similar to that found for African American students.

Extreme risk ratios: Reflections of inappropriate measurement. The analysis of risk ratios for the two groups of students revealed a few surprising values: between the two groups of students, there were five schools with risk ratios of 10 or higher – more than double the highest magnitude of risk reported in other studies. A closer examination of these schools found that in each case, fewer than 10 students (total, across all race/ethnicities) received suspension, and only one or two of these students were identified as African American or Hispanic. While the representation of each group of students in these few cases was *mathematically* disproportional, and the disciplinary climate for these students may have been less than ideal, the magnitude of over-representation suggested by a risk ratio of 10 or higher doesn't seem to fit these schools' suspension profiles. Shedding some light on this conundrum, Oliver-Schmidt and Kohlmann (2008) reported that the risk ratio can be used when "meaningful" incidences are available – in other words, the risk ratio may not be the most appropriate measure when very small *n*'s are being investigated. Given that less than 10 students experienced suspension in these few schools, the extreme values for the risk ratios may be more of a reflection of the application of an inappropriate metric than gross disproportionality.

Research question two: Disproportionality and SW-PBS implementation. The second research question aspired to determine whether schools which implement SW-PBS with higher levels of fidelity tended to have lower levels of disproportionality in office referrals and suspensions for African American and Hispanic students. The hypotheses speculated that schools with higher fidelity would tend to have lower levels of disproportionate referrals and/or suspensions, but based on the current results of visual and statistical analyses, that relationship is not supported. For office referrals, visual analyses of scatter plots depicting schools' Benchmarks of Quality (BoQ) scores and their risk ratios for African American and Hispanic students offered hints at a potential relationship. In both of these plots, schools with lower risk ratios tended to have higher BoQ scores. However, the relationship was not overwhelming, and

the results of the chi-square analyses determined that the association was not statistically significant for either African American or Hispanic students. With suspensions, scatter plots revealed a flat distribution of data points for both groups of students, and a more quantitative analysis suggested that higher-implementing schools may have actually had higher risk ratios (indicating higher levels of disproportionality, or inappropriate metrics as discussed earlier). The chi-square test confirmed that there was no statistically significant relationship between implementation and disproportionality in suspensions for either group of students.

Implementation scores and outcomes. The descriptive analysis of schools' BoQ scores and their corresponding referral and suspension rates (for all students) may have foreshadowed this finding: looking at scatter plots of schools' BoQ scores and their overall referral or suspension rates, there was a high degree of overlap between high- and low-implementing schools. Given that BoQ scores didn't appear to have a strong relationship with overall referral rates or overall suspension rates, it seems less surprising that BoQ scores didn't have a strong relationship with the risk of referral or risk of suspension for specific groups of students.

The lack of a strong association between BoQ scores and risk ratios, overall referral rates, and overall suspension rates should not be taken to mean that the BoQ lacks value as an evaluation tool. For one, the BoQ scores used in the current study only reflect a school's level of implementation at a specific point in history. A school's BoQ score at this point may be higher than it was in prior years, or it may represent a *decrease* from prior years' implementation. If a number of the BoQ scores from the current study were lower than these schools' scores in the past, this would indicate that SW-PBS was being implemented with less fidelity, and poorer student outcomes (including over-representation) would not be unexpected.

Additionally, differences between high- and low-implementing schools in the current study were detected in *average* referral rates and *average* suspension rates, suggesting that distinctions between these groups existed on a larger (aggregate) scale. When schools' rates were averaged by implementation level, high-implementing schools had ODR rates that were 25% lower and OSS rates that were 46% lower than the average for low-implementing schools. Existing studies and statewide evaluations (Cohen et al., 2007; Florida's Positive Behavior

Support Project, 2009) have likewise shown that higher-implementing schools have stronger behavioral and academic outcomes. With regards to school-level application, the process of completing the BoQ provides schools with prescriptive steps to continuously improve their implementation, so it's value as a diagnostic tool should not be overlooked. Given that the BoQ is a research-validated evaluation tool with demonstrated ability to differentiate student outcomes in larger-scale evaluations, future studies may choose to incorporate change scores when assessing schools' implementation status, so that the possibility of increases or decreases in fidelity may be accounted for.

Limitations

Missing data. One of the most troubling findings of this study was the overwhelming lack of referral-level race/ethnicity data maintained by the schools in this sample. Over half of the sample was missing this information on more than 10% of their data, with little difference by implementation level. By not keeping track of this information, schools are blind to potential problems. In light of more than 30 years of consistent findings of over-representation in school discipline, educators' reluctance to engage in an action as basic as acknowledging a student's race on an office referral boarders on negligence. Referral-level race/ethnicity data is necessary for any calculation of disproportionality, and if many referrals lack this information, the picture within any one school will quickly become distorted. In the current study, schools that were missing more than 10% of their referral-level race/ethnicity data were excluded from all analyses involving race or ethnicity. While removing these schools from the analyses greatly reduced the sample size and statistical power of the study, they nonetheless had to be eliminated in order to provide some validity to the picture of relative risk for African American and Hispanic students in this investigation.

Cut scores and categorization. The process of identifying high- and low- implementing schools, and high-, moderate- and low levels of disproportionality may also have been problematic. The split scores used in this analysis have not been applied to other research or evaluation, so the reliability and validity of these particular high- and low-implementing criteria (scores of 87 and higher, and scores of 79 and lower, respectively) haven't been established.

Likewise, the criteria used to define “low,” “moderate,” and “high” disproportionality groups were based on general guidelines gleaned from state definitions of “significant disproportionality” and advocate recommendations. The lack of association may be less due to the variables than to the way in which they were grouped.

Representation. The degree to which the current sample was representative of other schools which implement SW-PBS may also give pause. The databases used to collect referral and implementation data were utilized on a volunteer basis, and some states contributed considerably more volunteers than others. Two states in particular contributed more schools than the other six schools combined, so it cannot be assumed that the nationwide sample was truly representative.

Accuracy. The volunteer nature of the databases, as well as the fact that both the implementation and outcome measures were self-reported data, meant that the accuracy of the information was not controlled. In one example, a school’s self-reported school-level demographic information didn’t match their referral records, which indicated significantly more African American students were enrolled than were reported. While it is possible that in this example, the school-level demographic information was accurate and the higher number of African American students found in the referral-level information was due to several different students enrolling, being given office referrals, and then matriculating from the school throughout the year, this theory could not be verified. Ultimately, this school had to be excluded from all analyses that included race/ethnicity, and while the information provided by the school may have been valid, it serves to demonstrate the potential for inaccuracy inherent in the data set.

Considerations for future research and practice.

Continued promise. Given that proponents of culturally-responsive classroom management (CRCM) turn to the evidence base of SW-PBS as support for their recommended strategies (Cartledge & Kourea, 2008), the results of the current study – including the preponderance of missing data – may seem surprising. Certainly, one would hope that an approach which shares so much overlap with current recommendations for culturally-responsive practices would result in at least some improvement in disproportional discipline. However, the

current study does not close the door on the possibility that SW-PBS offers an avenue for improvement. The limitations noted above, particularly concerning the small number of schools involved in the analysis (and the resulting lack of statistical power), may have contributed to the lack of association between implementation and disproportionality. With so few examples, especially with what may have been tenuous categorization criteria and a less-than-representative sample of schools, it is difficult to determine the role implementation fidelity plays in disproportional referral and suspension outcomes with any finality. In addition, finding instances of under-representation in the different disciplinary consequences, and finding that several schools reported very few to no suspension for their students offers some hope that the proactive and positive approach of SW-PBS may ultimately impact disproportionality in student discipline.

Support for the theory of cultural mismatch. The idea that disproportionality may be tied to a mismatch between the culture of the school and the culture of the students and families it serves (one of the tenets of CRCM) may be indirectly supported through the findings of this analysis. The fact that disproportionate outcomes persisted in many of the schools in this study in spite of staff from these schools receiving professional development that was targeted to improving the degree to which discipline was administered consistently from staff member to staff member, suggests that either the professional development did not carry over into practice, or that the information was applied in a way that did not address differences between groups of students. Given the large segment of schools with higher BoQ scores, the professional development likely carried over into practice, but was applied in a way that did not address differences between groups of students. As Utley et al. (2002) suggest, schools which implement SW-PBS must go beyond agreement between their staff members on which of their values and beliefs will be addressed in the school-wide plan, and move to incorporate the values and beliefs of the families they serve into everyday practice. One example of this can be found at Chi-Dodge Elementary in Yahtahey, New Mexico (Skiba, 2007). Rather than base the school-wide system on the values and beliefs of just their staff, educators at Chi-Dodge found ways to incorporate students' background into the expectations and lesson plans. The school-wide expectations incorporated the majority of students' first language, Dine, and the lessons used to teach the

expectations featured historical figures from the Dine culture. Parents were welcomed into the school as partners in their children's education, and made to feel comfortable at the school. When surveyed, 80% of students could recite the school-wide expectations, and the referral rate at Chi-Dodge was the third lowest in the state of New Mexico.

The need for deliberate examination. However, in schools with small minority populations, non-majority students may still be left out: educators may incorporate the values and beliefs of the families they serve, but miss the small proportion of families who are in the minority. In addition, if professional development curricula teach that SW-PBS is intended to benefit 80-90% of the student body, it may be possible that when school-based PBS teams see this proportion of their students benefitting from their system, they do not look further to see if students from different groups are benefitting in an equitable fashion. In the current study, many of the minority rates for schools were small, with half of the sample reporting rates of 28% or less. Based on the available information, it could not be determined how frequently school teams analyzed or applied disaggregated information, or if they ever did. The findings of over-representation in this study, and the lack of association between implementation and disproportionality, may have been due to the fact that the schools weren't trying to correct the problem of over-representation in discipline, because they did not see that it existed.

The core features of SW-PBS do not currently focus on ensuring that SW-PBS teams disaggregate their school-wide data, and it is possible that schools do not receive this message until they undergo more advanced training (e.g., Tier 2 PBS; Florida's Positive Behavior Support: Response to Intervention Project, 2010). Therefore, explicit direction for schools to screen their discipline data for evidence of disproportionality at Tier 1 may be needed in order for SW-PBS to deliver on the promise supporters of CRCM have offered.

Clarifying the different measures. To complicate matters, there is no "one way" for schools to disaggregate their data, and the different measures can result in different interpretations. For example, in the comparison analysis for referrals for Hispanic students, there was a noticeable jump in the number of schools that showed evidence of over-representation when the unduplicated count was used compared to the number of schools showing over-

representation when the duplicated count was used. Risk ratios can be employed, but can be calculated using "all other students" as the comparison group, or by selecting a specific group (such as only White students) to serve as the comparison group, which may then impact findings of disproportionality. Furthermore, measures used in research studies frequently (but not consistently) employ the use of yet another measure, the odds ratio, which has been demonstrated to produce values that are more extreme than those found through the risk ratio, and has an additional concern regarding the interpretation of the values (Oliver-Schmidt & Kohlmann, 2008). Specifically, lay-persons have the tendency to interpret odds ratios as if they *were* a risk ratio, interpreting the *odds* of something occurring as if they were the same as the *chance* of something occurring. Odds, however, compare the chance of something occurring to the chance of it *not* occurring, a point which is commonly overlooked in interpretation.

Add to this the fact that risk ratios, although considered by some to be "best practice" in measuring disproportionality (Kozleski, 2005), may not always be the most appropriate measure to use when the total number of events under investigation is small. In the current study, schools with fewer than 10 suspensions produced risk ratios that were more than double the most extreme values reported in other research, even though suspension was used with only one or two students from the target group. When all of this information is considered together, it begins to seem plausible that the reason disproportionality continues to exist may have something to do with the fact that educators and researchers haven't been consistent in the way they define it. With so many options available, each providing valuable information as well as pitfalls, the identification of "disproportionality" has the potential to cause more confusion and misunderstanding than it does a path towards targeted interventions.

Applying multiple measures. Take the example provided in Chapter 4, where three schools shared a risk ratio that was very close in value. Table 11 is provided again for reference below. In all three schools, African American students are approximately 50% more likely to receive a referral than all other students, but noticeably more African American students were involved in the over-representation in two of the three schools (the unduplicated comparison was close to 50% higher than expected). In addition, one of the schools reflected a high value for the

duplicated comparison, suggesting that African American students at this particular school were very much over-represented in the degree to which they received multiple incidents – almost 150% higher than their proportion of the school population, and much higher than what was seen at the other two schools. When the information is studied on a school by school basis, it becomes clear that although the relative risk of African American students is approximately the same at all three schools, the ways in which each school needs to intervene will likely be different.

Table 29

Relationship of Risk Ratio and Comparison Analysis Findings for Referrals, African American students (reprint)

School ID	Risk Ratio	Unduplicated Comparison	Duplicated Comparison
171149	1.49	46.5%	148.6%
140486	1.50	22.8%	32.7%
4878	1.54	43.5%	50%

Note: Unduplicated Comparison and Duplicated Comparison reflect the percent difference of each of these measures compared to African American students' proportion of the student population (i.e., "magnitude of disproportionality").

In school 171149, African American students' representation among those who receive referrals and among students who receive multiple referrals is relatively high, suggesting that this school would need relatively broad and/or intensive intervention to address the disparity. With more students involved and receiving so many more referrals than would be expected, it would not be surprising to learn that student-teacher relationships for African American students at this school were suffering, and that a number of staff would benefit from increased support in addressing these students' behavior. At school 140486, African American students' representation among those who receive referrals and among students who receive multiple referrals is comparatively low, suggesting that this school may not need to intervene as intensively – and may only need to engage in individual consultation with a limited number of staff. School 4878 lies somewhere in-between. The current approach to professional

development for schools with disproportionality – blanket in-services for the entire school staff – have so far been limited in their effectiveness; teaching schools to intervene in a targeted manner will result in better deployment of resources, and offers the possibility for schools to identify strategies their staff is already using that result in equitable discipline. Clearly, using multiple data sources together will result in a more accurate picture and plan for intervention than using any one measure alone.

A hierarchy of analysis. Because the risk ratio provides an easily understandable way of summarizing relative risk compared to all other groups of students, this metric may serve as a good starting point for identifying over-representation in school discipline. If risk ratios for different groups were lower than one, school teams could move on to their next problem solving target. If risk ratios were greater than one, the total number of students experiencing the consequence of concern could then be examined (to alert personnel to the possibility of an invalid risk ratio), along with the percent differences for the unduplicated and duplicated counts. This information is easily obtainable, easy to interpret, and together provides a comprehensive picture of disproportionality for a group of students. Once this information is considered, thorough problem analysis will likely lead to targeted and effective interventions.

Researchers' choice of metric. On larger scales (district, state, or national), investigators can support practitioners' efforts first by focusing their research on school-by-school findings of disproportionality, as opposed to aggregated findings based on a number of schools. While reports based on aggregated numbers can be easier to calculate and can help garner attention for a cause, they are limited in the degree to which they can generate information for effective hypotheses generation and intervention planning. As can be seen through the wide range of school-level findings in the present study, basing the unit of analysis on the number or percent of schools reflecting disproportionality of *x* degree offers a much clearer picture of the circumstances facing our schools. Agreeing on a standard metric (or set of metrics) for reporting disproportionality will facilitate inter-study comparisons and potentially deepen readers' understanding of relevant variables. For example, extending the current study's approach to the middle and high school levels could result in a more complete picture of how disproportionate

outcomes can change across settings and developmental levels. Consistency in procedures and measures will similarly improve investigations of the types of behaviors leading to referrals and suspensions, as well as the behaviors of educators who write the referrals. Once researchers are speaking a common language, a common understanding of the variables is sure to follow.

Creating the groundwork for change. Before any of these recommendations can be utilized, it is vital for schools to improve their track record with maintaining accurate data. Inaccuracies in data collection limit educators' ability to intervene effectively and can delay students' access to necessary supports. Given the extensive history of disproportionality across a number of different outcome measures (such as academic achievement, placement in special education, and discipline), maintaining accurate record-level race/ethnicity information should not be an option. At the same time, widespread findings of incomplete data on office referrals suggest the need for broad, systems-level supports to make the desired behavior more likely to occur. Advances in technology have made it easier for separate databases to communicate and share information, but the processes involved in accomplishing this feat continue to be complex and frequently expensive. However, the trade off may well be worth it: with improved data collection and problem identification, educators can move to repair decades of imbalance and take real steps towards helping *all* of their students succeed.

List of References

Achilles, G., McLaughlin, M., & Croninger, R. (2007). Sociocultural correlates of disciplinary exclusion among students with emotional, behavioral, and learning disabilities in the SEELS national dataset. *Journal of Emotional and Behavioral Disorders, 15*(1), 33-45.

Baker, B., Hendricks, J., McGowan, J., & McKechnie, M. (2004). Eliminating the Achievement Gap: Reducing minority overrepresentation in school discipline. Juvenile Rights Project. Retrieved from the Juvenile Rights Project website: www.jrplaw.org/Overrep.htm

Barrett, S., Bradshaw, C., & Lewis-Palmer, T. (2008). Maryland statewide PBS initiative. *Journal of Positive Behavior Interventions, 10*(2), 105-114.

Bazron, B., Osher, D., & Fleischman, S. (2005). Creating culturally responsive schools. *Educational Leadership, 63*(1), 83-84.

Blanchett, W.J. (2006). Disproportionate representation of African American students in special education: Acknowledging the role of White privilege and racism. *Educational Researcher, 35*(6), 24-28.

Blanchett, W.J, Mumford, V., & Beachum, F. (2005). Urban school failure and disproportionality in a Post- *Brown* era. *Remedial and Special Education, 26* (2), 70-81.

Bohanon, H., Fenning, P., Carney, K., Minnis-Kim, M., Anderson-Harriss, S., Moroz, K.,…Pigott, T. (2006). Schoolwide application of positive behavior support in an urban high school. *Journal of Positive Behavior Interventions, 8*(3), 131-145.

Bollmer, J., Bethel, J., Garrison-Mogren, R., & Brauen, M. (2007). Using the risk ratio to assess racial/ethnic disproportionality in special education at the school-district level. *Journal of Special Education, 41*(3), 186-198.

Bondy, E., Ross, D., Gallingane, C., & Hambacher, E. (2007). Creating environments of success and resilience: Culturally responsive classroom management and more. *Urban Education, 42*(4), 326-348.

Brophy, J. (1988). Research linking teacher behavior to student achievement: Potential implications for instruction of Chapter 1 students. *Educational Psychologist, 23*(3), 235-286.

Brown v. Board of Educ., 347 U.S. 483 (1954). Retrieved from http://www.nationalcenter.org/brown.html

Burdette, P. (2007, July). State definitions of significant disproportionality. *In Forum* (Brief policy analysis). Retrieved from http://www.projectforum.org/docs/StateDefinitionsofSignificantDisproportionality.pdf

Cartledge, G., & Kourea, L. (2008). Culturally responsive classrooms for culturally diverse students with and at risk for disabilities. *Exceptional Children, 74*(3), 351-371.

Cartledge, G., Tillman, L.C., & Talbert Johnson, C. (2001). Professional ethics within the context of student discipline and diversity. *Teacher Education and Special Education, 24*(1), 25-37.

Christle, C., Nelson, M., & Jolivette, K. (2004). School characteristics related to the use of suspension. *Education and Treatment of Children, 27*(4), 509-526.

Cohen, R., Kincaid, D., & Childs, K.E. (2007). Measuring school-wide positive behavior support implementation: Development and validation of the Benchmarks of Quality. *Journal of Positive Behavior Interventions, 9*(4), 203-213.

Costenbader, V., & Markson, S. (1998). School suspension: A study with secondary school students. *Journal of School Psychology, 36*(1), 59-82.

Delpit, L. (1995). *Other People's Children.* New York: New Press.

De-Navas-Walt, C., Proctor, D., & Smith, J. (2009). U.S. Census Bureau, Current Population Reports, P60-236. *Income, Poverty, and Health Insurance Coverage in the United States: 2008.* U.S. Government Printing Office, Washington, D.C.

Douglas, D. (2005). *Jim Crow Moves North: The Battle Over Northern School Segregation, 1865 – 1954* (pp. 1-11). New York: Cambridge University Press.

Dunlap, G., Sailor, W., Horner, R., & Sugai, G. (2009). Overview and history of Positive Behavior Support. *Handbook of Positive Behavior Support* (pp. 3-16). New York: Springer Publishing.

Durham, J.T. (2003). The other side of the story: The world of African American academies in the South after the Civil War. *The Negro Educational Review, 54*(1-2), 3-16.

Educational and Community Supports (2010). About PBS Surveys [webpage describing the PBS Surveys web application]. Retrieved from www.pbssurveys.org

Educational and Community Supports (2010b). About SWIS [webpage describing the SWIS web application]. Retrieved from www.swis.org

Fairclough, A. (2004). *The* costs of Brown: Black teachers and school integration. *The Journal of American History, June 2004,* 43-55.

Farkas, G. (2003). Racial disparities and discrimination in education: What do we know, how do we know it, and what do we need to know? *Teachers College Record, 105*(6), 1119-1146.

Farkas, G.. Lleras, C., & Maczuga, S. (2002). Does oppositional culture exist in minority and poverty peer groups? *American Sociological Review, 67*(1), 148-155.

Field, A. (2009). *Discovering Statistics Using SPSS, 3rd Edition* (pp. 686-724). Thousand Oaks, CA: Sage Publications.

Fenning, P. & Rose, J. (2007). Overrepresentation of African American students in exclusionary discipline. *Urban Education, 42*(6), 536-559.

Ferri, B.A., & Conner, D.J. (2005). In the shadow of *Brown*: Special Education and students of color. *Remedial and Special Education, 26*(2), 93-100.

Fisher, D. (2009). The use of instructional time in the typical high school classroom. *The Educational Forum, 73,* 168-176.

Florida's Positive Behavior Support Project. (2008). *School-Wide Positive Behavior Support Team Training manual.* Training manual, College of Behavior and Community Sciences, University of South Florida, Tampa, Florida.

Florida's Positive Behavior Support Project. (2009). Outcome and Evaluation Data for Annual Report, 2008-2009. Retrieved from the FLPBS website: http://flpbs.fmhi.usf.edu/pdfs/DOE%20Final%20Report%201.08.10.pdf

Florida's Positive Behavior Support: Response to Intervention Project. (2010). *Team Training on Tier 2 Positive Behavior Support: Response to Intervention for Behavior.* Training manual, College of Behavior and Community Sciences, University of South Florida, Tampa, Florida.

Gilbertson, D., Duhon, G., Witt, J., & Dufrene, B. (2008). Effects of academic response rates on time-on-task in the classroom for students at academic and behavioral risk. *Education and Treatment of Children, 31*(2), 153-165.

Gordon, R., Piana, L.D., & Keleher, T. (2000). Facing the consequences: An examination of racial discrimination in U.S. public schools. Applied Research Center. Retrieved from the Applied Research Center website: www.arc.org/pdf/196apdf.pdf

Green, T. D. (2005). Promising prevention and early intervention strategies to reduce overrepresentation of African American students in special education. *Preventing School Failure, 49*(3), 33-41.

Greenwood, C.R., Horton, B.T., & Utley, C.A. (2002). Academic engagement: Current perspectives on research and practice. *School Psychology Review, 31*(3), 328-349.

Horner, R., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A., & Esperanza, J. (2009). A randomized, wait-list controlled effectiveness trial assessing school-wide positive behavior support in elementary schools. *Journal of Positive Behavior Interventions, 11*(3), 133-144.

Horsford, S.D., & McKenzie, K.B. (2008). 'Sometimes I feel like the problems started with desegregation': exploring Black superintendent perspectives on desegregation policy. *International Journal of Qualitative Studies in Education, 21*(5), 443-455.

Irvin, L., Tobin, T., Sprague, J. , Sugai, G., & Vincent, C. (2004). Validity of office discipline referral measures as indices of school-wide behavioral status and effects of school-wide interventions. *Journal of Positive Behavior Interventions, 6* (3), 131-147.

Johns, B., Crowley, P., & Guetzloe, E. (2008). Engaged time in the classroom. *Focus on Exceptional Children, 41*(4), 1-6.

Kaufman, J., Jaser, S., Vaughan, E., Reynolds, J., Donato, J., Bernard, S., Hernandez-Brereton, M. (2010). Patterns in office referral data by grade, race/ethnicity, and gender. *Journal of Positive Behavior Interventions, 12* (1), 44-54.

KewalRamani, A., Gilbertson, L., Fox, M., and Provasnik, S. (2007). *Status and Trends in the Education of Racial and Ethnic Minorities* (NCES 2007-039). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Kourea, L., Cartledge, G., & Musti-Rao, S. (2007). Improving the reading skills of urban elementary students through total class peer tutoring. *Remedial and Special Education, 28*(2), 95-107.

Kozleski, E. (2005 February 17). *The disproportionate representation of culturally and linguistically diverse students in special education* [Transcript of National Center on Secondary Education and Transition conference call presentation]. Retrieved from http://www.ncset.org/teleconferences/transcripts/2005.02.asp

Krezmien, M.P., Leone, P.E., & Achilles, G.M., (2006). Suspension, race and disability: Analysis of statewide practices and reporting. *Journal of Emotional and Behavioral Disorders, 14*(4), 217-227.

Kunjufu, J. (2006). *An African Centered Response to Ruby Payne's Poverty Theory.* Chicago, IL: African American Images.

Ladson-Billings, G. (2011). Boyz to men? Teaching to restore Black boys' childhood. *Race, Ethnicity and Education, 14*(1), 7-15.

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 32*(3), 465-491.

Lassen, S., Steele, M., & Sailor, W. (2006). The relationship of school-wide positive behavior support to academic achievement in an urban middle school. *Psychology in the Schools, 43*(6), 701-712.

Lee, J., Grigg, W., & Dion, G. (2007). *The Nation's Report Card: Mathematics 2007* (NCES 2007–494). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

Lee, J., Grigg, W., & Donahue, P. (2007). *The Nation's Report Card: Reading 2007* (NCES 2007–496). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

Lewis-Palmer, T., Horner, R., Todd, A., & Sugai, G. (2005). School-Wide Evaluation Tool (Version 2.1) [Evaluation tool for SW-PBS implementation]. Retrieved from http://www.pbis.org/pbis_resource_detail_page.aspx?Type=4&PBIS_ResourceID=222

Lohrmann-O'Rourke, S., Knoster, T., Sabatine, K., Smith, D., Horvath, B., & Llewellyn, G. (2000). School-Wide application of PBS in the Bangor area school district. *Journal of Positive Behavior Interventions, 2*(4), 238-240.

Losen, D. & Orfield, G. (2002). Racial Inequity in Special Education: Executive Summary for Federal Policy Makers. Retrieved from the Civil Rights Project website: http://www.civilrightsproject.ucla.edu/research/specialed/IDEA_paper02.php

Luiselli, J., Putnam, R., & Sunderland, M. (2002). Longitudinal evaluation of behavior support intervention in a public middle school. *Journal of Positive Behavior Interventions, 4*(3), 182-188.

McComas, J., Hoch, H., Paone, D., & El-Roy, D. (2000). Escape behavior during academic tasks: A preliminary analysis of idiosyncratic establishing operations. *Journal of Applied Behavior Analysis, 33*(4), 479-493.

McCurdy, B., Mannella, M., & Eldridge, N. (2003). Positive behavior support in urban schools: Can we prevent the escalation of antisocial behavior? *Journal of Positive Behavior Interventions, 5*(3), 158-170.

McFadden, A.C., Marsh II, G.E., Price, B.J., & Hwang, Y. (1992). A study of race and gender bias in the punishment of handicapped school children. *The Urban Review, 24*(4), 239-251.

McFadden, A.C., & Marsh II, G.E. (1992). A study of race and gender bias in the punishment of school children. *Education and Treatment of Children, 15*(2), 140-147.

Mickelson, R.A. (2001). Subverting Swann: First- and second-generation segregation in the Charlotte-Mecklenburg Schools. *American Educational Research Journal, 38*(2), 215-252.

Monroe, C.R., & Obidah, J.E. (2004). The influence of cultural synchronization on a teacher's perceptions of disruption. *Journal of Teacher Education, 55*(3), 256-268.

Monroe, C.R. (2006). African American boys and the discipline gap: Balancing educators' uneven hand. *Educational Horizons, Winter 2006*, 102-111.

National Assessment of Educational Progress, National Center for Educational Statistics, National Assessment Governing Board, Institute of Education Sciences, U.S. Department of Education (2009). *Data tools for the National Assessment of Educational Progress* [Data tables for national minority trends in Reading and Math, 4th and 8th grades]. Retrieved from http://nationsreportcard.gov/data_tools.asp

National Center for Education Statistics. (n.d.). *Appendix D: NCES Locale Codes.* Retrieved from the National Center for Education Statistics website: http://nces.ed.gov/programs/handbook/data/pdf /appendix_d.pdf

Nelson, J.L., Palonsky, S.B., & McCarthy, M.R. (2004). The academic achievement gap: Old remedies or new. *Critical Issues in Education: Dialogues and Dialectics, 5th Edition* (pp. 103-126). Boston: McGraw-Hill.

Netzel, D., & Eber, L. (2003). Shifting from reactive to proactive discipline in an urban school district: A change of focus through PBS implementation. *Journal of Positive Behavior Interventions, 5*(2), 71-79.

Newman, L., Wagner, M., Cameto, R., Knokey, A.-M. (2009). *The Post-High School Outcomes of Youth With Disabilities up to 4 Years After High School*. *A Report From the National Longitudinal Transition Study-2 (NLTS2)* (NCSER 2009-3017). Menlo Park, CA: SRI International.

Newport, F. (2007). Black or African American? Gallup News Service. Retrieved from http://www.gallup.com/poll/28816/black-african-american.aspx

Oliver Schmidt, C. , & Kohlmann, T. (2008). When to use the odds ratio or relative risk? *International Journal of Public Health, 53*, 165-167.

Pane, D.M. (2009). Viewing classroom discipline as a negotiable social interaction: A communities of practice perspective. *Teaching and Teacher Education,* doi:10.1016/j.tate.2009.05.002.

Peshak-George, H., Kincaid, D., & Pollard-Sage, J. (2009). Primary-Tier interventions and supports. In W. Sailor, G. Dunlap, G. Sugai, & R. Horner (Eds.), *Handbook of Positive Behavior Support* (pp. 375-394). New York: Springer Publishing.

Putnam, R., McCart, A., Griggs, P., & Hoon Choi, J. (2009). Implementation of schoolwide positive behavior support in urban settings. In W. Sailor, G. Dunlap, G. Sugai, & R. Horner (Eds.), *Handbook of Positive Behavior Support* (pp. 443-463). New York: Springer Publishing

Raffaele-Mendez, L. & Knoff, H. (2003). Who gets suspended from school and why: A demographic analysis of schools and disciplinary infractions in a large school district. *Education and Treatment of Children, 26*(1), 30-51.

Scott, T., & Barrett, S. (2004). Using staff and student time engaged in disciplinary procedures to evaluate the impact of School-Wide PBS. *Journal of Positive Behavior Interventions, 6*(1), 21-27.

Skiba, R. (2007, October). *Changing the data, changing our minds: Disproportionality and SW-PBS.* Keynote presentation presented at the 2007 PBIS National Implementers' Forum, Chicago, IL. Power Point retrieved from the PBIS website: http://www.pbis.org/common/pbisresources/presentations/RSkiba_keynote.pdf

Skiba, R. (2000). Zero Tolerance, Zero Evidence: An Analysis of School Disciplinary Practice. Policy Research Report #SRS2, Indiana Education Policy Center. Retrieved from http://www.indiana.edu/~safeschl/ztze.pdf

Skiba, R., Michael, R., Nardo, A., & Peterson, R. (2000). The Color of Discipline: Sources of Racial and Gender Disproportionality in School Punishment. Policy Research Report #SRS1, Indiana Education Policy Center. Retrieved from the Education Commission of the States website: http://www.ecs.org/html/offsite.asp?document=http%3A%2F%2Fwww%2Eindiana%2Edu%2F%257Esafeschl%2Fpublication%2Ehtml

Skiba, R., Poloni-Staudinger, L., Simmons, A., Feggins-Azziz, R., & Chung, C. (2005). Unproven links: Can poverty explain ethnic disproportionality in special education? *Journal of Special Education, 39*(3), 130-144.

Skiba, R., & Rausch, M.K. (2006). School disciplinary systems: Alternatives to suspension and expulsion. In G.G. Bear & K.M. Minke (Eds.),*Children's Needs III: Development, Prevention, and Intervention* (pp. 87-102). Bethesda, MD: National Association of School Psychologists.

Skiba, R., Simmons, A., Ritter, S., Gibb, A., Rausch, M., Cuadrado, J., & Chung, C. (2008). Achieving equity in special education: History, status, and current challenges. *Exceptional Children, 74*(3), 264-288.

Smith, B. (2000). Quantity matters: Annual instructional time in an urban school system. *Educational Administration Quarterly, 36*(5), 652-682.

Spaulding, S., Irvin, L., Horner, R., May, S., Emeldi, M., Tobin, T., & Sugai, G. (2010). School-Wide social behavioral climate, student problem behavior, and related administrative decisions: Empirical patterns from 1,510 schools nationwide. *Journal of Positive Behavior Interventions, 12* (2), 69-85.

Stichter, J.P., Stormont, M, & Lewis, T.J. (2009). Instructional practices and behavior during reading: A descriptive summary and comparison of practices in Title 1 and non-Title elementary schools. *Psychology in the Schools, 46*(2), 172-183.

Sugai, G., Horner, R.H., Dunlap, G., Hieneman, M., Lewis, T.J., Nelson, C.M.,…Ruef, M. (2000). Applying Positive Behavior Support and Functional Behavioral Assessment in schools. *Journal of Positive Behavior Interventions, 2*(3), 131-143.

Sugai, G. & Horner, R. (2009). Defining and describing school-wide positive behavior support. In W. Sailor, G. Dunlap, G. Sugai, & R. Horner (Eds.), *Handbook of Positive Behavior Support* (pp. 307-326). New York: Springer Publishing.

Sugai, G., & Horner, R.H. (2007). Advanced Topics in PBS: Secondary/Tertiary Interventions. Power Point presentation retrieved from the OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports website: http://pbis.org/pbis_resource_detail_page.aspx?Type=1&PBIS_ResourceID=205

Sugai, Lewis-Palmer, Todd, & Horner. (2005). Schoolwide Evaluation Tool version 2.1. Educational and Community Supports: University of Oregon.

Talbert-Johnson, C. (2004). Structural inequities and the achievement gap in urban schools. *Education and Urban Society,* 37(1). 22-36).

Tomlinson, J. (2004). Number of instructional days/hours in the school year. Downloaded from the Education Commission of the States website: http://www.ecs.org/clearinghouse/78/24/7824.pdf

Ukpokodu, N. (2002). Breaking through pre-service teachers' defense dispositions in a multicultural education course: A reflective practice. *Promising Practices, Spring*, 25-33.

University of Oregon (2010). *School-Wide Information System: Dataset 002.* [Database and Codebook]. Eugene, OR, University of Oregon.

U.S. Department of Education, Office of Special Education and Rehabilitative Services, Office of Special Education Programs. (2007). *27th Annual (2005) Report to Congress on the Implementation of the Individuals with Disabilities Education Act,* Volume 1, Washington, D.C.

Utley, C., Kozleski, E., Smith, A., & Draper, I. (2002). Positive Behavior Support: A proactive strategy for minimizing behavior problems in urban multicultural youth. *Journal of Positive Behavior Interventions, 4*(4), 196-207.

Vanneman, A., Hamilton, L., Baldwin Anderson, J., and Rahman, T. (2009). *Achievement Gaps: How Black and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress*, (NCES 2009-455). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Vavrus, F., & Cole, K. (2002). "I didn't do nothin'": The discursive construction of school suspension. *The Urban Review, 34*(2), 87-111.

Walker, V.S. (2001). African American teaching in the South: 1940-1960. *American Educational Research Journal, 38*(4), 751-779.

Ware, F. (2006). Warm demander pedagogy: Culturally responsive teaching that supports a culture of achievement for African American students. *Urban Education, 41*(4), 427-456.

Warren, J., Bohanon-Edmonson, H., Turnbull, A., Sailor, W., Wickham, D., Griggs, P., & Beech, S. (2006). School-Wide positive behavior support: Addressing behavior problems that impede student learning. *Educational Psychology Review, 18*, 187-198.

Weinstein, C., Curran, M., Tomlinson-Clarke, S. (2003). Culturally responsive classroom management: Awareness into action. *Theory into Practice, 42*(4), 269-276.

Zirkel, S. (2005). Ongoing issues of racial and ethnic stigma in education 50 years after Brown v. Board. *The Urban Review, 37*(2), 107-126.

Zirkel, S., & Cantor, N. (2004). 50 Years after *Brown v. Board of Education*: The promise and challenge of Multicultural Education. *Journal of Social Issues, 60*(1), 1-15.

Appendices

Appendix A. Reporting of 2004 Special Education Placement Data by Race

In 2007, the National Center for Education Statistics published a report entitled, “Status and Trends in the Education of Racial and Ethnic Minorities” (KewalRamani, Gilbertson, Fox, & Provasnik, 2007). At the time this proposal was written, this report supplied the most up to date data that “examined the educational progress and challenges that racial and ethnic minorities face in the United States” (introduction, iii). In reporting the number and percentage of children receiving special education services, the authors divided the ESE populations into two age categories: ages 3-5 and ages 6-21 (Table 8.1b, page 40). In order for this author to compare the racial composition of each ESE population to the racial composition in the total K-12 population, the age categories from the ESE populations were combined.

Additionally, in reporting the enrollment status of students by racial group in the K-12 student population, the KewalRamani et al. report (2007) only provided percentages of students for each racial category, and the number of students for total enrollment (Table 7.2, page 29). Numbers of students for each racial category’s composition in the K-12 population were therefore calculated by this author.

Appendix B. School-Level Characteristics

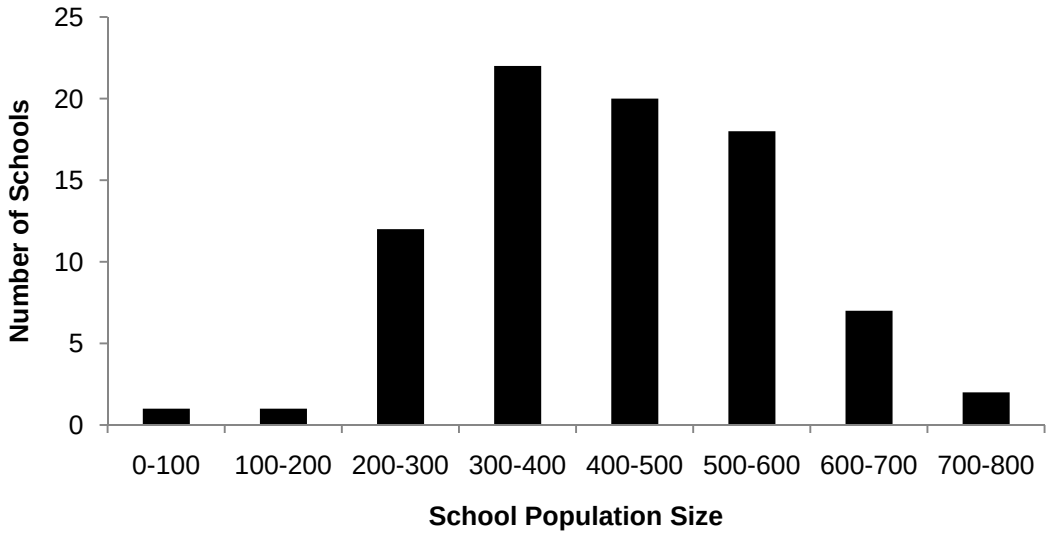

Figure 43. Distribution of school populations included in the sample (n=83).

School Population. School enrollment ranged from 88 to 769 students (*M* = 424, *SD* = 133). The majority of the schools in the sample (72%) had populations between 300-600 students, with relatively fewer schools with populations in the extreme ends of the distribution (skewness = .17, kurtosis = -.15). Only two schools had populations larger than 700 students, and only two schools had populations smaller than 200 students.

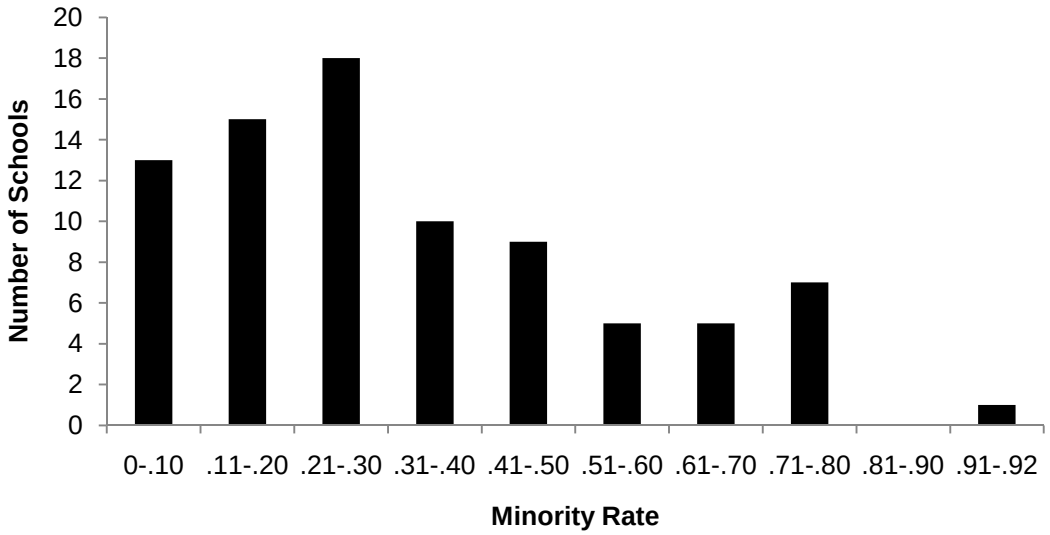

Figure 44. Distribution of minority rates for schools in the sample (n=83).

Minority rate. In order to get a sense of the amount of diversity within each school, the minority rate was calculated. This rate was established by determining the number of non-white students at each school (total enrollment minus the number of students identified as white) and dividing the result by the school's total enrollment. Schools' minority rates ranged from 4% to 92% (*M* = .33, *SD* = .22). At more than twice the standard error of skewness, the distribution of minority rates was found to be significantly positively skewed (skewness = .69; *SE* of skewness = .264), but relatively normally distributed in terms of the concentration of rates around the mean (kurtosis = -.43; *SE* of kurtosis = .523). The majority of the schools in the sample (78%, or 65 schools) had minority rates of 50% or less, with half of the sample containing schools with even lower minority rates (28% or less). Only eight schools had minority rates higher than 70%, and only one had a minority rate higher than 80%. With exceptions in only 22% of the sample, it appears that most commonly, students who were identified as white made up the majority of each school's population.

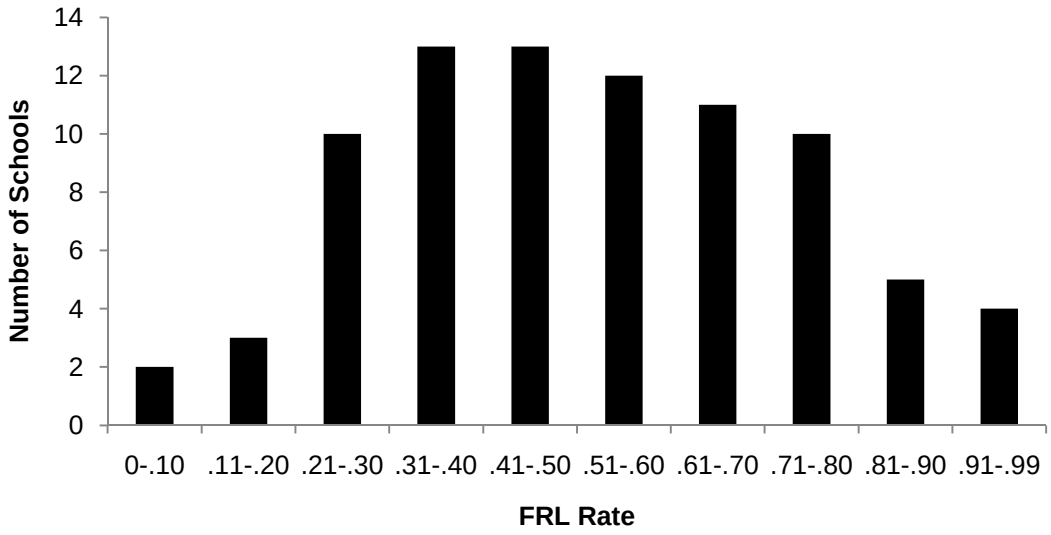

Figure 45. Distribution of Free/Reduced Lunch rates for schools in the sample (n=83).

Title 1 and Free/Reduced Lunch enrollment. A total of 81% of the schools were Title 1 eligible, and all schools reported Free and Reduced Lunch (FRL) enrollment. Given that the final sample had variation in the sizes of each school's population, the usefulness of an analysis of the number of students who were enrolled in the FRL program would be limited unless it took school size into account. Therefore, the percent of the school's population enrolled in the FRL program (the FRL rate) was calculated for each school. The percentage of FRL-enrolled students ranged from 7% of the school's population to 99% of the school's population (*M* = .52, *SD* = .23). Every school had a different percentage of students enrolled in the FRL program, so a mode could not be calculated for this characteristic. The skewness (.16) and kurtosis (-.68) reflected a relatively balanced and broad distribution of schools for this characteristic, which is supported by the finding that only five schools had FRL rates lower than 21%, and only four schools had FRL rates higher than 90%. Eighty-Three percent of the sample had FRL rates between 20 and 80 percent.

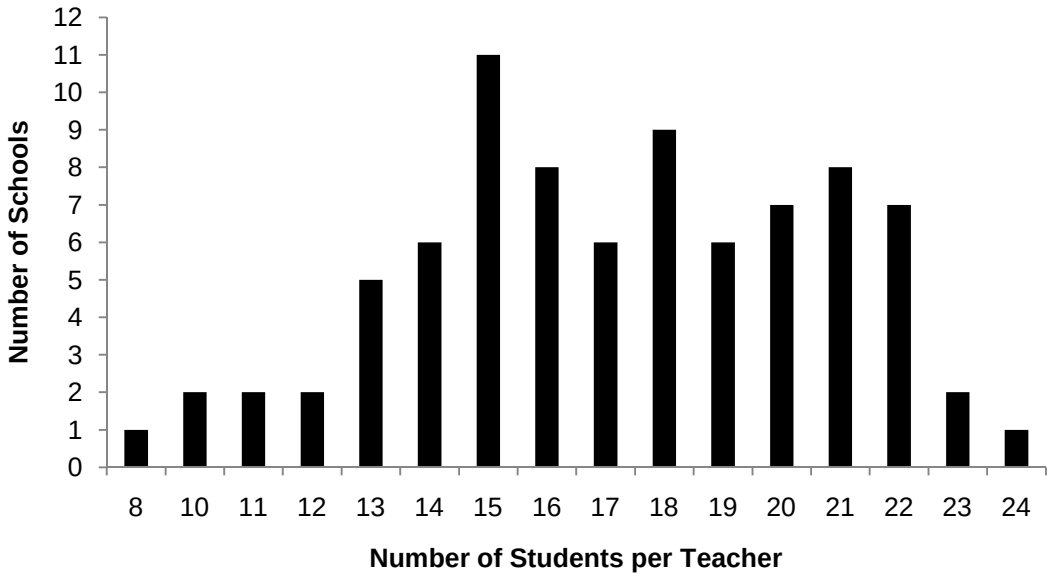

Figure 46. Distribution of student to teacher ratios for schools in the final sample.

Student-to-Teacher ratio. As with the number of students enrolled in the FRL program, the amount of classroom teacher FTE at any one school was not helpful for understanding the overall characteristics of the sample unless the metric included information about the school's enrollment. Therefore, a student-to-teacher ratio was calculated by dividing each school's enrollment by their classroom teacher FTE. The resulting number gave an indication of approximately how many students may be assigned to one classroom teacher at a particular school. The average student-to-teacher ratio was 17:1 (*SD* = 3.5), with the smallest ratio being 8:1 and the largest ratio being 24:1. The distribution of schools' ratios was fairly normally distributed (skewness = -.24; kurtosis = -.52), with just over 20% of the schools having ratios of 14:1 or smaller, and 75% of the schools having ratios between 15:1 and 22:1. Only three schools had student-to-teacher ratios higher than 22:1.

Appendix C: Rank-Order of Schools' Percent of Referrals Accounted for by Boys (duplicated count) and Boys' Composition among Students who Received Referrals (unduplicated count)

School ID	Duplicated	Unduplicated
182	66.67	61.07
184	67.07	63.43
3052	67.20	63.60
167675	68.71	73.50
183	69.34	70.68
310	69.59	62.96
4342	70.44	67.82
3489	71.99	63.50
4913	72.46	76.12
2666	72.65	67.02
194501	72.73	63.28
3494	73.15	77.97
212	73.63	69.50
5071	74.17	67.62
138770	74.31	74.19
134014	74.33	66.25
1962	74.57	60.94
3497	75.14	68.35
345	75.38	56.59
129708	75.71	65.91
4878	75.76	67.11
2659	76.33	70.89
4675	76.43	69.86
140486	76.86	65.08
3496	77.12	67.53
135954	77.20	76.32
3410	77.24	68.84
2006	77.28	73.79
4087	77.66	74.16
1139	77.78	67.70
1919	77.79	65.75
4061	78.04	67.45
209	78.14	70.19
129709	78.64	68.66
210	78.66	65.83
195182	79.47	67.82
138769	79.50	72.57
3553	79.90	69.47
1973	79.96	69.78
4062	80.09	71.43
3492	80.11	72.73
185	80.45	72.31
177	80.45	69.05
3498	80.53	73.79
139339	80.82	74.58
3163	80.88	75.00
171993	81.06	74.71
129695	81.10	77.05
1313	81.79	75.45
138771	81.80	64.15
5038	81.92	65.64
1010	82.19	73.73
140488	82.22	79.07
168728	82.67	72.46
599	82.83	76.19
1982	82.92	74.10
186	83.56	80.00
2310	83.62	75.00
206	83.67	76.52
4734	83.72	75.00
167670	83.94	68.33
140210	84.51	80.95
311	84.52	70.94
167666	84.90	72.32
141185	85.57	67.59
200	85.58	74.29
5069	85.73	73.33
3934	85.90	79.75
2664	86.08	82.35
3548	86.57	77.61
1140	86.80	82.52
201	87.19	73.55
2434	87.38	79.57
7	87.56	77.71
170	88.68	74.39
1981	89.92	83.67
1883	90.77	82.61
171149	91.10	71.79
167669	91.29	82.46
3491	91.32	85.09
1086	92.31	86.36
3401	92.90	85.71
135953	94.17	90.00

Appendix D: Risk Ratio and Comparison Analysis Findings for Referrals, African American Students

School ID	Risk Ratio	Unduplicated Comparison	Duplicated Comparison
2664	0	---	---
3494	0	---	---
3489	0.28	-72.2%	-55.6%
345	0.40	-60.8%	-64.3%
599	0.60	-38.9%	-23.1%
2434	0.78	-21.3%	-30.5%
212	0.92	-11.3%	-37.5%
1883	0.92	-9.4%	30.6%
186	1.08	8.4%	34.1%
3491	1.12	9.6%	75.0%
3934	1.17	15.1%	54.5%
167670	1.20	19.0%	7.1%
167666	1.21	19.0%	146.7%
1313	1.28	29.9%	50.0%
170	1.35	28.4%	140.1%
182	1.46	46.8%	80.8%
206	1.48	44.9%	21.4%
183	1.49	44.0%	-25.5%
171149	1.49	46.5%	148.6%
140486	1.50	22.8%	32.7%
4878	1.54	43.5%	50.0%
1962	1.65	52.2%	72.8%
171993	1.69	22.6%	24.1%
140488	1.70	50.8%	-3.8%
3052	1.71	26.8%	45.1%
4342	1.74	34.3%	59.5%
5038	1.76	14.4%	12.6%
4061	1.82	38.6%	86.1%
3553	1.83	58.5%	141.4%
177	1.86	25.9%	10.8%
1140	1.87	51.6%	79.6%
5071	2.08	50.7%	83.1%
2006	2.29	92.4%	128.6%
185	2.36	130.8%	70.0%
2310	2.41	84.3%	82.9%
129709	3.43	224.5%	217.4%
2659	3.53	261.7%	1114.3%
1086	3.89	278.8%	116.7%
1981	3.91	189.1%	231.7%

Note: Unduplicated Comparison and Duplicated Comparison reflect the percent difference of each of these measures compared to African American students' proportion of the student population (i.e., "level of disproportionality").

Appendix E: Relationship of Risk Ratio and Comparison Analysis Findings for Referrals, Hispanic Students

School ID	Risk Ratio	Unduplicated Comparison	Duplicated Comparison
599	0	---	---
1086	0	---	---
171993	0.24	-73.1%	-74.0%
3052	0.44	-50.8%	-66.3%
3489	0.48	-49.5%	-36.4%
1981	0.50	-47.9%	-66.3%
5071	0.59	-39.7%	-78.5%
2434	0.60	-38.6%	-54.3%
4878	0.64	-33.4%	-55.4%
167666	0.67	-21.6%	-44.1%
183	0.68	-26.0%	-56.6%
2006	0.70	-25.4%	-50.9%
4342	0.72	-25.7%	-47.0%
140486	0.73	-22.4%	-40.9%
3553	0.74	-21.2%	-57.3%
5038	0.74	-24.6%	-47.1%
2310	0.78	-19.9%	32.7%
177	0.79	-20.0%	-65.7%
212	0.80	-17.9%	-2.1%
1313	0.94	-5.3%	13.1%
129709	0.97	-2.7%	-4.3%
185	0.98	-2.3%	-34.9%
170	0.99	-0.4%	-22.4%
345	1.01	0.6%	-33.9%
171149	1.01	1.2%	7.9%
4061	1.02	2.7%	-35.7%
1962	1.07	6.5%	0.0%
1140	1.08	6.4%	-55.5%
167670	1.10	7.9%	17.0%
1883	1.11	9.4%	-21.4%
3494	1.15	10.5%	-12.3%
2659	1.16	15.1%	-36.4%
3491	1.22	20.2%	71.2%
182	1.26	21.4%	-9.1%
186	1.31	26.5%	56.1%
140488	1.52	47.4%	21.1%
206	1.57	51.8%	85.7%
2664	1.77	73.0%	608.8%
3934	1.79	66.9%	44.0%

Note: Unduplicated Comparison and Duplicated Comparison reflect the percent difference of each of these measures compared to Hispanic students' proportion of the student population (i.e., "level of disproportionality").

Appendix F. Rank-Order of Schools' Percent of Suspension Events Accounted for by Boys

School ID	Duplicated	Unduplicated
200	0.0%	0.0%
3934	0.0%	0.0%
2664	14.3%	33.3%
209	50.0%	50.0%
129695	50.0%	50.0%
138769	50.0%	50.0%
138770	50.0%	50.0%
184	62.5%	62.5%
4342	63.5%	71.1%
167675	66.7%	66.7%
183	72.7%	75.0%
212	75.0%	70.0%
3494	75.0%	50.0%
3498	75.0%	70.0%
134014	76.5%	70.7%
3410	78.3%	67.7%
182	78.6%	77.8%
4734	78.6%	75.0%
4061	79.5%	76.2%
1919	79.5%	75.9%
1973	79.6%	75.9%
167670	80.0%	85.7%
140486	80.2%	79.0%
3489	81.0%	81.8%
5038	81.1%	68.9%
4675	81.8%	79.2%
185	83.3%	75.0%
140488	84.8%	83.3%
177	85.1%	86.4%
4878	86.0%	85.7%
1313	86.4%	80.8%
168728	86.5%	90.9%
171993	86.7%	86.8%
4087	87.5%	87.5%
167666	87.5%	75.0%
194501	87.5%	80.0%
2006	87.8%	82.1%
186	88.2%	80.0%
1883	88.6%	92.9%
3496	88.9%	87.5%
4062	88.9%	83.8%
2434	89.7%	80.0%
5071	90.0%	81.3%
5069	90.3%	82.8%
195182	90.4%	85.7%
3052	91.2%	86.5%
345	92.9%	86.2%
135954	93.3%	90.9%
1140	93.8%	90.9%
139339	93.8%	87.0%
167669	95.5%	92.3%
311	95.8%	93.8%
1981	96.0%	93.8%
201	96.2%	92.3%
1962	96.7%	94.7%
7	100%	100%
170	100%	100%
206	100%	100%
210	100%	100%
310	100%	100%
599	100%	100%
1010	100%	100%
1086	100%	100%
1139	100%	100%
1982	100%	100%
2310	100%	100%
2659	100%	100%
2666	100%	100%
3163	100%	100%
3401	100%	100%
3491	100%	100%
3492	100%	100%
3548	100%	100%
3553	100%	100%
4913	100%	100%
129709	100%	100%
135953	100%	100%
138771	100%	100%
141185	100%	100%
171149	100%	100%

Appendix G: Risk Ratio and Comparison Analysis Findings for Suspensions, African American Students

School ID	Total # Students Receiving Suspension	Risk Ratio	Unduplicated Comparison	Duplicated Comparison
183	8	0	---	---
185	4	0	---	---
206	4	0	---	---
212	10	0	---	---
345	29	0	---	---
1086	4	0	---	---
1313	26	0	---	---
2659	3	0	---	---
2664	3	0	---	---
3489	11	0	---	---
3491	7	0	---	---
3494	2	0	---	---
3934	4	0	---	---
129709	3	0	---	---
2434	15	0.8	-18.7%	-57.9%
171993	38	1.27	10.7%	6.6%
177	44	1.29	11.0%	-3.8%
140486	100	1.40	18.8%	18.2%
5038	135	1.58	11.9%	7.3%
140488	18	1.69	50.2%	-1.7%
3052	37	1.90	31.8%	31.5%
4342	38	2.08	45.5%	63.5%
4878	28	2.50	107.0%	135.9%
186	10	2.59	143.9%	43.5%
4061	84	2.63	62.1%	105.4%
1962	19	2.92	132.2%	169.6%
2006	28	3.22	143.0%	132.3%
5071	32	3.48	85.4%	85.6%
2310	4	3.61	130.4%	130.4%
3553	4	4.38	168.8%	222.6%
1981	16	4.41	212.5%	200.0%
1883	14	4.53	296.8%	58.7%
167666	8	4.68	316.7%	108.3%
171149	6	5.52	376.2%	1042.9%
1140	11	7.26	170.4%	132.3%
182	9	10.54	754.7%	1548.4%
170	3	11.12	338.6%	463.9%
167670	7	11.83	920.4%	614.3%
599	3	12.20	754.7%	1182.1%

Note: Unduplicated Comparison and Duplicated Comparison reflect the percent difference of each of these measures compared to African American students' proportion of the student population (i.e., "level of disproportionality").

Appendix H: Risk Ratio and Comparison Analysis Findings for Suspensions, Hispanic Students

School ID	Total # Students Receiving Suspension	Risk Ratio	Unduplicated Comparison	Duplicated Comparison
170	3	0	---	---
182	9	0	---	---
185	4	0	---	---
599	3	0	---	---
1086	4	0	---	---
1140	11	0	---	---
1981	16	0	---	---
2310	4	0	---	---
2659	3	0	---	---
2664	3	0	---	---
3489	11	0	---	---
3494	2	0	---	---
3553	4	0	---	---
3934	4	0	---	---
5071	32	0	---	---
167666	8	0	---	---
171149	6	0	---	---
171993	38	0	---	---
3052	37	0.25	-71.1%	-84.3%
4878	28	0.27	-70.5%	-80.8%
1883	14	0.42	-53.6%	-81.4%
183	8	0.44	-48.8%	-25.5%
345	29	0.56	-38.4%	-52.2%
2006	28	0.58	-35.6%	-56.1%
140486	100	0.58	-36.4%	-44.7%
167670	7	0.64	-30.7%	-51.5%
177	44	0.66	-32.2%	-52.4%
140488	18	0.76	-21.8%	-57.3%
4342	38	0.77	-21.1%	-36.5%
4061	84	0.84	-15.0%	-74.8%
5038	135	1.08	8.9%	-0.1%
1313	26	1.12	9.3%	16.2%
1962	19	1.66	59.5%	1.0%
2434	15	1.94	90.5%	-1.5%
212	10	4.07	215.8%	163.2%
186	10	4.74	225.2%	186.9%
206	4	4.96	296.8%	217.5%
3491	7	5.07	291.4%	413.7%
129709	3	10.38	624.6%	443.5%

Note: Unduplicated Comparison and Duplicated Comparison reflect the percent difference of each of these measures compared to African American students' proportion of the student population (i.e., "level of disproportionality").