Multilevel Exploration of Factors Contributing to the Overrepresentation of Black Students in Office Disciplinary Referrals

Catherine P. Bradshaw, Mary M. Mitchell, Lindsey M. O'Brennan, and Philip J. Leaf
Johns Hopkins University

Although there is increasing awareness of the overrepresentation of ethic minority students—particularly Black students—in disciplinary actions, the extant research has rarely empirically examined potential factors that may contribute to these disparities. The current study used a multilevel modeling approach to examine factors at the child (e.g., teacher-rated disruptive behavior problems) and classroom or teacher levels (e.g., teacher ethnicity, level of disruptive behavior in classroom) that may contribute to the overrepresentation of minority students in office disciplinary referrals (ODRs). Data come from 6,988 children in 381 classrooms at 21 elementary schools. The analyses indicated that even after controlling for the student's level of teacher-rated behavior problems, teacher ethnicity, and other classroom factors, Black students were significantly more likely than White students to receive ODRs. Results also suggested that ethnic match between students and their teachers did not reduce the risk for referrals among Black students.

Keywords: Black, office discipline referrals, behavior problems, teachers, multilevel modeling

Although the overall rates of serious forms of school violence have declined over the past decade, disruptive behavior problems continue to be a great concern in the classroom (DeVoe et al., 2004). Disruptive behavior problems are of particular concern in school settings because they are associated with attendance problems, lower levels of academic success, greater risk for placement in special education programs, and later school failure (Gottfredson, 2001; Ialongo, Poduska, Werthamer, & Kellam, 2001; Ialongo et al., 1999; Patterson, DeBaryshe, & Ramsey, 1989). Frequently, youths who exhibit disruptive and aggressive behaviors are suspended from school and as a result are typically segregated with other delinquent youths upon their return, which consequently can increase their deviant behavior (Atkins et al., 2002; Gottfredson, 2001; Wu, Pink, Carin, & Moles, 1982). Students of color, specifically Black students, are overrepresented in suspensions and office referrals, as well as referrals to special education (Eitle & Eitle, 2004; Raffaele Mendez & Knoff, 2003; Skiba et al., 2008; Skiba, Michael, Nardo, & Peterson, 2002). However, the extant research has rarely empirically examined potential factors at both the student and classroom levels that may contribute to these disparities.

The current article used a multilevel approach to explore the extent to which a child's behavior problems, as well as classroom characteristics and teacher ethnicity, may contribute to the overrepresentation of Black students in office disciplinary referrals (ODRs). Having an enhanced understanding of factors at multiple levels that influence the overrepresentation of ethnic minority students in ODR data may prove helpful in identifying potential avenues for reducing these discrepancies.

Use of Office Disciplinary Referrals in Schools

ODRs are widely used by schools to monitor student discipline problems, evaluate the impact of school-based interventions and policies (Irvin, Tobin, Sprague, Sugai, & Vincent, 2004; Sugai, Sprague, Horner, & Walker, 2000), and make programmatic decisions regarding support services, both school-wide and for individual students (Irvin et al., 2006). ODRs have been correlated with rebellious behavior, drug use, and increased classroom disorderliness (Nelson & Roberts, 2000; Sprague et al., 2001). Moreover, Tobin and Sugai (1999) found that ODR data predicted disruptive behavior in later grades, as well as problem behavior in adulthood. Despite the increasing reliance on ODR data as a measure of student behavior problems, some researchers have questioned the validity of ODRs, because disciplining practices can vary depending on preexisting school administration practices and implementation of rules (Morrison, Redding, Fisher, & Peterson, 2006; Morrison & Skiba, 2001) and may vary in response to administrative changes, programs, or initiatives. Furthermore, the referral process is based solely on teachers' subjective appraisal of the situation (Irvin et al., 2004). As such, ODRs may be susceptible to contextual factors or potential bias. This is particularly disconcerting, given that ODRs can have negative effects for students, such as reduced opportunities for learning (Scott & Barrett, 2004) and increased risk for truancy, suspension, affiliation with deviant peers, and subsequent school failure (Morrison & Skiba, 2001; Skiba & Peterson, 1999).

Overrepresentation of Black Students in ODRs

Of great concern is the growing number of studies documenting the overrepresentation of Black students in disciplinary data. For example, Black students are from 3 to 7 times more likely than White students to be suspended (Raffaele Mendez & Knoff, 2003; Wald & Losen, 2003). Similarly, Black students are more likely than White students to receive an ODR (Skiba et al., 2008).

Recent findings suggest that Black students are often referred to the office for less serious offenses than their White counterparts (Monroe, 2005; Skiba et al., 2008). Skiba et al. (2002) found that Black students were more likely to be referred to the office for subjective offenses (e.g., disrespect, threat, excessive noise), whereas White youth were more likely to be referred for objective behavior problems (e.g., smoking, vandalism). Furthermore, Gottlieb et al. (1991) found that teachers were more likely to refer Black students for special education primarily for behavioral problems, whereas White youth were more likely referred for academic difficulties (Gottlieb et al., 1991). This disproportionate representation of ethnically diverse students in disciplinary data contradicts the increased efforts to reduce the academic gaps between cultural groups through national mandates, such as the No Child Left Behind Act (Donovan & Cross, 2002). Thus, it is imperative to further understand the context that supports disproportionate referrals to the office, which may in turn result in inappropriate referrals to special education or academic problems for minority youth.

Potential Contributing Factors

Although there has been increased concern about the overrepresentation of Black students in ODR and other disciplinary data, there have been relatively few empirical studies examining possible factors that contribute to these discrepancies. One potential hypothesis is that the increased rates of ODRs among Black youth are a consequence of higher base rates of behavior problems. Prior research on teacher ratings of student behavior has shown small but significantly higher levels of disruptive behavior and other externalizing behavior problems among Black students compared to their White peers (Epstein et al., 2005; Koth, Bradshaw, & Leaf, 2009; Sbarra & Pianta, 2001). However, several researchers have theorized that the overrepresentation of minority students in ODR and other disciplinary data (e.g., suspensions) reflects a potential cultural bias embedded in school discipline practices (Monroe, 2005; Townsend, 2000). More specifically, a potential cultural disparity between students and their teachers may contribute to the overrepresentation of Black students in ODRs. Prior research has suggested that cultural values such as communalism (i.e., social identity is linked with group membership), verve (i.e., proclivity toward sensation stimulation), and use of movement are particularly salient within the Black population (see Tyler et al., 2008, for full review of the literature). However, studies have shown that public school teachers rarely endorse these culturally valued behaviors in the classroom (Tyler, Boykin, Miller, & Hurley, 2006), despite research supporting the use of culturally sensitive teaching strategies as a means of improving academic achievement among minority populations (American Psychological Association [APA], 2002; Boykin, Lilja, & Tyler, 2004). Furthermore, competitive and individualistic instructional methods that are commonly used in most classrooms have been shown to be less effective with Black students than White students (APA Task Force on Resilience and Strength in Black Children and Adolescents, 2008).

This possible mismatch in values can increase the likelihood of a discrepancy between what minority students perceive as being "appropriate" behavior and what teachers and administrators hold as acceptable standards for student behavior (Skiba et al., 2008). For example, teachers working in a school that subscribes to mainstream cultural values may interpret culturally normative behaviors of Black youth (e.g., freedom of expression) as being disrespectful, combative, or argumentative (Monroe, 2005; Weinstein, Curran, & Tomlinson-Clarke, 2004). This potential cultural incongruity has also been shown to affect how teachers view students' academic capabilities. For example, Neal, McCray, Webb-Johnson, & Bridgest (2003) found that Black students who walked with a "stroll" were more likely to be perceived by teachers as being lower in achievement, higher in aggression, and more likely to need special education services. A similar effect was observed among White students who were perceived by their teachers as "acting Black" (Neal et al., 2003).

Given the potential cultural discontinuity between home and school for Black youth, there is an increased likelihood for a cycle of misconduct. When students perceive a lack of respect, differential treatment, and low expectations from teachers, they are less likely to feel connected to their school and are at a greater risk for future misconduct and school dropout (Furrer & Skinner, 2003; Roeser, Midgley, & Urdan, 1996). For example, Sheets (1996) found that students of color perceived that the disproportionality in discipline practices was deliberate and biased. Likewise, in accordance with the self-fulfilling prophecy, youth who perceive that teachers view them as hostile may in turn become more hostile or aggressive in the school environment because of these negative interactions. Further evidence that minority students tend to feel disconnected from their school environment comes from research on school climate, which has shown that Black students tend to feel less supported and less connected to school (Koth, Bradshaw, & Leaf, 2008). Student disengagement is particularly common when the school's values do not match those being modeled in the student's home or community (APA Task Force, 2008; Ford, 1993; Tyler et al., 2008). Given their increased risk of feeling less connected to school (Greenwood, Horton, & Utley, 2002; Hawkins, Doueck, & Lishner, 1988; Koth et al., 2008), it is not surprising that Black youth are also at increased risk for academic failure (National Center for Educational Statistics, 2004a, 2004b), which may be due in part to the disproportionate use of ODRs with these students.

The use of different disciplinary practices may also vary by both the student's and teacher's ethnicity (Murray & Murray, 2004). There is increasing interest in promoting a racial and/or ethnic match between students and their teachers, on the theory that teachers who are more similar—ethnically and culturally—to their students may better identify with and relate to their students (Ladson-Billings, 1997, 2002). Yet few studies have directly examined the extent to which teachers' use of ODRs varies as a function of their ethnicity, and more specifically, whether Black students are overrepresented in the ODRs made by Black teachers. Together, the extant research suggests that there may be a potential bias among teachers and administrators regarding normative cul-

tural behavior, which in turn may contribute to the overrepresentation of minority students in ODR data (Skiba et al., 2008).

Overview of the Current Study

Although there is an increasing number of studies documenting ethnic disparities in ODR data and other discipline data, few studies have empirically examined other factors that may contribute to these discrepancies. Therefore, we considered the student's overall level of behavior problems, characteristics of the classroom (i.e., overall level of disruption), and the teacher's ethnicity as potential factors that may contribute to the overrepresentation of Black students in ODR data. For example, in light of the increasing emphasis on the ethnic match between the students and teachers (Murray & Murray, 2004), we explored whether teachers who are also Black are as likely as their White colleagues to refer Black students to the office for discipline problems. A unique aspect of this study is that we were able to control for the individual child's teacher rating of disruptive behavior, which is a source of information absent in previous studies on ODRs. We also examined aspects of the classroom environment that may increase the odds of receiving an ODR (Koth et al., 2008), such as the average level of behavior problems occurring within the classroom, as well as the percentage of students in the class receiving an ODR. An enhanced understanding of the factors associated with an increased risk of ODRs among Black students may guide policies or professional development that aim to reduce disproportionality in ODR data.

Method

Sample

The data for this study come from a sample of 6,988 children enrolled at 21 elementary (K–5) schools participating in a randomized trial of School-Wide Positive Behavioral Interventions and Supports (SWPBIS). All of the 21 schools were randomized to the intervention condition in which they received training in SWPBIS and were

implementing the program with high fidelity; the remaining 16 schools in the trial that were randomized to the comparison condition did not systematically collect data on ODRs. Therefore, the 16 comparison schools were not included in the current study (for additional information on the trial, see Bradshaw, Mitchell, & Leaf, 2009). The students in the current analyses were nested within 381 classrooms in the 21 intervention schools. The participating elementary schools had an average enrollment of 472 students, and 42.9% of the student body received free or reduced-cost meals.

The original sample included 422 teachers and 8,645 students from the 21 schools. However, given our interest in contrasting Black and White teachers and students, we excluded data for 41 teachers because they were neither Black nor White or were missing data on their race or ethnicity. As a result, the approximately 840 students nested within those 41 classrooms were dropped from the analyses. In addition, approximately 810 students were excluded from the analyses because they were neither Black nor White or were missing data on race or ethnicity. Our final sample included 381 teachers and 6,988 children, all of whom were either Black or White. There were too few staff or students of other ethnicities (e.g., $n = 62$ American Indian students; $n = 339$ Asian or Pacific Islander students; $n = 286$ Hispanic students) to examine these groups separately. Student and teacher demographics for the current sample of participants are provided in Table 1. This study was approved by the institutional review board at Johns Hopkins University.

Measures

ODRs. ODR data were obtained from two sources: the classroom teachers and the School-Wide Information System (SWIS; May et al., 2003), an Internet-based data system used to manage student discipline referral data. With regard to the teacher-reported ODR data, each teacher indicated on a student demographic form at the end of the school year whether or not the child had been referred to the principal's office during that school year. This item

Table 1

Demographic Characteristics of the Participating Students and Classrooms

Variable	M	SD	%
Student level ($n = 6,988$)			
Male			52.9
Ethnicity			
Black			51.8
White			48.2
Disruptive Behavior rating by teacher (TOCA-C)	1.95	0.84	
Classroom level ($n = 381$ classrooms/teachers)			
Teacher ethnicity			
Black			10.5
White			89.5
Average Disruptive Behavior rating for class (TOCA-C)	1.91	0.40	
% Children in class with a teacher-reported ODR	14.15	13.26	
% Children within class with an ODR from SWIS			
Any ODR	17.67	14.44	
Major ODR	13.47	12.33	
Minor ODR	6.68	12.57	
Fighting ODR	7.59	8.78	
Defiance ODR	4.90	6.88	

Note. TOCA-C = Teacher Observation of Classroom Adaptation–Checklist; ODR = office discipline referral; SWIS = School-Wide Information System.

was dichotomized into no referrals (0) and one or more referrals (1) and is denoted in the current article as *teacher-reported ODR*. With regard to the SWIS ODR data, the teachers in all 21 schools received training from the school administration in the fall of the school year on how to complete a standardized ODR form to document each major or minor ODR. The referrals were entered daily into the online SWIS program by the schools' secretarial staff. We created the following five dichotomous SWIS ODR indicators based on the student-level SWIS data: any type of ODR (major or minor), a major ODR (e.g., fighting, defiance, abusive language, bullying or harassment), a minor ODR (e.g., disruption, property misuse), a major ODR for fighting, and a major ODR for defiance. Each student received a dichotomous code of 0 = none or 1 = one or more for each of the five SWIS ODR categories.

Teacher ratings of student behavior. Teacher ratings of each student's classroom behaviors were obtained using the Teacher Observation of Classroom Adaptation–Checklist (TOCA-C; Koth et al., 2009), a checklist version of the Teacher Observation of Classroom Adaptation–Revised (TOCA-R; Werthamer-Larsson, Kellam, & Wheeler, 1991). The TOCA-C contains a series of brief items that are rated by the teacher on a 6-point scale from *never* (1) to *almost always* (6). The Disruptive Behavior subscale included nine items (e.g., breaks rules, fights, harms property, teases class-mates; α = .90), with higher ratings indicating more problematic behavior. The psychometric properties of the TOCA have been examined extensively, with test–retest correlations over a 4-month interval being .75 or higher (Werthamer-Larsson et al., 1991). The Disruptive Behavior subscale in elementary school also has been shown to have high predictive validity, as evinced by the strong association with violence among adolescents (Petras, Chilcoat, Leaf, Ialongo, & Kellam, 2004) and criminality in young adults (Schaeffer et al., 2006; Schaeffer, Petras, Ialongo, Poduska, & Kellam, 2003). Individual ratings on the TOCA-C were modeled as a student-level covariate, whereas aggregated subscale scores were modeled as classroom-level covariates to adjust for the levels of class-wide behavior problems.

Student-level demographic covariates. Each student's gender (1 = male vs. 0 = female), ethnicity (1 = Black vs. 0 = White), and grade level (represented by a continuous variable from 0 [kindergarten] through Grade 5) were included in the multilevel models as student-level covariates.

Classroom-level variables. Characteristics of the teacher and/or classroom were entered as classroom-level covariates. Teacher ethnicity was dummy coded as 1 = Black versus 0 = White. As noted above, the classroom average ratings for Disruptive Behavior from the TOCA-C were included as a classroom contextual risk factor. In addition, the percentage of students in the classroom with one or more of the corresponding ODRs (e.g., percentage with any ODR; percentage with major ODR; percentage with minor ODR; percentage with ODR for fighting; and percentage with ODR for defiance) were included as covariates at the classroom level to adjust for the teacher's overall use of ODRs.

Overview of the Analyses

The purpose of the study was to examine the extent to which student and teacher characteristics were associated with a student's risk for an ODR. We modeled these associations using a two-level modeling procedure with student-level receipt of an ODR (i.e.,

referral or no referral) as the dichotomous dependent variable. The results are reported in terms of odds ratios, which indicate the odds of a student receiving an ODR while adjusting for the covariates of interest. All data were collected within a single school year.

After significant data cleaning and preliminary analyses, we found that the assumptions of logistic regression (e.g., linearity in the logit, absence of multicolinearity and outliers) were tenable (Tabachnick & Fidell, 2001). Linearity of the logit was achieved by using a Laplace transform of the dependent variable with the penalized quasilikelihood estimator for logistic regression (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2004). Minimal bivariate correlations between independent variables indicated the absence of multicollinearity (e.g., $r < .2$ for student-level variables, and $r < .4$ for classroom-level variables), and an examination of residuals indicated few outliers with minimal leverage. Subsequently, we fit a series of two-level models in the HLM software (Version 6.01; Raudenbush et al., 2004) to examine the influence of student- and classroom-level factors on students' odds of receiving each of the six ODR types.

The student-level covariates included ethnicity (i.e., Black vs. White), gender, grade, and teachers' ratings of each student's disruptive behaviors on the student-level TOCA-C, whereas co-variates at the classroom-level included the teacher's race or ethnicity (i.e., Black vs. White), the percentage of the students in the class receiving the corresponding ODR, and the mean class-room ratings of disruptive behaviors (via the TOCA-C; Rauden-bush & Bryk, 2002). Continuous covariates were centered using grand-mean centering, whereas dichotomous covariates were re-tained as uncentered. Because all six outcomes were dichotomous, we fit the models with a Bernoulli distribution for the outcome variables in HLM. After testing the main effects of the student- and classroom-level covariates (Model 1), we tested whether there was a significant interaction between student ethnicity and being a male student (Model 2). We then refitted the models with a cross-level interaction between student ethnicity and teacher eth-nicity (Model 3). Finally, we fitted a model that included a three-way interaction between student ethnicity, teacher ethnicity, and student gender (Model 4). Fit indices including the Akaike infor-mation criterion and the Bayesian information criterion were cal-culated for the models (Akaike, 1974; Schwarz, 1978).

Results

In our first set of models, we examined the main effect of the student- and classroom-level covariates on the receipt of each of the six types of ODRs. We found that after controlling for the student- and classroom-level covariates, Black children had sig-nificantly greater odds of receiving a teacher-reported ODR (co-variate adjusted odds ratio [AOR] = 1.35, $p < .01$), any ODR (AOR = 1.24, $p < .01$), a minor ODR (AOR = 1.82, $p < .001$), and an ODR for fighting (AOR = 1.26, $p < .05$) compared with White children (see Tables 2, 3, 5, and 6). However, there were no significant effects of student ethnicity on the odds of receiving a major ODR (see Table 4) or an ODR for defiance (see Table 7). Student gender also had a consistent effect across ODR categories, such that boys had more than twice the odds of receiving any ODR type or having a teacher-reported ODR, with AORs ranging from 2.16 to 3.69 ($p < .001$). The effects of TOCA-C disruptive behavior problems scores on the odds of a referral also were in the

Table 2

Multilevel Analyses Examining the Association Between Student- and Classroom-Level Characteristics and the Odds of Receiving a Teacher-Reported Office Discipline Referral (ODR)

	Teacher-reported ODRs					
	Main effects model		Child gender × Ethnicity interaction		Three-way interaction	
Variable	AOR	95% CI	AOR	95% CI	AOR	95% CI
Child level						
Male	3.40***	[2.63, 4.39]	2.37***	[1.66, 3.37]	2.37***	[1.68, 3.34]
Ethnicity: Black[a]	1.35**	[1.12, 1.62]	0.88	[0.61, 1.26]	0.90	[0.64, 1.27]
Grade	1.02	[0.98, 1.07]	1.02	[0.98, 1.07]	1.03	[0.98, 1.08]
Disruptive behavior	11.67***	[9.75, 13.97]	11.82***	[9.86, 14.16]	11.81***	[9.92, 14.05]
Male × Ethnicity			1.80**	[1.15, 2.83]	1.77*	[1.14, 2.73]
Classroom level						
Teacher ethnicity: Black[a]	0.96	[0.76, 1.22]	0.96	[0.81, 1.175]	1.00	[0.74, 1.34]
% children in class with the corresponding ODR	1.12***	[1.11, 1.13]	1.12***	[1.11, 1.14]	1.12***	[1.11, 1.13]
Average disruptive behavior rating for class	0.04***	[0.03, 0.06]	0.04***	[0.03, 0.06]	0.04***	[0.03, 0.06]
Child Ethnicity × Boy × Teacher Ethnicity					0.90	[0.42, 1.91]
Fit statistics						
AIC	15,535.42		15,530.22		15,535.24	
BIC	15,597.08		15,598.74		15,624.32	

Note. N = 6,988 students within 381 classrooms. AOR = adjusted odds ratio; CI = confidence interval; AIC = Akaike information criterion; BIC = Bayesian information criterion.

[a] Indicates dummy code for ethnicity (Black = 1; White = 0). Disruptive behavior data obtained from the teacher-rated Teacher Observation of Classroom Adaptation–Checklist.

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

expected directions. For example, for each 1-point increase in Disruptive Behavior scores, students' odds of receiving an ODR increased by a factor between 4.72 and 11.67 ($p < .001$; see Tables 2–7).

With regard to classroom-level covariates in this first set of models, the classroom-level mean scores on disruptive behavior problems were significantly associated with reduced odds of receiving a teacher-reported ODR or any of the five categories of ODRs from the SWIS. For example, each 1-point decrease in teachers' average Disruptive Behavior ratings was associated with an increase of 8.3 times the odds of receiving any ODR (AOR = .12, $p < .001$; see Tables 2 and 3). The average level of ODRs per classroom also was associated with an increase in the odds of a student receiving an ODR, such that for each increase in the percentage of children per class with the corresponding ODR, there was a 9%–16% increase in their odds of receiving an ODR (AORs ranged from 1.09 to 1.16, $p = .001$). In addition, teacher ethnicity was associated with ODRs, such that having a Black teacher was associated with a 28% increase in the odds of receiving a major ODR (see Table 4) compared to having a White teacher. In contrast, students in classrooms with White teachers had more than twice the odds (AOR = 2.22, $p < .01$) of receiving a minor ODR (see Table 5).

Interactions

We then refit the models described above to test for the following three hypothesized interactions.

Child gender by child ethnicity. We first fit a within-level interaction between the child's ethnicity and the child's gender to explore whether Black boys were more likely to receive ODRs than other children, while controlling for the student- and classroom-level covariates. The analyses indicated a significant interaction between child gender and ethnicity for teacher-reported ODRs (AOR = 1.80, $p < .01$) (see Table 2). Further analyses revealed that although there were no significant differences between Black and White girls, Black boys had 55% greater odds of receiving a teacher-reported ODR compared with White boys (AOR = 1.55, $p < .001$). Surprisingly, none of the other five ODR outcomes were significant for the child gender by child ethnicity interaction term.

Child ethnicity by teacher ethnicity. We then examined whether there was a significant cross-level interaction between child ethnicity and teacher ethnicity for each of the six ODRs. We found that none of the six ODR outcomes were significantly associated with this interaction term, which suggests that Black students were as likely as White children to receive each of the ODRs in classrooms with White teachers as they were in classrooms with Black teachers. Results are not reported due to nonsignificance.

Child ethnicity by child gender by teacher ethnicity. Finally, we tested a series of three-way interactions between child gender and ethnicity with teacher ethnicity and found significant interaction effects for receiving any ODRs (AOR = .58, $p < .05$; see Table 3) and major ODRs (AOR = .51, $p < .05$; see Table 4). The interaction term for the other four ODR outcomes was nonsignificant.

Inspection of the significant three-way interaction for any ODR receipt depicted in Figure 1A suggests that Black boys with Black teachers had the greatest odds of receiving any type of ODR. To further explore this three-way interaction, we stratified the sample

Table 3

Multilevel Analyses Examining the Association Between Student- and Classroom-Level Characteristics and the Odds of Receiving Any Office Discipline Referral (ODR)

	Any SWIS ODRs					
	Main effects model		Child gender × Ethnicity interaction		Three-way interaction	
Variable	AOR	95% CI	AOR	95% CI	AOR	95% CI
Child level						
Male	2.81***	[2.35, 3.36]	2.71***	[2.06, 3.57]	2.68***	[2.07, 3.47]
Ethnicity: Black[a]	1.24**	[1.06, 1.44]	1.18	[0.87, 1.61]	1.16	[0.85, 1.57]
Grade	1.01	[0.98, 1.05]	1.01	[0.98, 1.05]	1.02	[0.98, 1.05]
Disruptive behavior	6.74***	[5.83, 7.80]	6.75***	[5.83, 7.81]	6.77***	[5.85, 7.84]
Male × Ethnicity			1.06	[0.73, 1.55]	1.09	[0.75, 1.58]
Classroom level						
Teacher ethnicity: Black[a]	1.12	[0.96, 1.32]	1.12	[0.95, 1.33]	1.28*	[1.01, 1.61]
% children in class with the corresponding ODR	1.09***	[1.09, 1.10]	1.09***	[1.08, 1.10]	1.09***	[1.09, 1.10]
Average disruptive behavior rating for class	0.12***	[0.09, 0.15]	0.12***	[0.09, 0.15]	0.11***	[0.09, 0.15]
Child Ethnicity × Boy × Teacher Ethnicity					0.58*	[0.34, 0.98]
Fit statistics						
AIC	16,896.63		16,898.50		16,896.22	
BIC	16,958.29		16,967.02		16,971.59	

Note. N = 6,988 students within 381 classrooms. Any ODR includes both major and minor events; examples of major ODRs include referrals for fighting, defiance, lying, cheating; examples of minor ODRs include inappropriate verbal language, physical contact, and property misuse. SWIS = School-wide Information System; AOR = adjusted odds ratio; CI = confidence interval; AIC = Akaike information criterion; BIC = Bayesian information criterion.
[a] Indicates dummy code for ethnicity (Black = 1; White = 0). Disruptive behavior data obtained from the teacher-rated Teacher Observation of Classroom Adaptation–Checklist.
* $p < .05$.   ** $p < .01$.   *** $p < .001$.

(by student gender, student ethnicity, and teacher ethnicity) and conducted a series of post hoc two-way interactions; however, none of the two-way interactions in the stratified samples yielded statistically significant results at $p < .05$.

With regard to the three-way interaction predicting major ODRs, inspection of the data reported in Figure 1B suggested that students in classrooms with Black teachers were more likely to receive a major ODR. Our post hoc stratified analyses revealed a marginally significant two-way interaction between student gender and teacher ethnicity among Black students (AOR = .52, $p = .052$) and a trend for Black girls, such that girls who had Black teachers had marginally greater odds of receiving an ODR than Black girls with White teachers (AOR = 1.63, $p = .066$). Given the relatively large effect sizes (i.e., AORs) observed with the marginally significant p values, we hypothesized that we may have limited power to detect a significant cross-level interaction within the stratified analyses. Therefore, we also tested an interaction between child gender and teacher ethnicity using the full sample and found that regardless of student ethnicity, students in classrooms with Black teachers had greater odds of receiving a major ODR compared with students in classrooms with White teachers (AOR = .57, $p < .05$). Furthermore, the odds of receiving a major ODR were greatest among boys in classrooms with Black teachers (see Figure 1B).

Discussion

The current study aimed to identify factors at the student and classroom levels that may contribute to the overrepresentation of Black students in ODRs. The availability of data from the teachers regarding their perceptions of the child's level of disruptive behavior problems is a unique feature of the current study, for it allowed us to explore ethnic disparities in ODRs while controlling for the child's level of behavior problems. Consistent with prior research, the results of the current study indicate that Black students have significantly greater odds of receiving four of the six ODR types compared to their White peers. In contrast to prior work by Skiba et al. (2008), we did not find that Black students were more likely to receive an ODR for defiance, which is likely a more subjective assessment of the student's behavioral infraction. It is possible that such an effect diminished as a result of controlling for the teacher's ratings of the students' behavior problems. To further explore these associations, additional research is needed with larger, more diverse samples of students who have higher rates of behavior problems.

We also examined specific interactions, including the within-level interaction between student ethnicity and gender. Although we did observe a robust main effect for being male on the odds of receiving all types of ODRs and for being Black on the odds of receiving four of the six types of ODRs, we only observed a significant interaction between being male and student ethnicity for teacher-reported ODRs. It is quite possible that much of this effect was accounted for in the teacher ratings of disruptive behaviors on the TOCA-C.

With regard to the overall trend for the increased odds of ODRs among Black students (for four of the six ODR types), it is important to note that these findings take into account student- and classroom-level characteristics, as well as teachers' behavior ratings of the individual students. The inclusion of

Table 4

Multilevel Analyses Examining the Association Between Student- and Classroom-Level Characteristics and the Odds of Receiving a Major Office Discipline Referral (ODR)

| | SWIS major ODRs | | | | | |
| | Main effects model | | Child gender × Ethnicity interaction | | Three-way interaction | |
Variable	AOR	95% CI	AOR	95% CI	AOR	95% CI
Child level						
Male	3.21***	[2.60, 3.97]	2.97***	[2.12, 4.17]	2.97***	[2.12, 4.16]
Ethnicity: Black[a]	1.15	[0.98, 1.35]	1.05	[0.73, 1.52]	1.02	[0.70, 1.47]
Grade	1.03	[0.99, 1.07]	1.03	[0.99, 1.07]	1.03	[0.99, 1.07]
Disruptive behavior	6.47***	[5.58, 7.50]	6.48***	[5.59, 7.51]	6.51***	[5.62, 7.55]
Male × Ethnicity			1.13	[0.73, 1.76]	1.17	[0.76, 1.82]
Classroom level						
Teacher ethnicity: Black[a]	1.28**	[1.10, 1.50]	1.28**	[1.10, 1.50]	1.52***	[1.25, 1.85]
% children in class with the corresponding ODR	1.10***	[1.09, 1.11]	1.10***	[1.09, 1.11]	1.10***	[1.09, 1.12]
Average disruptive behavior rating for class	0.11***	[0.09, 0.15]	0.11***	[0.09, 0.15]	0.11***	[0.09, 0.15]
Child Ethnicity × Boy × Teacher Ethnicity					0.51**	[0.31, 0.85]
Fit statistics						
AIC	16,281.76		16,604.39		16,279.43	
BIC	16,343.43		16,672.90		16,354.81	

Note. $N = 6,988$ students within 381 classrooms. Examples of major ODRs include referrals for fighting, defiance, lying, cheating. AOR = adjusted odds ratio; CI = confidence interval; SWIS = School-Wide Information System; AIC = Akaike information criterion; BIC = Bayesian information criterion.
[a] Indicates dummy code for ethnicity (Black = 1; White = 0). Disruptive behavior data obtained from the teacher-rated Teacher Observation of Classroom Adaptation–Checklist.
* $p < .05$.   ** $p < .01$.   *** $p < .001$.

the TOCA-C teacher rating data allows us to adjust for the teacher's perception of the level of disruption. Thus, it appears that Black students had greater odds of being referred to the office, even after controlling for the child's level of behavior problems and classroom-level covariates. If the increased odds of ODRs among Black students were the result of higher base rates of behavior problems among Black students, then we would have expected the effect of ethnicity (i.e., Black) to be nonsignificant when the teachers' ratings of the students' level of disruption were controlled for in the analyses. However, Black ethnicity remained a significant main effect. More specifically, these findings suggest that if two students—one Black and one White—had identical ratings on the TOCA-C and all other measures in the study, the Black student had a 24%–80% increase in the odds of receiving an ODR (depending on the type of ODR) compared to his/her White counterpart. The fact that we still observed a significantly higher risk for ODRs among Black students, even after controlling for teachers' perceptions of the level of disruptive behavior and other student and classroom covariates, lends support for the hypothesis that there is a bias against Black students in the use of ODRs (Skiba et al., 2008).

Teacher Ethnicity

We also explored whether the teacher's ethnicity was associated with students' odds of receiving an ODR. The main effects analyses indicated that students in classrooms with Black teachers were more likely to receive a major ODR and less likely to receive a minor ODR than their peers in classrooms with White teachers. Interestingly, the interaction analyses suggested that boys in class-

rooms with Black teachers had greater odds of receiving major ODRs compared to the other students. However, for the other four ODR types, students in classrooms with Black teachers had similar odds as their peers of receiving an ODR. It is important to bear in mind that all of these analyses controlled for student- and classroom-level covariates, including teachers' ratings of students' behavior problems, as well as classroom-level behavior problems. Additional research is needed to understand why students in classrooms with Black teachers were more likely to receive major ODRs (especially among boys), but less likely to receive minor ODRs.

To further explore whether the influence of student ethnicity on the odds of receiving an ODR varied by teacher ethnicity, we examined the interaction between teacher ethnicity and student ethnicity. Despite increased interest in promoting cultural continuity between students and teachers (APA, 2002; Ladson-Billings, 1997, 2002; Murray & Murray, 2004; Tyler et al., 2008), we found no significant interactions between student and teacher ethnicity in the odds of receiving an ODR. However, when we incorporated the three-way interaction term (Student Ethnicity × Student Gender × Teacher Ethnicity), the results suggested that Black male students in classrooms with Black teachers had greater odds of receiving any ODRs than the other students. Results from the significant three-way interaction in the major ODRs model suggested that students in classrooms with Black teachers had greater odds of receiving a major ODR than students with White teachers, and these effects were greatest among boys, regardless of student ethnicity.

Although we interpret these two three-way interactions with caution, it is possible that the Black teachers may have a higher

Table 5

Multilevel Analyses Examining the Association Between Student- and Classroom-Level Characteristics and the Odds of Receiving a Minor Office Discipline Referral (ODR)

| | SWIS minor ODRs | | | | | |
| | Main effects model | | Child gender × Ethnicity interaction | | Three-way interaction | |
Variable	AOR	95% CI	AOR	95% CI	AOR	95% CI
Child level						
Male	2.16***	[1.59, 2.92]	2.24***	[1.41, 3.54]	2.23***	[1.41, 3.52]
Ethnicity: Black[a]	1.82***	[1.34, 2.48]	1.91**	[1.19, 3.06]	1.89**	[1.18, 3.03]
Grade	0.93	[0.86, 1.02]	0.94	[0.86, 1.02]	0.94	[0.86, 1.02]
Disruptive behavior	4.72***	[4.00, 5.57]	4.72***	[4.00, 5.57]	4.72***	[4.00, 5.57]
Male × Ethnicity			0.94	[0.51, 1.71]	0.92	[0.51, 1.69]
Classroom level						
Teacher ethnicity: Black[a]	0.45**	[0.26, 0.78]	0.45**	[0.26, 0.78]	0.50	[0.25, 1.01]
% children in class with the corresponding ODR	1.14***	[1.13, 1.16]	1.14***	[1.13, 1.16]	1.14***	[1.13, 1.16]
Average disruptive behavior rating for class	0.11***	[0.07, 0.18]	0.11***	[0.07, 0.18]	0.11***	[0.07, 0.18]
Child Ethnicity × Boy × Teacher Ethnicity					0.67	[0.18, 2.56]
Fit statistics						
AIC	14,594.47		14,596.38		14,597.89	
BIC	14,656.13		14,664.90		14,673.26	

Note. N = 6,988 students within 381 classrooms. Examples of Minor ODRs include inappropriate verbal language, physical contact, and property misuse. SWIS = School-Wide Information System; AOR = adjusted odds ratio; CI = confidence interval; AIC = Akaike information criterion; BIC = Bayesian information criterion.
[a] Indicates dummy code for ethnicity (Black = 1; White = 0). Disruptive behavior data obtained from the teacher-rated Teacher Observation of Classroom Adaptation–Checklist.
* p < .05. ** p < .01. *** p < .001.

standard for the behavior of their students and thus may be more inclined to use major ODRs with their students than other teachers. Support for this interpretation comes from research on parenting within Black communities, whereby Black parents may be inclined to use more authoritarian parenting styles with their children—especially with boys—than typically observed among White parents (Portes, Dunham, & Williams, 1986). It is believed that these slightly more authoritarian parenting practices displayed by the Black parents may prepare their children for the challenges they may face in a historically racist society (Bradley, 1998). Again, we are cautious about the interpretation of these three-way interactions, especially because they only occurred for two of the six ODR types. Furthermore, there was just one significant main effect of teacher ethnicity on the receipt of a major ODR, but a reduced risk of receiving a minor ODR among students in classrooms with Black teachers.

It is important to note, however, that although the students and teachers in the current sample may both be Black, there may be other cultural, contextual, or economic factors that could contribute to the use of ODRs across ethnic groups. As such, ethnicity may represent a poor proxy for cultural match between students and their teachers (APA, 2002). Furthermore, teacher factors not assessed in this study, such as teachers' perceptions of acceptable classroom behaviors, their attitudes regarding the utility of documenting minor versus major ODRs, and their skills at handling disruptive behaviors may contribute to these differences by teacher ethnicity. Additional research is needed to further explore the issue of cultural match between students and teachers (see Tyler et al., 2008) and the potential influence of other teacher factors not assessed in this study.

Nevertheless, these findings do not suggest that a cultural or ethnic match between students and their teachers reduces the risk for ODRs among Black students.

Classroom Influences

Finally, we also explored select classroom-level factors on the odds of receiving an ODR. Surprisingly, we found that higher classroom-level ratings on disruptive behavior problems were associated with a reduction in the odds that a student would receive most types of ODRs. It is unclear why at the individual level, the TOCA-C ratings were associated with an increased risk, but at the classroom level they tended to have a suppression effect on the odds of receiving an ODR. A recent study on suspensions observed a similar effect, whereby in classrooms with higher levels of teacher-rated disruptive behavior, students were less likely to be suspended (Petras, Masyn, Buckley, Ialongo, & Kellam, 2009). It is possible that in more disruptive and chaotic classrooms, it may be more difficult for teachers to detect specific students who are displaying behavior problems. Furthermore, the overall level of disruption occurring in the classroom may raise the teacher's threshold for behavior warranting an ODR. Although this finding requires further investigation, these effects draw our attention to classroom contextual factors when considering a teacher's use of different management and disciplinary practices (Koth et al., 2008).

Limitations

It is important to consider some limitations when interpreting these findings. We focused exclusively on Black and White stu-

Table 6

Multilevel Analyses Examining the Association Between Student- and Classroom-Level Characteristics and the Odds of Receiving an Office Discipline Referral (ODR) for Fighting

| | SWIS ODR for fighting | | | | | |
| | Main effects model | | Child gender × Ethnicity interaction | | Three-way interaction | |
Variable	AOR	95% CI	AOR	95% CI	AOR	95% CI
Child level						
Male	3.69***	[2.87, 4.76]	3.04***	[2.00, 4.64]	3.03***	[1.99, 4.62]
Ethnicity: Black[a]	1.26*	[1.04, 1.53]	1.00	[0.62, 1.62]	0.99	[0.61, 1.60]
Grade	1.06*	[1.01, 1.12]	1.06*	[1.01, 1.12]	1.07*	[1.01, 1.12]
Disruptive behavior	4.78***	[4.13, 5.53]	4.79***	[4.14, 5.54]	4.80***	[4.15, 5.55]
Male × Ethnicity			1.34	[0.77, 2.32]	1.35	[0.78, 2.35]
Classroom level						
Teacher Ethnicity: Black[a]	1.23	[0.97, 1.58]	1.23	[.97, 1.57]	1.41*	[1.01, 1.96]
% children in class with the corresponding ODR	1.13***	[1.10, 1.15]	1.13***	[1.10, 1.15]	1.13***	[1.10, 1.15]
Average disruptive behavior rating for class	0.16***	[0.11, 0.22]	0.16***	[0.11, 0.22]	0.16***	[0.11, 0.22]
Child Ethnicity × Boy × Teacher Ethnicity					0.68	[0.36, 1.28]
Fit statistics						
AIC	15,280.94		15,281.84		15,282.62	
BIC	15,342.61		15,350.36		15,357.99	

Note. $N = 6,988$ students within 381 classrooms. SWIS = School-wide Information System; AOR = adjusted odds ratio; CI = confidence interval; AIC = Akaike information criterion; BIC = Bayesian information criterion.
[a] Indicates dummy code for ethnicity (Black = 1; White = 0). Disruptive behavior data obtained from the teacher-rated Teacher Observation of Classroom Adaptation–Checklist.
* $p < .05.$ ** $p < .01.$ *** $p < .001.$

dents and teachers, as there were too few Hispanic/Latino, Asian American, or Native American participants to examine these groups separately. Additional research is needed in larger and more diverse samples to better understand the pattern of findings for student and staff participants of other ethnicities. As noted above, we were only able to examine student and teacher ethnicity, which may be a poor proxy for cultural factors (Tyler et al., 2008). This may contribute to the current findings that suggest that continuity between teacher and student ethnicity does not reduce ODRs. In addition, other cultural, contextual, or economic factors not assessed in this study could also be contributing to this potential disconnect. Furthermore, data are not available on socioeconomic status or the community in which the students and teachers live. Information on these factors may inform our understanding of the extent to which the students and teachers have similar cultural and social-economic backgrounds. Moreover, culturally relevant factors such as communalism, verve, and movement among students and teachers may be more salient when examining school discipline within the Black community (Tyler et al., 2006; Tyler et al., 2008).

Whereas much of the previous research on the overrepresentation of ethnic minorities in ODRs has typically only reported one source of data regarding discipline problems (e.g., Skiba et al., 2008), a strength of this study is the availability of multiple sources of data (ODRs from teachers and SWIS, in addition to teachers' TOCA-C ratings). Nevertheless, several of the study variables were assessed via teacher-report, which may be associated with unmeasured teacher characteristics. Future studies may consider obtaining data from other sources, such as peers or outside observers, to reduce a reliance on teacher-reported measures. In

addition, other student-level factors (e.g., academic performance) and/or teacher-level factors (e.g., personality, skills to handle discipline problems) not assessed in the current study may have influenced the outcomes. However, a recent multilevel study of elementary school teachers found that neither teacher efficacy nor burnout was associated with ODRs, after controlling for the other student-, teacher-, and school-level variables (Pas, Bradshaw, Hershfeldt, & Leaf, 2009). Further exploration of these and other student, teacher, and school contextual factors may help us begin to understand why Black students continue to be at increased risk for ODRs.

The current study only included elementary school children, who on average had a relatively low base-rate of ODRs (Bradshaw et al., 2009). Future studies should examine whether similar effects occur in secondary schools, where the rates of disruptive behavior problems and ODRs tend to be higher. Because these data are cross-sectional, we are unable to examine previous factors, such as a history of ODRs on the students' current risk for an ODR. It is possible that students with a prior history of ODRs may develop a "reputation" for being disruptive, which could influence the teachers' likelihood of referring the student. Additional longitudinal research should be conducted to further explore these associations over time.

We also considered modeling school-level factors as a third level of influence; however, when we explored inclusion of school-level factors (e.g., school size, student mobility, faculty turnover), we did not find any significant effects for these factors. This is likely a result of the relatively small number of schools ($n = 21$). In fact, most researchers caution against using multilevel modeling with 20 or fewer clusters, as the robust

Table 7

Multilevel Analyses Examining the Association Between Student- and Classroom-Level Characteristics and the Odds of Receiving an Office Discipline Referral [ODR] for Defiance

	SWIS ODR for defiance					
	Main effects model		Child gender × Ethnicity interaction		Three-way interaction	
Variable	AOR	95% CI	AOR	95% CI	AOR	95% CI
Child level						
Male	2.88***	[2.00, 4.14]	3.17***	[1.64, 6.14]	3.16***	[1.63, 6.10]
Ethnicity: Black[a]	1.09	[0.84, 1.40]	1.22	[0.61, 2.45]	1.18	[0.60, 2.35]
Grade	1.06	[0.99, 1.12]	1.06	[0.99, 1.12]	1.05	[0.99, 1.12]
Disruptive behavior	5.17***	[4.48, 5.96]	5.17***	[4.48, 5.96]	5.18***	[4.49, 5.98]
Male × Ethnicity			0.87	[0.39, 1.91]	0.89	[0.41, 1.93]
Classroom level						
Teacher ethnicity: Black[a]	1.22	[0.90, 1.66]	1.22	[0.90, 1.66]	1.44	[0.94, 2.22]
% of children per class with the corresponding ODR	1.16***	[1.12, 1.21]	1.16***	[1.12, 1.21]	1.16***	[1.12, 1.21]
Average disruptive behavior rating for class	0.17***	[0.12, 0.26]	0.17***	[0.12, 0.26]	0.17***	[0.12, 0.26]
Child Ethnicity × Boy × Teacher Ethnicity					0.62	[0.25, 1.53]
Fit statistics						
AIC	14,533.59		14,535.41		14,536.22	
BIC	14,595.26		14,603.93		14,611.59	

Note. $N = 6,988$ students within 381 classrooms. SWIS = School-Wide Information System; AOR = adjusted odds ratio; CI = confidence interval; AIC = Akaike information criterion; BIC = Bayesian information criterion.
[a] Indicates dummy code for ethnicity (Black = 1; White = 0). Disruptive behavior data obtained from the teacher-rated Teacher Observation of Classroom Adaptation–Checklist.
* $p < .05$. ** $p < .01$. *** $p < .001$.

standard errors may be biased and there is limited power to detect small effects (Raudenbush & Bryk, 2002; Raudenbush et al., 2004). Furthermore, the pattern of findings for the student- and classroom-level covariates did not change substantially in the exploratory three-level models (i.e., those which included school-level covariates). Therefore, we questioned the extent to which this additional level of complexity would contribute to our interpretation of the findings (Raudenbush & Bryk, 2002). Future research with a larger sample of schools should explore characteristics at the school-level that may influence the risk for ODRs.

Implications for Research and Practice

Although a detailed discussion of the validity of ODRs is beyond the scope of the current study (for a review see Irvin et al., 2004; Morrison et al., 2006), ODRs are likely influenced by multiple factors including the student's behavior, the classroom and/or school rules, and teacher factors (efficacy, skills, perceptions of the behavior and/or student) (Morrison et al., 2006). As such, ODRs may not be truly objective indicators of student behavior problems, but rather an indicator of the teachers' use of removal from the classroom as a disciplinary strategy. Regardless of the validity of ODRs, they are a significant concern because they are indicators of missed class time for the student (estimated to be between 20 and 40 min; Scott & Barrett, 2004), which in itself may portend poor educational outcomes for the student. Furthermore, students who go on to receive multiple ODRs may be at risk for developing a reputation among other teachers or possibly among students for displaying problem behavior, which in turn

could increase their risk for subsequent discipline problems (Irvin et al., 2004).

Given the accumulating evidence of the overrepresentation of Black students among ODRs and other indicators of academic problems, additional research is needed to better understand why teachers make ODRs and why Black students are at greater risk for receiving them. The discrepancies in disciplinary practices observed in the current study are likely perceived by the students as biased and may lead to negative student-teacher interactions as well as a diminished sense of school climate (Koth et al., 2008). This in turn could increase the likelihood of attendance problems, lower levels of academic success, and increase risk for placement in special education programs and school dropout (DeRidder, 1990; Patterson et al., 1989; Wald & Losen, 2007).

Professional development should be provided to increase staff awareness of the overrepresentation of Black students in ODR data and to help staff identify potential factors within their schools and classrooms that may contribute to such discrepancies. Although much of the extant research on ethnic differences in ODR data has focused on increasing awareness of these discrepancies (Skiba et al., 2008), few researchers have developed skill-based strategies for addressing this issue. Research by Towns, Cole-Henderson, and Serpell (2001) suggests that aspects of school climate, such as strong administrative leadership, the communication of high expectations for achievement, and an emphasis on academic skill development and teamwork contribute to student success in urban schools. Therefore, providing training for teachers and staff on culturally relevant and respectful discipline practices should be a
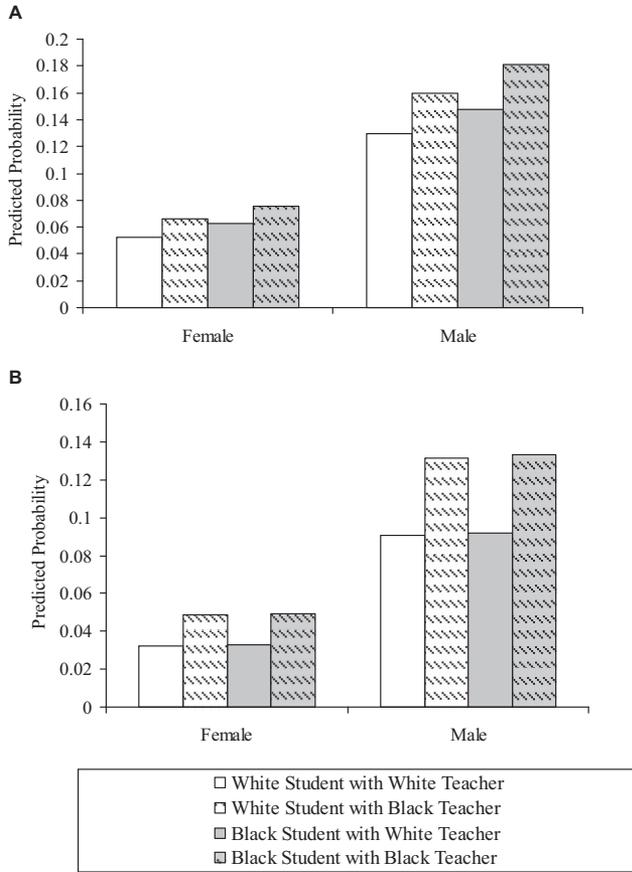
Figure 1. A: Interaction between gender, student ethnicity, and teacher ethnicity in predicting any office disciplinary referrals (ODRs). B: Interaction between gender, student ethnicity, and teacher ethnicity in predicting major ODRs.

core component of annual professional development (Ladson-Billings, 1997, 2002). Additional empirical work is also needed to develop evidence-based skills-focused programs to promote cultural competency among teachers in order to reduce disproportionality in ODRs and provide increased opportunities for student learning.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19,* 716–723.

American Psychological Association. (2002). *Guidelines on multicultural education, training, research, practice, and organizational change for psychologists.* Washington, DC: Author.

American Psychological Association Task Force on Resilience and Strength in Black Children and Adolescents. (2008). *Resilience in African American children and adolescents: A vision for optimal development.* Washington, DC: Author.

Atkins, M. S., McKay, M. M., Frazier, S. L., Jakobsons, L. J., Arvanitis, P., Cunningham, T., . . . Lambrecht, L. (2002). Suspensions and detentions in an urban, low-income school: Punishment or reward. *Journal of Abnormal Child Psychology, 30,* 361–371.

Boykin, A. W., Lilja, A. J., & Tyler, K. M. (2004). The influence of communal versus individual learning context on the academic perfor-

mance of African-American elementary school students. *Learning Environments Journal, 7,* 227–244.

Bradley, C. R. (1998). Child rearing in African American families: A study of the disciplinary practices of African American parents. *Journal of Multicultural Counseling and Development, 26,* 273–281.

Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2009). Examining the effects of School-Wide Positive Behavioral Interventions and Supports on student outcomes: Results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions.* Advance online publication. doi:10.1177/1098300709334798

DeRidder, L. M. (1990). How suspension and expulsion contribute to dropping out. *Education Digest, 56,* 44–47.

DeVoe, J. F., Peter, K., Kaufman, P., Miller, A. K., Noonan, M., Snyder, T. D., & Baum, K. (2004). *Indicators of school crime and safety: 2004* (NCES Report No. 2005–002/NCJ 205290). Washington, DC: Government Printing Office.

Donovan, S. M., & Cross, C. T. (2002). *Minority students in special and gifted education.* Washington, DC: National Academic Press.

Eitle, D. J., & Eitle, T. M. (2004). School and county characteristics as predictors of school rates of drug, alcohol, and tobacco offenses. *Journal of Health and Social Behavior, 45,* 408–421.

Epstein, J. N., Willoughby, M., Valencia, E. Y., Tonev, S. T., Abikoff, H. B., Arnold, L., & Hinshaw, S. P. (2005). The role of children's ethnicity in the relationship between teacher ratings of attention-deficit/hyperactivity disorder and observed classroom behavior. *Journal of Consulting and Clinical Psychology, 73,* 424–434.

Ford, D. Y. (1993). An investigation of the paradox of underachievement among gifted Black students. *Roeper Review, 16,* 78–84.

Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology, 95,* 148–162.

Gottfredson, D. (2001). *Schools and delinquency.* New York, NY: Cambridge University Press.

Gottlieb, J., Gottlieb, B. W., & Trongue, S. (1991). Parent and teacher referrals for a psychoeducational evaluation. *Journal of Special Education, 25,* 453–465.

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

Hawkins, J. D., Doueck, H. J., & Lishner, D. M. (1988). Changing teaching practices in mainstream classrooms to improve bonding and behavior of low achievers. *American Educational Research Journal, 25,* 31–50.

Ialongo, N., Poduska, J., Werthamer, L., & Kellam, S. (2001). The distal impact of two first-grade preventive interventions on conduct problems and disorder in early adolescence. *Journal of Emotional and Behavioral Disorders, 9,* 146–160.

Ialongo, N. S., Werthamer, L., Kellam, S. G., Brown, C. H., Wang, S., & Lin, Y. (1999). Proximal impact of two first-grade preventive interventions on the early risk behaviors for later substance abuse, depression, and antisocial behavior. *American Journal of Community Psychology, 27,* 599–642.

Irvin, L. K., Horner, R. H., Ingram, K., Todd, A. W., Sugai, G., Sampson, N. K., & Boland, J. B. (2006). Using office discipline referral data for decision making about student behavior in elementary and middle schools: An empirical evaluation of validity. *Journal of Positive Behavior Interventions, 8,* 10–23.

Irvin, L. K., Tobin, T. J., Sprague, J. R., Sugai, G., & Vincent, C. G. (2004). Validity of office discipline referral measures as indices of school-wide behavioral status and effects of school-wide behavioral interventions. *Journal of Positive Behavior Interventions, 6,* 3–12.

Koth, C. W., Bradshaw, C. P., & Leaf, P. J. (2008). Examining the relationship between classroom-level factors and students' perception of school climate. *Journal of Educational Psychology, 100,* 96–104.

Koth, C. W., Bradshaw, C. P., & Leaf, P. J. (2009). Teacher Observation

of Classroom Adaptation–Checklist (TOCA-C): Development and factor structure. *Measurement and Evaluation in Counseling and Development, 42,* 15–30.

Ladson-Billings, G. (1997). *The dreamkeepers: Successful teachers of African American children.* New York, NY: Jossey-Bass.

Ladson-Billings, G. (2002). But that's just good teaching! The case for culturally relevant pedagogy. In S. J. Denbo & L. M. Beaulieu (Eds.), *Improving schools for African American students: A reader for educational leaders* (pp. 95–102). Springfield, IL: Thomas.

May, S., Ard, W., Todd, A., Horner, R., Glasgow, A., Sugai, G., & Sprague, J. R. (2003). Schoolwide information system [Computer software]. Eugene, OR: University of Oregon, Educational and Community Supports.

Monroe, C. R. (2005). Why are "bad boys" always Black? Causes for disproportionality in school discipline and recommendations for change. *The Clearing House, 79,* 45–50.

Morrison, G. M., Redding, G. M., Fisher, E., & Peterson, R. (2006). Assessing school discipline. In S. R. Jimerson & M. J. Furlong (Eds.), *Handbook of school violence and school safety: From research to practice* (pp. 221–220). Mahwah, NJ: Erlbaum.

Morrison, G. M., & Skiba, R. (2001). Predicting violence from school misbehavior: Promises and perils. *Psychology in the Schools, 38,* 173–184.

Murray, C., & Murray, K. M. (2004). Child level correlates of teacher–student relationships: An examination of demographic characteristics, academic orientation, and behavioral orientations. *Psychology in the Schools, 41,* 751–762.

National Center for Educational Statistics. (2004a). *The nation's report card: Mathematics highlights 2003* (NCES Publication No. 2004-451). Jessup, MD: U.S. Department of Education.

National Center for Educational Statistics. (2004b). *The nation's report card: Reading highlights 2003* (NCES Publication No. 2004-452). Jessup, MD: U.S. Department of Education.

Neal, L. I., McCray, A. D., Webb-Johnson, G., & Bridgest, S. T. (2003). The effects of African American movement styles on teachers' perceptions and reactions. *Journal of Special Education, 37,* 49–57.

Nelson, J. R., & Roberts, M. L. (2000). Ongoing reciprocal teacher–student interactions involving disruptive behaviors in general education classrooms. *Journal of Emotional and Behavioral Disorders, 8,* 27–37.

Pas, E. T., Bradshaw, C. P., Hershfeldt, P. A., & Leaf, P. J. (2009). *A multilevel exploration of the influence of teacher efficacy and burnout on response to student problem behavior and school-based service use.* Manuscript submitted for publication.

Patterson, G. R., DeBaryshe, B. D., & Ramsey, E. (1989). A developmental perspective on antisocial behavior. *American Psychologist, 44,* 329–335.

Petras, H., Chilcoat, H. D., Leaf, P. J., Ialongo, N. S., & Kellam, S. G. (2004). Utility of TOCA-R scores during the elementary school years in identifying later violence among adolescent males. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*(1), 88–96.

Petras, H., Masyn, K. A., Buckley, J. A., Ialongo, N. S., & Kellam, S. (2009). *Who is most at risk for school removal? A multilevel discrete-time survival analysis of individual and contextual-level influences.* Manuscript submitted for publication.

Portes, P. R., Dunham, R. M., & Williams, S. (1986). Assessing child-rearing style in ecological settings: Its relation to culture, social class, early age intervention and scholastic achievement. *Adolescence, 21,* 723–735.

Raffaele Mendez, L. M., & Knoff, H. M. (2003). Who gets suspended from school and why: A demographic analysis of schools and disciplinary infractions in a large school district. *Psychology in the Schools, 39,* 259–277.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models:*

Applications and data analysis methods (2nd ed.). Thousand Oaks, CA: Sage.

Raudenbush, S. W., Bryk, A., Cheong, Y.-F., Congdon, R., & du Toit, M. (2004). HLM statistical software (Version 6) [Computer software]. Lincolnwood, IL: Scientific Software International.

Roeser, R. W., Midgley, C. M., & Urdan, T. C. (1996). Perceptions of the school psychological environment and early adolescents' psychological and behavioral functioning in school: The mediating role of goals and belonging. *Journal of Educational Psychology, 88,* 408–422.

Sbarra, D. A., & Pianta, R. C. (2001). Teacher ratings of behavior among African American and Caucasian children during the first two years of school. *Psychology in the Schools, 38,* 229–238.

Schaeffer, C. M., Petras, H., Ialongo, N., Masyn, K. E., Hubbard, S., Poduska, J., & Kellam, S. (2006). A comparison of girls' and boys' aggressive-disruptive behavior trajectories across elementary school: Prediction to young adult antisocial outcomes. *Journal of Consulting and Clinical Psychology, 74,* 500–510.

Schaeffer, C. M., Petras, H., Ialongo, N., Poduska, J., & Kellam, S. (2003). Modeling growth in boys' aggressive behavior across elementary school: Links to later criminal involvement, conduct disorder, and antisocial personality disorder. *Developmental Psychology, 6,* 1020–1035.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6,* 461–464.

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,* 21–27.

Sheets, R. H. (1996). Urban classroom conflict: Student–teacher perception: Ethnic integrity, solidarity, and resistance. *Urban Review, 28,* 165–183.

Skiba, R., Horner, R., Chung, C. G., Rausch, M. K., May, S. L., & Tobin, T. (2008, March). *Race is not neutral: A national investigation of African American and Latino disproportionality in school discipline.* Paper presented at the annual meeting of the American Educational Research Association, New York, NY.

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

Skiba, R. J., & Peterson, R. (1999). Zero-tolerance: Can punishment lead to safe schools? *Phi Delta Kappan, 80,* 372–376.

Sprague, J., Walker, H. M., Stieber, S., Simonsen, B., Nishioka, V., & Wagner, L. (2001). Exploring the relationship between school discipline referrals and delinquency. *Psychology in the Schools, 38,* 197–206.

Sugai, G., Sprague, J. R., Horner, R. H., & Walker, H. M. (2000). Preventing school violence: The use of office discipline referrals to assess and monitor school-wide discipline interventions. *Journal of Emotional and Behavioral Disorders, 8,* 94–101.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston, MA: Allyn & Bacon.

Tobin, T. J., & Sugai, G. M. (1999). Using sixth-grade school records to predict school violence, chronic discipline problems, and high school outcomes. *Journal of Emotional and Behavioral Disorders, 7,* 40–53.

Towns, D. P., Cole-Henderson, B., & Serpell, Z. (2001). The journey to urban school success: Going the extra mile. *Journal of Negro Education, 70,* 4–18.

Townsend, B. L. (2000). The disproportionate discipline of African American learners: Reducing school suspensions and expulsions. *Exceptional Children, 66,* 381–391.

Tyler, K. M., Boykin, A. W., Miller, O. A., & Hurley, E. A. (2006). Cultural values in the home and school experiences of low-income African American students. *Social Psychology of Education, 9,* 363–380.

Tyler, K. M., Uqdah, A. L., Dillihunt, M. L., Beatty-Hazelbaker, R., Conner, T., Gadson, N., . . . Stevens, R. (2008). Cultural discontinuity:

Toward a quantitative investigation of a major hypothesis in education. *Educational Researcher, 37,* 280–297.

Wald, J., & Losen, D. J. (2003). Defining and redirecting a school-to-prison pipeline. In J. Wald & D. J. Losen (Eds.), *New directions for youth development: 99. Deconstructing the school-to-prison pipeline* (pp. 9–15). San Francisco, CA: Jossey-Bass.

Wald, J., & Losen, D. J. (2007). Out of sight: The journey through the school-to-prison pipeline. In S. Books (Ed.), *Invisible children in the society and its schools* (3rd ed., pp. 23–27). Mahwah, NJ: Erlbaum.

Weinstein, C., Curran, M., & Tomlinson-Clarke, S. (2004). Toward a conception of culturally responsive classroom management. *Journal of Teacher Education, 55,* 25–38.

Werthamer-Larsson, L., Kellam, S., & Wheeler, L. (1991). Effect of first-grade classroom environment on shy behavior, aggressive behavior, and concentration problems. *American Journal of Community Psychology, 19,* 585–602.

Wu, S., Pink, W., Carin, R., & Moles, O. (1982). Student suspension: A critical reappraisal. *Urban Review, 14,* 245–303.