Vulnerable Decision Points for Disproportionate Office Discipline Referrals: Comparisons of Discipline for African American and White Elementary School Students

Keith Smolkowski
Oregon Research Institute

Erik J. Girvan, Kent McIntosh, Rhonda N. T. Nese, and Robert H. Horner
University of Oregon

ABSTRACT: Racial disparities in rates of exclusionary school discipline are well documented and seemingly intractable. However, emerging theories on implicit bias show promise in identifying effective interventions. In this study, we used school discipline data from 1,666 elementary schools and 483,686 office discipline referrals to identify specific situations in which disproportionality was more likely. Results were largely consistent with our theoretical model, indicating increased racial and gender disproportionality for subjectively defined behaviors, in classrooms, and for incidents classified as more severe. The time of day also substantially affected disproportionality. These findings can be used to pinpoint specific student–teacher interactions for intervention.

In the United States, racial disparities in rates of exclusionary discipline for students of color have been well documented, with differences most pronounced for African American students in particular (Losen & Gillespie, 2012; Losen, Hodson, Keith, Morrison, & Belway, 2015). For example, in the 2011–2012 school year, nationally, administrators used out-of-school suspensions to discipline 8% of African American elementary students and 23% of African American secondary students, compared to 2% of White elementary students and 7% of White secondary students (Losen et al., 2015). In addition, converging research provides evidence that disproportionality cannot be wholly attributed to structural factors associated with students or schools. Even controlling for poverty, participation in gifted-and-talented programs, student–teacher ratio, attendance rates, and other factors, African American students continue to be disciplined at higher rates than White students (Anyon et al., 2014; Fabelo et al., 2011; Skiba, Poloni-Staudinger, Simmons, Feggins, & Chung, 2005; Wallace, Goodkind, Wallace, & Bachman, 2008). As a result, although disparate treatment of any students by race is concerning, the disparities for African American students are most severe. Moreover, there is no evidence that disproportionality results from differences in levels of student behavior by race. To the contrary, research has shown that teachers are more likely to issue office discipline referrals (ODRs) to African American students even after controlling for their own ratings of the students' behaviors (Bradshaw, Mitchell, O'Brennan, & Leaf, 2010).

This research base provides a clear description of the extent of disproportionality in school discipline. By comparison, there is much less empirical research on interventions to reduce disproportionality, or even what variables should be targeted for intervention (Martinez, 2013; Staats, 2014). Thus, to lay a crucial foundation for addressing disproportionality, it is necessary to focus on developing and validating a theoretical framework that explains when and why disproportionality is most likely to occur and, more importantly, identifies malleable variables that can be used to reduce it.

A Model for Explaining Disproportionality in School Discipline

To address this need, McIntosh, Girvan, Horner, and Smolkowski (2014) proposed the Vulnerable Decision Points model. This model draws on psychological research to describe the conditions under which racial bias is most likely to influence decisions in the school

discipline context and highlights specific avenues for intervention. As such, it focuses on teacher and administrator perceptions and judgments within specific discipline decisions. Although explaining the entire model is beyond the scope of this article, two critical aspects of it are described here: (a) explicit versus implicit bias and (b) vulnerable decision points in school discipline.

Explicit Versus Implicit Bias

A large body of psychological research suggests that there are two distinct types of bias, explicit and implicit, which operate differently and can influence different types of decisions (Girvan, 2015; Girvan, Deason, & Borgida, 2015). This distinction is particularly important because interventions that have been shown to reduce the effects of explicit bias are not necessarily effective for implicit bias and vice versa (Lai, Hoffman, Nosek, & Greenwald, 2013; Pettigrew & Tropp, 2006). Explicit bias is what we typically think of as prejudice: ethnocentrism, racism, and other consciously endorsed attitudes or beliefs, such as the belief that African Americans are inherently criminal or lazy (Pearson, Dovidio, & Gaertner, 2009). By comparison, implicit bias is the automatic, often unconscious impact that stereotypic associations with racial and other groups can have on perceptions, judgments, decision-making, and behavior (Devine, 1989; Greenwald & Banaji, 1995; Pearson et al., 2009). Rather than conscious endorsement of beliefs or feelings, it has its roots in generalized associations formed from systematically repetitious or unique and limited experience or exposure. Thus, for example, regularly seeing images of African American but not White criminal offenders in the media or knowing only one person who was the victim of a carjacking, an incident in which the perpetrator happened to be African American, may lead even people with egalitarian values to automatically assume that a racially unidentified gang member is African American, presume that an area in which there are many African Americans living must have a crime problem, or lock their car door when seeing an African American man (Greenwald & Banaji, 1995). From a behavioral standpoint, implicit bias can be conceptualized as inappropriate stimulus control over an individual's responses to others' behavior that is based on irrelevant features of the behavior, as opposed to an objective view of the behavior. In school settings, implicit bias may be seen in staff decisions to send students of color to the office for relatively minor incidents of unwanted behavior.

An individual's levels of explicit and implicit biases are relatively independent of one another (Greenwald, Poehlman, Uhlmann, & Banaji, 2009). As a result, several combinations are possible: Those without explicit or implicit bias (the "truly nonprejudiced"), those with both explicit and implicit bias (the "truly prejudiced"), and those whose explicit and implicit biases are not aligned (Fazio, Jackson, Dunton, & Williams, 1995). Of the combinations, given evidence of both a substantial decline in explicit racial bias and the relative pervasiveness of implicit racial bias (Bobo, Charles, Krysan, & Simmons, 2012; Greenwald & Pettigrew, 2014; Nosek, Greenwald, & Banaji, 2007), the majority of U.S. adults are likely to express no explicit racial bias (i.e., have and report a belief in the value of diversity, equity, and inclusion in society) but have implicit racial biases favoring Whites over African Americans, a combination known as aversive racism (Dovidio & Gaertner, 2000; Pearson et al., 2009).

Under aversive racism theory, people are assumed to be highly motivated not to be, or appear not to be, racially biased. As such, when confronted with decisions in which incorrect responses are clear or particular responses would be seen as discriminatory, they will most likely select correct or nondiscriminatory responses. But in discretionary decisions and those with an unclear "right answer," decisions do not directly implicate or threaten decision-makers' egalitarian values or self-concepts. In those circumstances, values notwithstanding, the majority of people may act in ways that are discriminatory (Pearson et al., 2009). Research across a range of contexts outside of education supports these predictions. For example, White people are generally willing to offer help at the same rates to African Americans as to Whites, except when there are difficult circumstances, such as when helping is risky or time consuming, or in the presence of other factors that can be used to justify the failure to help (Saucier, Miller, & Doucet, 2005); people tend to discriminate in hiring recommendations against moderately but not highly or poorly qualified candidates (Dovidio & Gaertner, 2000); and incriminating evidence that is found to be inadmissible tends to influence jury decisions about

African American but not White defendants (Johnson, Whitestone, Jackson, & Gatto, 1995).

Translating this work into the school discipline context, disproportionality may come from an individual educator's explicit biases, but we posit it is more likely that it comes from implicit biases. School discipline data patterns can help identify which is more at play. If explicit bias is prominent, school discipline data would likely represent a consistent pattern of disproportionality across many circumstances. For example, analysis of discipline data might demonstrate that African American students are sent out of class regularly for behavior incidents, regardless of the situation. From the literature, effective top-down policies, such as evaluating administrators and teachers based on levels of disproportionality, are more likely to mitigate the effects of explicit bias (Lerner & Tetlock, 1999; Pettigrew & Tropp, 2006).

In contrast, an indicator of implicit bias in school discipline data would reflect peaks and valleys in disproportionality from the same teachers across different situations, with relative equity in some situations and high disproportionality in others, as predicted by psychological theory. The data might demonstrate, for example, that African American students receive disproportionately more ODRs for defiance or disrespect than White students because identifying these behaviors involves a discretionary decision for teachers (e.g., whether student behavior is acceptable or unacceptable to the teacher). Or it might show that consequences for the same behavior are more severe for African American students during times of the day when teachers are tired. The model indicates that effects of implicit bias can be reduced by making discipline procedures for these types of behaviors as objective as possible, and by examining staff expectations and providing training in how to respond instructionally to unexpected student behavior in these specific situations, without resorting to an ODR.

Vulnerable Decision Points in School Discipline

We use the term *vulnerable decision points* (VDPs) to describe specific situations in which increased disproportionality tends to occur. VDPs are contextual events or elements, such as those that increase the likelihood of implicit bias affecting discipline decision making, including a teacher's decision to issue an ODR or an administrator's decision to suspend the student.

Emerging research indicates the presence of some VDPs in education. The VDP with the strongest research support is a situation in which the student behavior is inherently subjective (i.e., when staff have to make a judgment call regarding whether the behavior is a violation; Skiba, Michael, Nardo, & Peterson, 2002). For example, defiance and disrespect are more ambiguously defined and allow more staff discretion than more objectively defined behaviors (Greflund, McIntosh, Mercer, & May, 2014). In these circumstances, educators must decide whether a student's behavior (e.g., a student sharing an opinion about an assignment) is disrespectful, whereas behaviors such as smoking or theft are far more easily determined. An analysis of school discipline outcomes for every ninth-grade student in Texas for three academic years showed that, after controlling for 83 student- and school-level factors, African American students had a 31% higher likelihood than White students of being disciplined for discretionary violations but a 23% lower likelihood of being disciplined for mandatory violations (Fabelo et al., 2011).

Other VDPs have theoretical but no empirical support to date. A previously untested extension of the subjectivity of VDPs is the discretionary judgment involved in classifying similar or borderline student behavior as severe (e.g., "fighting" warranting a major ODR) or less severe (e.g., "physical contact" warranting a minor ODR or warning slip). In keeping with our theory, educators may be more likely to overreact to minor behavior by African American students by classifying it as a more severe (major) incident. Another hypothesized but untested VDP is location. The nature of student–teacher interactions across different contexts may lead to different behaviors observed or varying risk for biased responding. For example, previous theories implicate a relation between discipline disproportionality and the academic achievement gap (Gregory, Skiba, & Noguera, 2010), which suggests that classrooms are themselves a potential VDP, especially during periods with a strong academic focus (e.g., literacy). There are several reasons why classrooms could be a VDP. For example, in classrooms, teachers provide more directions to complete tasks, which some students may perceive as controlling, too difficult, or irrelevant to their lives. There may also be less engaging instructional techniques used, a mismatch of teacher and student goals (e.g., instructional time versus

socialization), or a fear of "losing control" of a classroom when minor noncompliance occurs (Fenning & Rose, 2007; Okonofua, Walton, & Eberhardt, 2016). Evidence for other VDPs, such as time of day (Kouchaki & Smith, 2014; Linder et al., 2014) or decision-maker fatigue or hunger (Danziger, Levav, & Avnaim-Pesso, 2011; Gailliot, Peruche, Plant, & Baumeister, 2009; Kouchaki & Smith, 2014) align with the behavioral principle of a motivating (or establishing) operation (Laraway, Snycerski, Michael, & Poling, 2003). These VDPs are not related to student behavior, but rather the internal state of the decision maker. For example, when a teacher is exhausted, sending a student from a particular group out of the classroom becomes particularly reinforcing. The student's behavior is no different, but the internal state of the educator may make unexpected behavior more likely to be categorized as defiance. These VDPs have been found in other fields (e.g., law, medicine) but have yet to be explored in education.

Finally, research shows significant gender differences and possible race–gender interactions in rates of student discipline (Fabelo et al., 2011; Losen et al., 2015). For example, discipline rates for African American females can be much higher than for White females and in some cases are more similar to those of African American males (Blake, Butler, Lewis, & Darensbourg, 2011; Crenshaw, Ocen, & Nanda, 2015), suggesting potential interactions between students' behaviors and the salience of either their race or gender to teachers (Sinclair & Kunda, 1999, 2000). Such interactions would indicate that student gender may influence a teacher's decision to send a student to the office, making it a potential VDP as well.

Understanding the mechanisms by which implicit bias emerges and affects school discipline decision making will be important for developing effective interventions. If those conditions that are most likely to be influenced by implicit bias are also those most likely to produce disproportionality, research from other fields would indicate the following intervention approaches as promising: (a) making the specific decision more objective (e.g., creating operational definitions of behavior violations that should and should not result in an ODR), (b) teaching individuals to recognize these VDPs (including personal motivating operations), and (c) practicing and using alternative responses to behavior that are instructive and nonexclusionary (McIntosh, Girvan, Horner, & Smolkowski, 2014).

The Present Study

The main purpose of this study was to identify patterns in actual school discipline data that would support or disprove the VDP model. Our conceptual model predicts that, within the context of adult decisions about disciplinary actions, certain situations are more vulnerable to the impacts of implicit biases. This study tested two specific research hypotheses consistent with previous research and our VDP model:

1. Compared to White students, African American students receive ODRs at a higher rate for subjectively defined behaviors than for objectively defined behaviors.
2. The relative odds of receiving a subjective versus objective ODR will be greater for African Americans when associated with four potential VDPs: (a) at the end of the day versus earlier in the day, (b) in classroom settings versus nonclassroom settings, (c) for classifying incidents as severe versus minor, and (d) for African American females as opposed to White females.

To the extent patterns in discipline data suggest that disproportionality tends to be concentrated in particular situations, such as those discussed above, it indicates that the VDP approach may be an important strategy to improve equity in school discipline. If the model was supported, the results could then serve as a guide for understanding the larger problem of disproportionality as one that is, in fact, composed of smaller, more specific, and potentially more manageable situations that can be targeted for the development of effective interventions.

Method

Participants and Settings

The sample included 483,686 ODRs issued in the 2011–2012 academic year to 235,542 students by 53,030 educators in 1,666 elementary schools that were using the School-Wide Information System (SWIS; May et al., 2013), an online computer application for tracking and analyzing ODRs. Schools came from 45 states across the U.S. The average enrollment was 493 (SD = 184), the average percent of students receiving free or reduced price meals was

55% (*SD* = 24%), and the average percent of non-White students was 47% (*SD* = 26%).

Consistent with our goals, this sample included a number of restrictions. We included only elementary schools to examine relations in a relatively consistent student–teacher context, as opposed to middle and high schools, which tend to operate differently (e.g., different teachers each period). We included only schools that coded race or ethnicity for at least 80% of ODRs and with at least 10 African American and 10 White students to avoid using estimates with schools that have little or no racial diversity. For the present analysis, we included only ODRs delivered to African American or White students to narrow our focus to the most common comparison for disproportionality (Skiba et al., 2011). Because the analysis focused on subjectively versus objectively defined ODRs, the sample excluded ODRs for behavior types that could not reliably be classified as one or the other by a panel of educational experts, as described below (Greflund et al., 2014). The sample included 424,840 subjective ODRs and 58,846 objective ODRs.

Measures

Office Discipline Referrals

ODRs are standardized forms used to document incidents of problem behavior (Sugai, Sprague, Horner, & Walker, 2000). School personnel issue ODRs to students for a defined set of behavior violations (e.g., fighting). When the process and specific behaviors are operationally defined (as is required for the use of SWIS), ODRs can be reliable and valid indicators of problem behavior (Irvin, Tobin, Sprague, Sugai, & Vincent, 2004; McIntosh, Campbell, Carter, & Zumbo, 2009; Walker, Cheney, Stage, & Blum, 2005). SWIS ODRs include a range of fields (e.g., location, time of day, student, staff) that can be used to identify specific situations with elevated problem behavior. In this study, these fields were used to identify and test whether and to what extent specific situations had increased disproportionality (i.e., theorized VDPs).

Subjectivity of ODRs

Each ODR behavior type was classified as either subjective (e.g., defiance, disrespect, disruption) or objective (e.g., fighting, theft, truancy) by an expert panel, composed of four researchers in school discipline, racial/ethnic disproportionality, and/or culturally responsive behavior support, which rated the specific SWIS behavior definitions used for ODRs (Greflund et al., 2014). ODRs for behaviors in which the expert panel did not agree on a classification (e.g., dress code violation) or were not clearly attributable to student actions in elementary school (e.g., attendance, which is related to caregiver behavior) were removed from analyses.

Time of Day

Based on the time of the incident, ODRs were coded as occurring in 15-min intervals throughout the school day. According to our hypotheses regarding fatigue, ODRs issued in approximately the last hour of the school day (2:00 to 3:00 p.m.) were compared to ODRs issued earlier in the day (8:30 a.m. to 1:45 p.m.).

Severity of ODRs

In SWIS, school personnel determine whether each behavior incident is major (i.e., requiring administrator action) or minor (i.e., expected to be handled in the classroom). Except for certain major behaviors (e.g., arson, bomb threat), ODRs can be classified by staff as either major (i.e., severe) or minor (i.e., less severe).

School-Level Variables

School characteristics included proportion of students receiving free and reduced-price lunch, the proportion of African American students, and the proportion of minority students other than African American. These data were collected from the National Center for Educational Statistics (NCES) and were used as covariates to control for their influence on ODR patterns.

Procedure

ODR data from each school were extracted from the SWIS database. Each school in the study had signed a data-sharing agreement that allowed their data to be used for research purposes. Upon extraction, information entered for each ODR was used to identify student race, student gender, referring staff member, and hypothesized VDPs (e.g., subjectivity, time of day, location, severity).

Analytic Plan

Consistent with the VDP model, we hypothesized that the teacher decision to issue ODRs was subject to implicit bias, and thus that biases increased the odds of subjective compared to objective ODRs for African American students during specific VDPs. Accordingly, the dependent variable is the relative odds of an ODR for a subjectively defined behavior versus an objectively defined behavior. For ease of reference, in describing the analysis and results, we refer to this outcome as the odds of a subjective ODR or just use the term "subjective ODR."

To assess these hypotheses, we fit a series of multilevel logistic regression models with different predictors of the odds of a subjective ODR: (a) unconditional model with no predictors, (b) school-level covariates only, (c) African American recipient, (d) African American and end of day, (e) African American and classroom, (f) African American and major ODR, and (g) African American and female. We then fit three final models that included each of the three hypothesized VDPs with both African American and female as predictors. Models with multiple predictors also included all relevant interactions. Because ODRs were collected from different educators in different schools, we also included these two sources of random variation into all models (i.e., ODRs were nested within educators and schools).

For certain analyses, we further restricted the sample, so the sample size varied by model. First, missing data on covariates reduced the sample to 455,527 ODRs (94.2%) and 1,595 schools (95.7%). The model that tested major versus minor ODRs was restricted to schools that use both majors and minors, 368,692 subjective ODRs (76.2%). Models that tested time of day included only the 402,724 ODRs (83.3%) that occurred when classes are typically in session, from 8:30 a.m. to 3:00 p.m.

Interpretation of coefficients. Logistic regressions are particularly useful because the results allow the calculation of odds ratios, a form of effect size, from the raw parameter estimates (Judge & Cable, 2004). The odds ratio is an estimate of the increase in odds per unit change of the predictor, so if the model produced a raw coefficient for African Americans of 0.405, then the odds ratio = $e^{0.405} = 1.5$ and implies that African American students are 50% more likely to receive a subjective ODR as White students (holding all other predictors, covariates, and random effects constant). Consider, for example, a school with 100 African American and 100 White students. An odds ratio of 1.5 would indicate that, if the 100 White students received a total of 20 subjective ODRs, then the African American students would have received a total of 30, controlling for covariates. An odds ratio of 1 indicates that African American and White students are equally likely to receive a subjective ODR. An odds ratio of 0.5 indicates that the outcome is half as likely for that group. As such, the odds ratio for African American students is our primary indicator of disproportionality in this set of analyses.

Because the large sample made even trivial differences statistically significant, we focused on whether odds ratios in the hypothesized VDPs represent substantial increases in risk of disproportionality in the expected direction. In the United States, each state is left to determine its own criterion for significant racial disproportionality in education (U.S. Government Accountability Office, 2013), leaving little guidance for those seeking a benchmark for meaningful differences. To determine this threshold, we used the "four-fifths" or "80% rule" used by the Department of Justice and Equal Employment Opportunity Commission (EEOC) to identify employment practices that result in "serious discrepancies" based on race or other protected classes (Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, & Department of Justice, 1978). The rule translates to an odds ratio of 1.25 or greater or, equivalently, 0.80 or less. Referring again to the hypothetical school with 100 White and 100 African American students, an odds ratio of 1.25 implies that if 20 White students were sent to the office for a subjective offense, such as inappropriate language, 25 African American students would have been sent to the office, which would be problematic if the actual rate of behaviors were similar for both White and African American students (e.g., Bradshaw et al., 2010). This threshold for odds ratios is conservative and is thus not intended to suggest that odds ratios of, say, 1.10 to 1.25 are unimportant. To the contrary, the EEOC, for example, considers evidence of racial discrepancies that do not satisfy the "four-fifths" rule to constitute an adverse impact if the discrepancies are based on large (e.g., nationwide) samples and otherwise practically significant. Consistent with this, we considered those odds ratios that reveal disproportionality within the [0.80, 1.25] interval to be worth examination and odds ratios equal

to or outside of the [0.80, 1.25] interval to identify a situation in which disproportionality may be especially problematic.

Results

Table 1 describes the complete set of models and reports estimates, standard errors, and statistical significance indicators. This table provides coefficients in the log odds scale, which are not easily interpretable but fully describe our statistical models. We interpreted the results in terms of odds and odds ratios, which we present in *Tables 2* through *6*.

The analysis included five predictors associated with ODRs: African American (AA), female, end of day, classroom, and major ODR. The names denote the event (coded "1") compared to the converse (coded "0"; i.e., White, male, earlier in the day, and minor ODR). The sample contained 38% AA, 24% female, 19% end of day, 57% in the classroom setting, and 33% major ODRs. The sample ranged from 368,692 to 483,686 depending on the measure and model. For analyses of major ODRs, for example, schools were excluded if they did not use minor ODRs, leaving 368,692 ODRs for the analysis. End of day versus earlier ODRs included only those given during the regular portion of the day, from 8:30 a.m. to 3:00 p.m. (N = 402,724). The analyses included data from 1,100 to 1,595 schools, which should offer robust, generalizable results.

Subjective ODRs

The dependent variable for all analyses was the subjectivity of ODRs: subjective versus objective ODRs (subjective ODR). Approximately 88% (424,840) of the ODRs were subjective (58,846 were objective). Thus, we calculated 7.2:1 odds of a subjective ODR across educators and schools: For every one objective ODR in the dataset, there were 7.2 subjective ODRs. Once we account for clustering within educators and schools, the odds of a subjective ODR for a specific individual educator and school, on average, is 13.8:1 or simply 13.8.

Random (Clustering) Effects

The variances in *Table 1* represent random effects from the multilevel logistic regression models. We interpreted the educator- and school-level variances in terms of the median odds ratio (MOR), which is "the median value of the odds ratio between the [cluster] at highest risk and the [cluster] at lowest risk when randomly picking out two [clusters]" (Merlo et al., 2006, p. 292). The MOR for educator was 2.81, 95% CI [2.75, 2.86], and the MOR for school was 2.75, 95% CI [2.63, 2.86]. These MORs suggest considerable variability across teachers and schools in the odds of subjective ODRs compared to objective ODRs.

Fixed Effects

The fixed effects describe the odds ratios for race, gender, and the three VDPs, controlling for school-level variables and accounting for variability at the educator and school levels. The odds ratio for African American students was 1.20. With these controls, on average, African American students were 1.2 times more likely to receive a subjective ODR than White students from the same teacher in the same school. As an illustration, if a teacher issued 10 subjective and 10 objective ODRs to White students, we would expect that same teacher to have issued 12 subjective and 10 objective ODRs to African American students.

AA and VDPs: Two-Way Interactions

We estimated interactions between African American (AA) and end of day versus other times, classroom versus nonclassroom settings, major versus minor ODRs, and student gender. The results are presented in *Tables 2* through *6* and shown in *Figure 1*.

End of day. We classified end-of-day ODRs as those delivered between 2:00 and 3:00 p.m. and compared them to ODRs delivered between 8:30 a.m. and 1:45 p.m. *Table 2* partitions the effects from the fourth multilevel logistic regression model, which include predictors AA, end of day, and their interaction, in terms of odds ratios. The first four rows of the table give the odds (not odds ratios) of a subjective ODR for each subgroup. The odds represent the raw likelihood of a subjective referral, before comparing it to the odds of another situation (i.e., odds ratios). As shown in *Table 2*, the odds increase for White students at the end of the day but remain consistently high for African American students.

The next set of four rows, Rows 5 through 8, provide odds ratios associated with AA: White students for either early in the day or the end of day, and vice versa. Earlier in the day, African American students were 1.25

TABLE 1
Multilevel Logistic Regression Models Predicting the Receipt of a Subjective Referral by Demographic and Vulnerable Decision Point (VDP) Indicators

	(1) Unconditional	(2) Covariates Only	(3) African American (AA)	(4) AA & End of Day	(5) AA & Classroom	(6) AA & Major Referral	(7) AA & Female	(8) AA, Female, & End of Day	(9) AA, Female, & Classroom	(10) AA, Female, & Major Referral
Fixed effects										
Intercept	2.63*** (.05)	2.59*** (.05)	2.52*** (.05)	2.49*** (.05)	2.38*** (.05)	2.59*** (.05)	2.59*** (.05)	2.57*** (.05)	2.41*** (.05)	2.66*** (.05)
AA			0.18*** (.02)	0.23*** (.02)	0.14*** (.02)	0.09*** (.02)	0.14*** (.02)	0.18*** (.02)	0.11*** (.02)	0.07** (.02)
Female							−0.30*** (.02)	−0.33*** (.02)	−0.16*** (.02)	−0.31*** (.02)
VDP				0.19*** (.02)	0.29*** (.02)	−0.22*** (.02)		0.17*** (.02)	0.36*** (.02)	−0.23*** (.02)
AA × Female							0.20*** (.03)	0.22*** (.03)	0.13** (.04)	0.14*** (.04)
AA × VDP				−0.20*** (.03)	0.10*** (.02)	0.20*** (.03)		−0.16*** (.04)	0.06* (.03)	0.16*** (.03)
Female × VDP								0.08 (.04)	−0.28*** (.03)	0.01 (.04)
AA × Female × VDP								−0.15* (.07)	0.13* (.05)	0.19** (.06)
FRL %		0.09*** (.01)	0.09*** (.01)	0.10*** (.01)	0.09*** (.01)	0.07*** (.02)	0.09*** (.01)	0.10*** (.01)	0.09*** (.01)	0.08*** (.02)
Black %		0.03 (.02)	0.01 (.02)	0.01 (.02)	0.00 (.02)	−0.01 (.02)	0.01 (.02)	0.01 (.02)	0.00 (.02)	−0.01 (.02)
Non-Black %		−0.02 (.02)	−0.03* (.02)	−0.03 (.02)	−0.03* (.02)	−0.03 (.02)	−0.03* (.02)	−0.03 (.02)	−0.03* (.02)	−0.03 (.02)
Random effects										
Educator	1.17 (.02)	1.16 (.02)	1.16 (.02)	1.14 (.02)	1.16 (.02)	1.17 (.03)	1.15 (.02)	1.14 (.02)	1.16 (.02)	1.17 (.03)
School	1.12 (.05)	1.04 (.05)	1.04 (.05)	1.03 (.05)	1.04 (.05)	1.04 (.05)	1.05 (.05)	1.03 (.05)	1.04 (.05)	1.04 (.05)

Note. The table reports fixed-effect coefficients, β, in terms of the log of the odds ratio, and e^{β} would be interpreted as an odds ratio. The coefficients are reported with their statistical significance (asterisks) and standard error (in parentheses). Although we conducted *t* tests to test parameters, the degrees of freedom ranged from 579 and 1665, which do not differ appreciably from the standard normal distribution (e.g., *Z* tests). Variances were not tested for statistical significance. AA = African American students; FRL = Percent of students receiving free and/or reduced price lunches.
* $p < .05$. ** $p < .01$. *** $p < .001$.

TABLE 2
Odds and Odds Ratios (ORs) of a Subjective Referral for Specific Contrasts between African American (AA) versus White and Three Vulnerable Decision Points (VDPs) from Multilevel Logistic Regression

Analysis Focus	Student Race	VDP	Odds or OR	95% CI		Four-Fifths Rule
				Lower	Upper	
End of Day	White	Earlier	12.11	11.38	12.89	
	White	End of day	14.62	13.64	15.67	
	AA	Earlier	15.17	14.22	16.18	
	AA	End of day	14.97	13.86	16.16	
	AA:White	Earlier	1.25	1.21	1.30	>
	AA:White	End of day	1.02	0.96	1.09	↔
	White	End of day:Earlier	1.21	1.16	1.25	↔
	AA	End of day:Earlier	0.99	0.94	1.04	↔
	AA:White	End of day:Earlier	0.82	0.77	0.87	↔
Classroom	White	Other setting	10.80	10.14	11.50	
	White	Classroom	14.38	13.5	15.32	
	AA	Other setting	12.37	11.58	13.21	
	AA	Classroom	18.16	16.99	19.41	
	AA:White	Other setting	1.15	1.10	1.19	↔
	AA:White	Classroom	1.26	1.21	1.31	>
	White	Classroom:Other	1.33	1.29	1.37	>
	AA	Classroom:Other	1.47	1.41	1.53	>
	AA:White	Classroom:Other	1.10	1.05	1.16	↔
Major	White	Minor	13.28	12.32	14.31	
	White	Major	10.69	9.91	11.55	
	AA	Minor	14.55	13.46	15.73	
	AA	Major	14.37	13.25	15.58	
	AA:White	Minor	1.10	1.05	1.14	↔
	AA:White	Major	1.34	1.28	1.41	>
	White	Major:Minor	0.81	0.78	0.83	↔
	AA	Major:Minor	0.99	0.94	1.04	↔
	AA:White	Major:Minor	1.23	1.16	1.30	↔

Note. This table provides the odds or odds ratio from specific contrasts created from the models in *Table 1*. For rows that contain singular terms (e.g., White, AA, or Earlier), the table reports to odds of a subjective referral. For rows that contain comparison (e.g., AA:White, End of Day:Earlier), the rows provide ORs. Confidence intervals (CIs) that exclude 1.0 indicate a statistically significant result. The four-fifths rule indicates whether a particular OR equals or exceeds (>) four-fifths (1.25), its reciprocal (<; 0.80), or falls within the bounds of the four-fifths rule (↔).

times as likely as White students to receive a subjective ODR. Near the end of the day, however, African American students were 1.02 times as likely as White students to receive a subjective ODR. The odds ratio for earlier in the day meets the criteria for disproportionality according to the four-fifths rule.

Classroom. Continuing with the same approach to interpretation, *Table 2* shows that subjective ODRs are 1.26 times more likely to have been given to African American students than White students within classrooms. The model also shows that subjective ODRs were more likely in classrooms than other settings for both African American students, odds ratio = 1.47, and White students, odds ratio = 1.33.

Major ODRs. African American students were also much more likely to receive major subjective ODRs compared to White students, odds ratio = 1.34. For minor ODRs, the odds ratio associated with AA was 1.1, also greater than 1.0, likely representing important levels of disproportionality but below our four-fifths criterion level for indicating a serious problem.

Gender. Males received three fourths of the ODRs in the sample. *Table 3* presents the partitioned effects as odds ratios from multilevel logistic regression model, which included the predictors AA, female, and their interactions. The results suggest that African American males were more likely to receive subjective ODRs than White males, odds ratio = 1.15, but the

TABLE 3
Odds and Odds Ratios (ORs) of a Subjective Referral for Specific Contrasts between African American (AA) versus White and Females versus Males

Student Race	Student Gender	Odds or OR	95% CI		Four-Fifths Rule
			Lower	Upper	
White	Male	13.27	12.47	14.11	
White	Female	9.83	9.21	10.50	
AA	Male	15.28	14.33	16.30	
AA	Female	13.80	12.87	14.81	
AA:White	Male	1.15	1.11	1.19	↔
AA:White	Female	1.40	1.34	1.47	>
White	Female:Male	0.74	0.72	0.76	<
AA	Female:Male	0.90	0.87	0.94	↔
AA:White	Female:Male	1.22	1.16	1.29	↔

Note. This table provides the odds or odds ratio from specific contrasts created from the models in *Table 1*. For rows that contain singular terms (e.g., White, AA, or Male), the table reports to odds of a subjective referral. For rows that contain comparison (e.g., AA:White, Female:Male), the rows provide ORs. Confidence intervals (CIs) that exclude 1.0 indicate a statistically significant result. The four-fifths rule indicates whether a particular OR equals or exceeds (>) four-fifths (1.25), its reciprocal (<; 0.80), or falls within the bounds of the four-fifths rule (↔).

risk was much higher for African American females compared to White females, odds ratio = 1.40. Females also tended to receive fewer subjective ODRs than males, and more so for White females, odds ratio = 0.74, than African American females, odds ratio = 0.90.

Gender, AA, and VDPs: Three-Way Interactions

Although the two-way interactions demonstrated disproportionality among the delivery of subjective ODRs to African American students, the differences by gender suggested a more complex interaction. We therefore fit the data to three additional models to tease out differences by gender.

End of day. *Table 4* provides the odds ratios associated with the three-way AA × Female × End-of-Day Interaction and all specific subgroup contrasts. The first eight rows give the odds of a subjective ODR by ODR subgroup. The middle 12 rows give odds ratios for one contrast (e.g., AA:White or female:male) within subgroups defined by the other two predictors, and the first four of these rows represent the contrasts of most interest for the present analysis. The final seven rows provide odds ratios that represent the two-way interactions and, lastly, the three-way interaction.

This analysis demonstrated that African American females were 1.49 times more likely to receive a subjective ODR before the end of the day than White females. African American males were also more likely to receive subjective ODRs than White males, odds ratio = 1.19, earlier in the day. African American females, odds ratio = 1.10, and males, odds ratio = 1.01, were both more likely to receive subjective ODRs at the end of the day as well, but these risks were much smaller.

Classroom. African American females were more likely to receive subjective ODRs in the classroom, odds ratio = 1.54, and in other settings, odds ratio = 1.27, than White females. Again, African American males were also at more risk of subjective ODRs in both settings, but at a lower level (see *Table 5*).

Major ODRs. The final analyses specifically examined risk by rated severity of behavior (major vs. minor). As noted above, African American students were more likely to receive subjective major ODRs compared to White students, odds ratio = 1.34. The disaggregation of effects by gender in *Table 6* shows that the risk for subjective majors was particularly prominent for African American females, odds ratio = 1.73, although still problematic for African American males, odds ratio = 1.25.

Supplemental Analyses

We failed to confirm our hypothesis that the end of the day represented a VDP for African American students. Instead, end of day appeared to relate to increased odds of subjective ODRs for White students relative to African American students, which had the effect of reducing disproportionality itself. To further explore the impact of time of day, we tested a competing

TABLE 4
Odds and Odds Ratios (ORs) of a Subjective Referral for Specific Contrasts between African American (AA) versus White, Females versus Males, and End of Day versus Earlier in the Day

Student Race	Student Gender	VDP	Odds or OR	95% CI		Four-Fifths Rule
				Lower	Upper	
White	Male	Earlier	13.01	12.22	13.85	
White	Male	End of day	15.38	14.31	16.53	
White	Female	Earlier	9.38	8.77	10.05	
White	Female	End of day	12.07	11.01	13.23	
AA	Male	Earlier	15.53	14.53	16.60	
AA	Male	End of day	15.60	14.37	16.94	
AA	Female	Earlier	14.02	13.01	15.11	
AA	Female	End of day	13.26	11.93	14.73	
AA:White	Male	Earlier	1.19	1.15	1.24	↔
AA:White	Male	End of day	1.01	0.94	1.09	↔
AA:White	Female	Earlier	1.49	1.41	1.58	>
AA:White	Female	End of day	1.10	0.98	1.23	↔
White	Female:Male	Earlier	0.72	0.70	0.75	<
White	Female:Male	End of day	0.78	0.73	0.85	<
AA	Female:Male	Earlier	0.90	0.86	0.95	↔
AA	Female:Male	End of day	0.85	0.77	0.94	↔
White	Male	End of day:Earlier	1.18	1.13	1.23	↔
White	Female	End of day:Earlier	1.29	1.19	1.39	>
AA	Male	End of day:Earlier	1.00	0.94	1.07	↔
AA	Female	End of day:Earlier	0.95	0.86	1.04	↔
AA:White	Female:Male	Earlier	1.25	1.18	1.33	>
AA:White	Female:Male	End of day	1.08	0.95	1.23	↔
AA:White	Male	End of day:Earlier	0.85	0.79	0.92	↔
AA:White	Female	End of day:Earlier	0.73	0.65	0.83	<
White	Female:Male	End of day:Earlier	1.09	1.00	1.19	↔
AA	Female:Male	End of day:Earlier	0.94	0.84	1.05	↔
AA:White	Female:Male	End of day:Earlier	0.86	0.75	1.00	↔

Note. This table provides the odds or odds ratio from specific contrasts created from the models in *Table 1*. For rows that contain singular terms (e.g., White, AA, or Male), the table reports to odds of a subjective referral. For rows that contain comparison (e.g., AA:White, Female:Male), the rows provide ORs. Confidence intervals (CIs) that exclude 1.0 indicate a statistically significant result. The four-fifths rule indicates whether a particular OR equals or exceeds (>) four-fifths (1.25), its reciprocal (<; 0.80), or falls within the bounds of the four-fifths rule (↔). VDP = vulnerable decision point.

VDP hypothesis: the start of the school day. Early mornings represent a stressful time of day for teachers, as many struggle to organize their students for academic instruction. Hence, we explored the beginning of the day as an alternative VDP, defining it as the first 90 minutes of the school day (i.e., ODRs delivered from 8:30 to 10:00 a.m.) compared to those delivered later (i.e., 10:15 a.m. to 3:00 p.m.).

These results showed that African American students were considerably more likely to receive subjective ODRs in the first 90 min compared to White students, odds ratio = 1.40. After the school day is underway (i.e., after the first 90 min), African American students continued to receive more subjective ODRs than White students, odds ratio = 1.19, but at a reduced level of disproportionality. The three-way interactions with gender showed that the level of disproportionality was high for both males and females. African American males received more subjective referrals than White males in the first 90 min, odds ratio = 1.32, but the risk was smaller throughout the rest of the day, odds ratio = 1.13. The risks were more troublesome for African American females, who were much more likely to receive subjective ODRs than White girls in the first 90 min, odds ratio = 1.72. African American females, compared to White females, also continued to exceed the four-fifths criterion for disproportionality during the remainder of the day, odds ratio = 1.35.

TABLE 5
Odds and Odds Ratios (ORs) of a Subjective Referral for Specific Contrasts between African American (AA) versus White, Females versus Males, and Classroom versus Other Settings

Student Race	Student Gender	VDP	Odds or OR	95% CI		Four-Fifths Rule
				Lower	Upper	
White	Male	Other setting	11.14	10.45	11.87	
White	Male	Classroom	15.97	14.97	17.03	
White	Female	Other setting	9.51	8.85	10.23	
White	Female	Classroom	10.35	9.65	11.11	
AA	Male	Other setting	12.44	11.62	13.32	
AA	Male	Classroom	19.03	17.76	20.39	
AA	Female	Other setting	12.05	11.12	13.05	
AA	Female	Classroom	15.92	14.72	17.21	
AA:White	Male	Other setting	1.12	1.07	1.17	↔
AA:White	Male	Classroom	1.19	1.14	1.25	↔
AA:White	Female	Other setting	1.27	1.18	1.36	>
AA:White	Female	Classroom	1.54	1.44	1.64	>
White	Female:Male	Other setting	0.85	0.82	0.89	↔
White	Female:Male	Classroom	0.65	0.62	0.68	<
AA	Female:Male	Other setting	0.97	0.91	1.03	↔
AA	Female:Male	Classroom	0.84	0.79	0.89	↔
White	Male	Classroom:Other setting	1.43	1.39	1.48	>
White	Female	Classroom:Other setting	1.09	1.03	1.15	↔
AA	Male	Classroom:Other setting	1.53	1.46	1.60	>
AA	Female	Classroom:Other setting	1.32	1.23	1.42	>
AA:White	Female:Male	Other setting	1.13	1.05	1.22	↔
AA:White	Female:Male	Classroom	1.29	1.20	1.39	>
AA:White	Male	Classroom:Other setting	1.07	1.01	1.13	↔
AA:White	Female	Classroom:Other setting	1.21	1.11	1.33	↔
White	Female:Male	Classroom:Other setting	0.76	0.71	0.81	<
AA	Female:Male	Classroom:Other setting	0.86	0.80	0.94	↔
AA:White	Female:Male	Classroom:Other setting	1.14	1.03	1.26	↔

Note. This table provides the odds or odds ratio from specific contrasts created from the models in *Table 1*. For rows that contain singular terms (e.g., White, AA, or Male), the table reports to odds of a subjective referral. For rows that contain comparison (e.g., AA:White, Female:Male), the rows provide ORs. Confidence intervals (CIs) that exclude 1.0 indicate a statistically significant result. The four-fifths rule indicates whether a particular OR equals or exceeds (>) four-fifths (1.25), its reciprocal (<; 0.80), or falls within the bounds of the four-fifths rule (↔). VDP = vulnerable decision point.

Discussion

The results demonstrated that, overall, African American students were more likely to receive subjective ODRs than White students (*Figure 1*). In addition, African American students were at greater risk for subjective ODRs than White students in the classroom compared to other settings and when the ODR was perceived as a major offense rather than minor. Rather than finding that disproportionality was stable and pervasive throughout all situations, we found specific situations in which disproportionality in subjective ODRs was more pronounced, as well as situations in which the provision of subjective ODRs was not inequitable. These patterns lend support to the VDP model (McIntosh, Girvan, Horner, & Smolkowski, 2014) and are consistent with research suggesting that some decisions and decision contexts are more influenced by implicit racial biases. Our findings highlight the staff decision to issue major ODRs in the classroom, particularly in the first 90 min of the day, as particularly disproportionate and worthy of focused equity intervention.

The results for the end of the day versus earlier, in particular, were not consistent with our hypotheses. The end of the day may have been a poor choice as a VDP given its competition with other times that may have also been VDPs, such as the early morning. Subjective ODRs became more likely at the end of the day for White students, which is consistent with the possibility that teachers accumulate information about particular students over the

TABLE 6
Odds and Odds Ratios (ORs) of a Subjective Referral for Specific Contrasts between African American (AA) versus White, Females versus Males, and Major Referrals versus Minor Referrals

Student Race	Student Gender	VDP	Odds or OR	95% CI		Four-Fifths Rule
				Lower	Upper	
White	Male	Minor	14.30	13.26	15.43	
White	Male	Major	11.40	10.54	12.32	
White	Female	Minor	10.45	9.64	11.32	
White	Female	Major	8.38	7.66	9.17	
AA	Male	Minor	15.30	14.12	16.58	
AA	Male	Major	14.26	13.12	15.50	
AA	Female	Minor	12.79	11.71	13.98	
AA	Female	Major	14.50	13.12	16.02	
AA:White	Male	Minor	1.07	1.02	1.12	↔
AA:White	Male	Major	1.25	1.18	1.32	>
AA:White	Female	Minor	1.22	1.15	1.31	↔
AA:White	Female	Major	1.73	1.58	1.89	>
White	Female:Male	Minor	0.73	0.70	0.76	<
White	Female:Male	Major	0.74	0.69	0.78	<
AA	Female:Male	Minor	0.84	0.79	0.89	↔
AA	Female:Male	Major	1.02	0.94	1.10	↔
White	Male	Major:Minor	0.80	0.77	0.83	<
White	Female	Major:Minor	0.80	0.75	0.86	↔
AA	Male	Major:Minor	0.93	0.88	0.99	↔
AA	Female	Major:Minor	1.13	1.04	1.23	↔
AA:White	Female:Male	Minor	1.14	1.06	1.23	↔
AA:White	Female:Male	Major	1.38	1.25	1.53	>
AA:White	Male	Major:Minor	1.17	1.10	1.25	↔
AA:White	Female	Major:Minor	1.41	1.27	1.57	>
White	Female:Male	Major:Minor	1.01	0.94	1.08	↔
AA	Female:Male	Major:Minor	1.22	1.10	1.34	↔
AA:White	Female:Male	Major:Minor	1.21	1.07	1.36	↔

Note. This table provides the odds or odds ratio from specific contrasts created from the models in *Table 1*. For rows that contain singular terms (e.g., White, AA, or Male), the table reports to odds of a subjective referral. For rows that contain comparison (e.g., AA:White, Female:Male), the rows provide ORs. Confidence intervals (CIs) that exclude 1.0 indicate a statistically significant result. The four-fifths rule indicates whether a particular OR equals or exceeds (>) four-fifths (1.25), its reciprocal (<; 0.80), or falls within the bounds of the four-fifths rule (↔). VDP = vulnerable decision point.

course of the day, an effect that actually attenuated the substantial racial differences seen earlier in the day. Instead, we found support in our exploratory analysis for the first 90 min of the school day as a possible VDP. Although speculative, it is possible that early morning teacher stress, a stronger academic focus at the start of the day, increased disorder from transitioning between home and school environments and corresponding changes in behavioral expectations, or some combination of these features may explain this finding.

Our gender analyses indicated that subjective ODR rates differed by gender and the intersection of race and gender. The risk of disproportionate ODRs associated with VDPs reached problematic levels among African American males versus White males. However, evidence of disproportionality was particularly strong for African American females compared to White females. One example is in major ODRs, an area with particularly important implications for educational outcomes, because their receipt is more likely to have been associated with the student's removal from the classroom, leading to less instructional time spent with peers. African American students were 1.34 times as likely to receive a major subjective ODR. African American males were given major ODRs at a rate of 1.25 times as often as their White counterparts, but African American females in particular were much more likely, odds ratio = 1.73, to receive major ODRs when compared to White females, consistent with previous research (Blake et al., 2011). This finding implies that African American

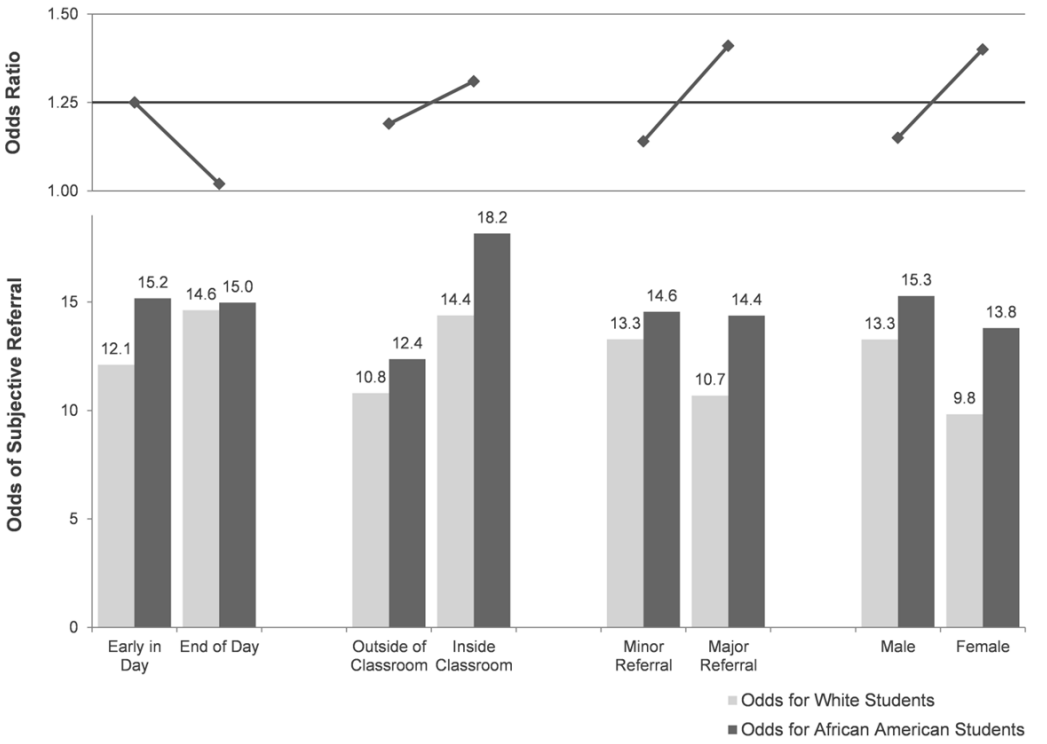

Figure 1. Comparisons of potential VDPs. The columns at the bottom represent the average odds of an ODR for a subjectively versus objectively defined behavior by student race for each clustered condition. The diamonds linked by lines represent the odds ratio within each cluster (data for odds and odds ratios come from the first column of data in* Tables 2 *and 3). Odds ratios of 1 indicate no disproportionality in the clustered condition. The darker horizontal line at an odds ratio of 1.25 indicates the threshold above which we interpret the magnitude of disproportionality to be particularly problematic.

females were nearly twice as likely as White females to be removed from the classroom during instruction, which could negatively affect academic achievement.

The pattern of disproportionality we observed at the intersection of race and gender also suggests that the most influential biases may involve paternalistic gender bias (i.e., overlooking violations of female students) and in-group preferences rather than deliberate hostility toward students who look less like teachers (Greenwald & Pettigrew, 2014). Specifically, the result of our analysis suggest that a substantial proportion of the disproportionality in elementary schools is a function of teachers having very low odds, relatively speaking, of making referrals of White female students for subjectively defined compared to objectively defined behaviors. Other researchers using different criterion measures have reported similar overall patterns of rates of student discipline by race and gender, suggesting the pattern is fairly robust (Fabelo et al., 2011; Losen et al., 2015).

The observed gender pattern is also consistent with the weight of research on decisions in legal settings involving adults. Female criminal defendants and grievants in labor disputes have been found to be treated more favorably than their male counterparts, an effect attributable by psychologists directly to benevolent forms of implicit and explicit sexism (Girvan et al., 2015). Further, although one often thinks of discrimination in terms of members of a dominant group seeking to harm members of a minority group, recent examinations of social psychological research on bias suggest that most discrimination occurs because of explicit or implicit motivations to favor in-group members (Greenwald & Pettigrew, 2014; Smith, Levinson, & Robinson, 2014). Indeed, many of the interventions that have been shown to be the most effective at reducing implicit bias work primarily because they reduce positive implicit

in-group associations (Lai et al., 2013). Because roughly 76% of teachers are female and 82% are White (NCES, 2014), we estimate that about 62% of teachers were White females and 20% White males, which contrasts sharply with the estimated 7% African American teachers, or 5% African American female teachers and 2% African American male teachers. Due to their group membership or paternalistic attitudes towards certain groups in certain contexts (Sinclair & Kunda, 1999, 2000), teachers may be less inclined to categorize the behavior of White female students in particular as meriting a disciplinary response than they would African American female students or male students in general. In any case, the operation of these biases appeared to be strongest in situations that aligned with VDPs. Further, these findings point to another possible VDP, that teachers may need to be particularly careful when assessing the behavior of students who are further from their in-group, in terms of race, gender, or other characteristics.

Limitations and Future Research

This study was limited by extant data from a nonrandom sample of elementary schools. Analyses should be repeated with samples of middle and secondary schools to assess the consistency of VDPs across settings. Although schools contributed data from nearly all the U.S. states, we suggest caution in generalizing results, particularly beyond schools using a standardized system for tracking ODRs. Likewise, our analyses were limited to recorded ODRs, and therefore we could not measure behavior that met criteria for ODR but did not result in ODRs. Future research could include direct observation of behavior and analyses of whether students received ODRs at all. The analysis was also correlational in nature and does not allow for a causal interpretation. Nonetheless, the research hypotheses proposed specific patterns that were, within the precision of our VDP measures, falsifiable and subsequently supported. In addition, the VDPs represented in this research consist of proxies for potentially more accurately defined VDPs. Subsequent research should address these limitations and include more information about teachers' race and gender and the school day (e.g., exactly when teachers may be most fatigued or hungry). Schools may also have unique patterns of VDPs that do not conform to these general VDPs (e.g., different schedules). Finally, the promise of the VDP model is limited until intervention research can confirm its intervention validity. Such research is currently underway.

Implications for Practice

Despite these limitations, the results provide tentative support for the importance of considering VDPs as important variables in understanding and reducing school discipline disparities. The findings suggest two avenues. First, improving the specificity of definitions of subjective ODRs such as defiance and disrespect (i.e., providing definitions that reduce ambiguity as much as possible) could attenuate the influence of implicit bias on discipline decisions. School personnel can decrease (but not eliminate) subjectivity by creating and using operational definitions of each behavior, as well as the thresholds for no ODR, a minor ODR, and a major ODR. Second, these general VDPs (i.e., first 90 min of the day, classroom, assessing severity) could be used to help educators identify specific decisions that are vulnerable to bias and use alternative responses in place of issuing ODRs that perpetuate disproportionality. Once they are aware of these VDPs, teachers may be trained in responses that are more instructional than exclusionary, such as teaching or reteaching expectations or visibly modeling cool-down strategies for students (McIntosh, Ellwood, & McCall, 2016; McIntosh, Girvan, Horner, Smolkowski, & Sugai, 2014). Administrators can be encouraged to use more instructional or restorative alternatives to suspension (Nese, Massar, & McIntosh, 2015) and use interventions such as Check-in Check-out, which have been shown to be effective with African American students (Vincent, Tobin, Hawken, & Frank, 2012). Schools can also use their own discipline data to identify school-specific VDPs (McIntosh, Barnes, Morris, & Eliason, 2014), which could be even more effective. In addition, professional development can help educators identify and counteract their own personal VDPs. For example, if individuals make less equitable discipline decisions when they are stressed in the early morning, they can use this knowledge as a cue to slow the decision-making process during these moments. Finally, preventive approaches, such as proactively teaching classroom routines, using acknowledgement systems equitably, and enhancing the level of student engagement in classroom instruction, may prevent VDPs in the first place (Chaparro,

Nese, & McIntosh, 2015; Tobin & Vincent, 2011). Although promising, these implications should be tested through rigorous intervention research.

REFERENCES

Anyon, Y., Jenson, J. M., Altschul, I., Farrar, J., McQueen, J., Greer, E., … Simmons, J. (2014). The persistent effect of race and the promise of alternatives to suspension in school discipline outcomes. *Children and Youth Services Review*, *44*, 379–386. doi:10.1016/j.childyouth.2014.06.025

Blake, J. J., Butler, B. A., Lewis, C., & Darensbourg, A. (2011). Unmasking the inequitable discipline experiences of urban Black girls: Implications for urban educational stakeholders. *The Urban Review*, *43*, 90–116. doi:10.1007/s11256-009-0148-8

Bobo, L., Charles, C., Krysan, M., & Simmons, A. (2012). The real record on racial attitudes. In P. V. Marsden (Ed.), *Social trends in American life: Evidence from the General Social Survey since 1972* (pp. 38–83). Princeton, NJ: Princeton University Press.

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

Chaparro, E. A., Nese, R. N. T., & McIntosh, K. (2015). *Examples of engaging instruction to increase equity in education*. Eugene, OR: Center on Positive Behavioral Interventions and Supports, University of Oregon.

Crenshaw, K. W., Ocen, P., & Nanda, J. (2015). *Black girls matter: Pushed out, overpoliced, and underprotected*. New York: African American Policy Forum.

Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, *108*, 6889–6892. doi:10.1073/pnas.1018033108

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18. doi:10.1037/0022-3514.56.1.5

Dovidio, J. F., & Gaertner, S. L. (2000). Aversive racism and selection decisions: 1989 and 1999. *Psychological Science*, *11*, 315–319. doi:10.1111/1467-9280.00262

Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, & Department of Justice. (1978). Uniform guidelines on employee selection procedures. *Federal Register*, *43*, 38290–38315.

Fabelo, T., Thompson, M. D., Plotkin, M., Carmichael, D., Marchbanks, M. P. I., & Booth, E. A. (2011). *Breaking schools' rules: A statewide study of how school discipline relates to students' success and juvenile justice involvement*. New York: Council of State Governments Justice Center.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027. doi:10.1037/0022-3514.69.6.1013

Fenning, P., & Rose, J. (2007). Overrepresentation of African American students in exclusionary discipline: The role of school policy. *Urban Education*, *42*, 536–559. doi:10.1177/0042085907305039

Gailliot, M. T., Peruche, B. M., Plant, E. A., & Baumeister, R. F. (2009). Stereotypes and prejudice in the blood: Sucrose drinks reduce prejudice and stereotyping. *Journal of Experimental Social Psychology*, *45*, 288–290. doi:10.1016/j.jesp.2008.09.003

Girvan, E. J. (2015). On using the psychological science of implicit bias to advance anti-discrimination law. *George Mason University Civil Rights Law Journal*, *26*, 1–86.

Girvan, E. J., Deason, G., & Borgida, E. (2015). The generalizability of gender bias: Testing the effects of contextual, explicit, and implicit theories of sexism on labor arbitration decisions. *Law and Human Behavior*, *39*, 525–537. doi:10.1037/lhb0000139

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27. doi:http://dx.doi.org/10.1037/0033-295X.102.1.4

Greenwald, A. G., & Pettigrew, T. F. (2014). With malice toward none and charity for some: Ingroup favoritism enables discrimination. *American Psychologist*, *69*, 669–684. doi:http://dx.doi.org/10.1037/a0036056

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*, 17–41. doi:http://dx.doi.org/10.1037/a0015575

Greflund, S., McIntosh, K., Mercer, S. H., & May, S. L. (2014). Examining disproportionality in school discipline for Aboriginal students in schools implementing PBIS. *Canadian Journal of School Psychology*, *29*, 213–235. doi:10.1177/0829573514542214

Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher*, *39*, 59–68. doi:10.3102/0013189x09357621

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*, 131–147. doi:10.1177/10983007040060030201

Johnson, J., Whitestone, E., Jackson, L., & Gatto, L. (1995). Justice is still not colorblind: Differential racial effects of exposure to inadmissible evidence. *Personality and Social Psychology Bulletin*, *21*, 893–898. doi:10.1177/0146167295219003

Judge, T. A., & Cable, D. M. (2004). The effect of physical height on workplace success and income: preliminary test of a theoretical model. *Journal of Applied Psychology*, *89*, 428–441. doi: http://dx.doi.org/10.1037/0021-9010.89.3.428

Kouchaki, M., & Smith, I. H. (2014). The morning morality effect: The influence of time of day on unethical behavior. *Psychological Science*, *25*, 95–102. doi:10.1177/0956797613498099

Lai, C. K., Hoffman, K. M., Nosek, B. A., & Greenwald, A. G. (2013). Reducing implicit prejudice. *Social and Personality Psychology Compass*, *7*, 315–330. doi:10.1111/spc3.12023

Laraway, S., Snycerski, S., Michael, J., & Poling, A. (2003). Motivating operations and terms to describe them: Some further refinements. *Journal of Applied Behavior Analysis*, *36*, 407–414. doi:10.1901/jaba.2003.36-407

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, *125*, 255–275. doi:http://dx.doi.org/10.1037/0033-2909.125.2.255

Linder, J. A., Doctor, J. N., Friedberg, M. W., Nieva, H. R., Birks, C., Meeker, D., & Fox, C. R. (2014). Time of day and the decision to prescribe antibiotics. *JAMA Internal Medicine*, *174*, 2029–2031. doi:10.1001/jamainternmed.2014.5225

Losen, D. J., & Gillespie, J. (2012). *Opportunities suspended: The disparate impact of disciplinary exclusion from school*. Los Angeles, CA: Center for Civil Rights Remedies at The Civil Rights Project at UCLA.

Losen, D. J., Hodson, C., Keith, M. A., Morrison, K., & Belway, S. (2015). *Are we closing the school discipline gap?* Los Angeles, CA: Center for Civil Rights Remedies at The Civil Rights Project at UCLA.

Martinez, T. (2013, January). *Difference-making: Why education policymakers should attend to frequent and disparate disciplinary exclusion from school*. Symposium presented at the Closing the School Discipline Gap: Research to Practice Conference, Washington, DC.

May, S. L., Ard, W. I., Todd, A. W., Horner, R. H., Glasgow, A., Sugai, G., & Sprague, J. R. (2013). *School-wide information system*. Eugene, OR: Educational and Community Supports, University of Oregon. Retrieved from http:/www.pbisapps.org

McIntosh, K., Barnes, A., Morris, K., & Eliason, B. M. (2014). *Using discipline data within SWPBIS to identify and address disproportionality: A guide for school teams*. Eugene, OR: OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports, University of Oregon.

McIntosh, K., Campbell, A. L., Carter, D. R., & Zumbo, B. D. (2009). Concurrent validity of office discipline referrals and cut points used in schoolwide positive behavior support. *Behavioral Disorders*, *34*, 100–113.

McIntosh, K., Ellwood, K., & McCall, L. (2016). *Using discipline data within a PBIS framework to enhance equity in school discipline*. Manuscript submitted for publication.

McIntosh, K., Girvan, E. J., Horner, R. H., & Smolkowski, K. (2014). Education not incarceration: A conceptual model for reducing racial and ethnic disproportionality in school discipline. *Journal of Applied Research on Children*, *5*(2), 1–22.

McIntosh, K., Girvan, E. J., Horner, R. H., Smolkowski, K., & Sugai, G. (2014). *Recommendations for addressing discipline disproportionality in education*. Eugene, OR: OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports.

Merlo, J., Chaix, B., Ohlsson, H., Beckman, A., Johnell, K., Hjerpe, P., … Larsen, K. K. (2006). A brief conceptual tutorial of multilevel analysis in social epidemiology: Using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *Journal of Epidemiology and Community Health*, *60*, 290–297. doi:10.1136/jech.2004.029454

National Center for Education Statistics. (2014). *Digest of education statistics*. Washington, DC: U.S. Department of Education.

Nese, R. N. T., Massar, M. M., & McIntosh, K. (2015). Alternatives to suspension: Strategies for school principals. *Principal Leadership*, *16*, 52–56.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes* (pp. 265–292). New York: Psychology Press.

Okonofua, J. A., Walton, G. M., & Eberhardt, J. L. (2016). A vicious cycle: A social–psychological account of extreme racial disparities in school discipline. *Perspectives on Psychological Science*, *11*, 381–398. doi:10.1177/1745691616635592

Pearson, A. R., Dovidio, J. F., & Gaertner, S. L. (2009). The nature of contemporary prejudice: Insights from aversive racism. *Social and Personality Psychology Compass*, *3*, 314–338. doi:10.1111/j.1751-9004.2009.00183.x

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*, 751–783. doi: http://dx.doi.org/10.1037/0022-3514.90.5.751

Saucier, D. A., Miller, C. T., & Doucet, N. (2005). Differences in helping whites and blacks: A meta-analysis. *Personality And Social Psychology Review*, *9*, 2–16. doi:10.1207/s15327957pspr0901_1

Sinclair, L., & Kunda, Z. (1999). Reactions to a Black professional: Motivated inhibition and activation of conflicting stereotypes. *Journal of Personality and Social Psychology*, *77*, 885–904. doi: http://dx.doi.org/10.1037/0022-3514.77.5.885

Sinclair, L., & Kunda, Z. (2000). Motivated stereotyping of women: She's fine if she praised me but incompetent if she criticized me. *Personality and Social Psychology Bulletin*, *26*, 1329–1342. doi:10.1177/0146167200263002

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. *The Urban Review*, *34*, 317–342. doi:10.1023/A:1021320817372

Skiba, R. J., Poloni-Staudinger, L., Simmons, A. B., Feggins, L. R., & Chung, C. G. (2005). Unproven links: Can poverty explain ethnic disproportionality in special education? *Journal of Special Education*, *39*, 130–144. doi:10.1177/00224669050390030101

Smith, R. J., Levinson, J. D., & Robinson, Z. (2014). Implicit white favoritism in the criminal justice system. *Alabama Law Review*, *66*, 871–1169.

Staats, C. (2014). *State of the science: Implicit bias review 2014*. Columbus, OH: Kirwan Institute.

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. doi:10.1177/106342660000800205

Tobin, T. J., & Vincent, C. G. (2011). Strategies for preventing disproportionate exclusions of African American students. *Preventing School Failure*, *55*, 192–201. doi:10.1080/1045988X.2010.532520

U.S. Government Accountability Office. (2013). *Standards needed to improve identification of racial and ethnic overrepresentation in special education*. Report to the chairman, Committee on Health, Education, Labor, and Pensions, U.S. Senate (Report GAO-13-137). Washington, DC: Author.

Vincent, C. G., Tobin, T. J., Hawken, L. S., & Frank, J. L. (2012). Discipline referrals and access to secondary level support in elementary and middle schools: Patterns across African-American, Hispanic-American, and white students. *Education and Treatment of Children*, *35*, 431–458. doi:10.1353/etc.2012.0018

Walker, B., Cheney, D., Stage, S. A., & Blum, C. (2005). Schoolwide screening and Positive Behavior Supports: Identifying and supporting students at risk for school failure. *Journal of Positive Behavior Interventions*, *7*, 194–204. doi:10.1177/10983007050070040101

Wallace, J. M. J., Goodkind, S., Wallace, C. M., & Bachman, J. G. (2008). Racial, ethnic, and gender differences in school discipline among U.S. high school students: 1991–2005. *Negro Educational Review*, *59*, 47–62.

AUTHOR'S NOTE

The development of this chapter was supported by the Office of Special Education Programs, U.S. Department of Education (#H326S130004). The opinions expressed are those of the authors and do not represent views of the Office or U.S. Department of Education.

Address correspondences to Kent McIntosh, 1235 University of Oregon, Eugene, OR, 97403; E-mail: kentm@uoregon.edu.

MANUSCRIPT

Initial Acceptance: 6/11/2016
Final Acceptance: 6/23/2016