

# The Effects of Student and Teacher Characteristics on Teacher Impressions of – and Responses to – Student Behaviors

Amity Noltemeyer<sup>1</sup>, Claire Kunesh<sup>1</sup>, Cody Hostutler<sup>2</sup>, Patrick Frato<sup>3</sup> & Brianna J. Sarr-Kerman<sup>4</sup>

<sup>1</sup> Educational Psychology Department, Miami University, Oxford, Ohio, USA

<sup>2</sup> Lehigh University, Bethlehem, Pennsylvania, USA

<sup>3</sup> Massillon City Schools, Massillon, Ohio, USA

<sup>4</sup> Centerville City Schools, Centerville, Ohio, USA

Correspondence: Amity Noltemeyer, Miami University, Department of Educational Psychology, 201 McGuffey Hall, Oxford, OH, USA. Tel: 1-513-529-6632. E-mail: anoltemeyer@muohio.edu

Received: April 17, 2012 Accepted: May 7, 2012 Online Published: June 25, 2012

doi:10.5539/ies.v5n4p96

URL: <http://dx.doi.org/10.5539/ies.v5n4p96>

## Abstract

This study examined how student characteristics (e.g., race, gender) and teacher characteristics (e.g., race, gender, years of experience, confidence in behavior management) influence the way teachers perceive and respond to student behaviors in the U.S.A. A rigorous process was used to develop and pilot a survey consisting of questions about a defiant student behavioral incident that might be encountered in a school. This process involved systematically identifying student names that would imply different gender/ethnicity combinations, creating the instrument using these names, expert review, cognitive interviews, and a pilot study using 135 pre-service teachers. After refining the instrument based on feedback from each of these activities, we administered it to 57 practicing teachers. Participants were randomly assigned to one of four scenario conditions, each of which implied a student with a different gender/ethnicity combination (i.e., African American female student, African American male student, European American female student, European American male student). Although some interesting trends in responding emerged based on the implied student race and ethnicity, none were statistically significant. However, teacher characteristics significantly influenced responding, with less experienced teachers being less likely to ignore behaviors – and more likely to address them directly – than their more seasoned counterparts. This adds to the extant knowledge about how teachers in different phases of their careers may interpret and approach classroom situations, and reveals implications for teacher professional development efforts. Further implications, limitations, and future directions are also discussed.

**Keywords:** discipline, teachers, schools, bias, disproportionality, race, gender

## 1. Literature Review

Disciplinary disproportionality – also known as the “discipline gap” – refers to the overrepresentation of some student populations in school-based disciplinary proceedings. One of the most marked examples of this phenomenon in the U.S.A. is the disproportionate representation of African American students in school suspensions, expulsions, and other disciplinary actions. This phenomenon has been consistently documented beginning with the seminal Children’s Defense Fund study (1975), which found that African American students were two to three times as likely to be suspended from school as their White counterparts. Findings indicate that these trends have persisted (e.g., Noltemeyer & Mcloughlin, 2010; Skiba et al., 2011), although the source of the discrepancy continues to be debated. Disproportionality is problematic because exclusionary discipline is associated with negative outcomes including academic difficulties (e.g., Rausch & Skiba, 2004), high school drop-out (e.g., Costenbader & Markson, 1998), participation in criminal activity (e.g., Florida State Department of Education, 1995), and grade retention (e.g., Safer, 1986). Given the severity of these and other potentially negative correlates of suspensions and expulsions, this study aims to explore the role of teacher and student characteristics in explaining disciplinary disproportionality.

### 1.1 Student Characteristics and Discipline

Some have hypothesized that the discipline gap may be attributable to true differences in rates and/or types of behavior across ethnic groups. For example, Wallace, Goodkind, Wallace and Bachman (2008) found that there

were slight but significant differences with respect to drinking, drug use and bringing a gun to school across certain gender-ethnicity combinations. However, they concluded these differences were “relatively small” (p. 53). Wentzel (2002) also found that African American students were significantly less likely to report engaging in prosocial and “responsible” behavior such as trying to do what the teacher asks. However, according to experts in the field and to our knowledge, there is no evidence to suggest that any differences in behavior- if they exist at all- are large enough to account for the glaring discipline gap (Skiba et al., 2011).

On the contrary, research has shown that African American students are disciplined for minor and/or less clearly defined offenses. For example, Skiba, Michael, Nardo, and Peterson (2002) found that African American students were disciplined for more subjective offenses such as disrespect and threats, whereas European American students were disciplined for more objective offenses like vandalism, obscene language and smoking. Some of the offenses for which African American students were referred, such as loitering and “excessive noise,” also appeared to be less serious. Skiba et al. (2011) found similar results revealing that African American students were disciplined more severely for the same or similar types of behaviors engaged in by White students. This referral and consequent punishment of minority students in the absence of serious misbehavior was also found by Vavrus and Cole (2002). The authors analyzed videotapes of class sessions during which referrals leading to suspensions occurred, over a period of five years in a school with a zero-tolerance policy. Their analysis of videotapes, field notes and interview data led them to conclude that such discipline of minority students was largely unjustifiable. Minority students were very often removed from the class by teachers, and therefore suspended, after being singled out as “representatives” when a series of moderately disruptive, non-violent, non-threatening events occurred.

Gender is another aspect of the discipline gap that is important to consider. Male students consistently and disproportionately receive disciplinary actions in schools (Mendez & Knoff, 2003; Skiba, Peterson, & Williams, 1997). In fact, the rate of disciplinary actions for male students has been found to range between two (Mendez & Knoff, 2003) to four (Imich, 1994) times the rate of female students. African American males are at particular risk for receiving harsh disciplinary consequences (Losen & Skiba, 2010), with one study documenting this population to be 16 times more likely to receive corporal punishment than White females (Gregory, 1996). Unlike the research on racial disproportionality, there is some evidence suggesting the differential disciplinary response for male students may be at least in part attributable to actual differences in the range of problem behaviors exhibited between the genders (Skiba, Michael, Nardo, & Peterson, 2002).

### *1.2 Teacher Characteristics and Discipline*

Teachers play an important role in determining which students receive disciplinary referrals. In fact, it has been speculated that teacher bias against minorities may contribute to disproportionality. For example, Fisher, Wallace, & Fenton (2000) found that African American adolescents perceived racial bias as contributing to receiving lower grades and being wrongly disciplined in school, and were more likely to report such biases compared to Latino or Asian American adolescents. Although it is certainly possible that some teachers may demonstrate overtly discriminatory behavior, it is more likely that these biases occur unintentionally (Dovidio, Gaertner, Kawakami, Hodson, 2002). Moule (2009) suggested that such unconscious biases affect all relationships, including the teacher-student relationship. It has been documented that teachers rate African American students as more disruptive and more irresponsible than European American students (Wentzel, 2002). Downey and Pribesh (2004) examined this phenomenon and found that European American teachers rated African American students less favorably than European American students on measures of disruptive behaviors and noncognitive academic skills and habits. This was true for both children in kindergarten (n=12,989) and 8th grade (n=8,881) and held after socioeconomic status, gender and other potential confounds were controlled for. This pattern was not found, however, for African American teachers' ratings of students. Also, African American kindergartners' eagerness to attend school and their reported liking of their teacher was not affected by said teacher's ethnicity, so it does not appear that kindergartners acted differently according to their teacher's ethnicity.

Other studies have failed to find evidence of such biases. For example, one study revealed that the positive behavior of hypothetical Latino students is actually rated more favorably than that of European American students featured in identical scenarios (Chan, Lam, & Covault, 2009). However, Latino students may actually be underrepresented, compared to European American students, in disciplinary referrals from kindergarten until grade 6, at which point they begin to be referred at a disproportionately high rate (Skiba et al., 2011). Therefore, the absence of negative findings in the Chan et al. (2009) study should not be generalized without further investigation. Also, it is possible that the discipline gap does not solely originate in the classroom and teacher-issued office referral, as new evidence suggests that administrator decisions may also play a role (Skiba et al., 2011).

There are also contradictory findings related to gender. Shepherd (2011) found some evidence of teacher bias in a study revealing that teachers evaluated spoken responses of male students significantly less favorably than female students (even when the responses were identically worded). However, other studies have suggested that bias may not play a role. Skiba et al. (2002), as previously mentioned, found evidence that there may be actual differences in male student behavior that explain the discipline gap, which implies a minimized role of bias. Also, Green, Shriberg, & Farber (2008) found no differences in teacher responses to student behaviors (presented in vignettes) based on the gender of the student.

The role of teachers in the discipline gap is further complicated by the demographic characteristics of teachers themselves. Young teachers report that they would enact harsher punishments than older teachers (Salvano-Pardieu, Fontaine, Bouazzaoui, & Florer, 2009). In addition, male and female teachers may punish misbehavior differently, depending on a student's history of misbehavior and level of academic performance (Salvano-Pardieu et al., 2009). For example, although Green et al. (2008) found no differences in teacher responding to behaviors based on *student* gender, they did find differences based on *teacher* gender. Specifically, female teachers rated the behaviors as more severe than their male counterparts.

Furthermore, there may be interactions between teacher characteristics and student characteristics. For example, teacher age and various student characteristics interact, leading to different effects on students' grades (Mullola et al., 2011). In addition, research suggests that child temperament and teacher temperament may interact to influence teacher responses to child behavior (Oren & Jones, 2009). With respect to academic achievement, teacher characteristics such as years of experience have a different effect on European American and African American students (Kukla-Acevedo, 2009). Interaction effects between teacher and student ethnicity, as well as gender, have also been found for behavioral ratings such as a student's disruptiveness and inattention (Dee, 2005).

### 1.3 Rationale for Study

The existence of racial and gender disproportionality in discipline in the U.S.A. is indisputable. Although some research suggests there may be some differences in student misbehavior based on gender, no research has successfully attributed racial disparities to true differences in misbehavior or socioeconomic status. Therefore, it appears that teachers and/or administrators may react differently to students depending on their ethnicity, and perhaps also their gender. It also is possible that teacher characteristics such as years of experience further influence disciplinary choices and that there may be interaction effects amongst these characteristics and students' ethnicity and/or gender.

Due to social desirability and the often unconscious nature of bias (Dovidio, Gaertner, Kawakami, & Hodson, 2002), it can be difficult to accurately determine the extent to which bias is occurring in a given situation. Given this challenge, one line of inquiry designed to examine bias uses fictitious scenarios or situations designed to determine if there are differences in thoughts or responding based on a student's implied race or gender. Because respondents are unaware that race and/or gender is being manipulated, it is assumed that findings more accurately approximate reality than if the respondent were asked outright about biases.

A landmark study in racial discrimination by Bertrand and Mullainathan (2004) assessed bias in this indirect way. The researchers responded to real job listings with fictitious resumes, randomly assigning to each resume a typically African American or typically European American name. The researchers found that employers were significantly more likely to follow up on the resumes of individuals with typically European American names than those with typically African American names. This bias held across skill levels, occupational sectors, locations and types of businesses, including Equal Opportunity Employers. Although the chosen African American names were associated with lower education levels and the selected European American names with higher-than-average levels of education, rate of callback was not correlated with education across the individual names in the study.

Similar methodology has been utilized in a select number of educational studies in an attempt to identify or quantify biases in the school setting. However, such investigations have been limited in quantity and their findings have been somewhat contradictory. Some previous investigations in the school population have found a lack of bias. For example, Chang and Sue's (2003) research on bias in discipline found that across three scenarios teachers did not show a bias against or for African American and European American students for any type of problematic behavior. However, the influence of cultural miscommunication is suggested in work by Tyler, Boykin, and Walton (2006). Teachers read four scenarios, each of which depicted a student displaying behavior valued either by African American or mainstream European American culture. Teachers rated students presenting mainstream cultural values as having significantly higher classroom motivation and academic

achievement than students exhibiting African American values. In sum, these studies suggest that although teachers may not exhibit outright racial bias, they may show a preference for traditional European American-valued behaviors over conduct that may be more typical or valued by African Americans. These findings suggest that further research is needed to explore the interactions between race, gender and behavior and how they influence teachers' impressions towards student behavior (Chang & Sue, 2003).

#### *1.4 Purpose and Research Questions*

This study was designed to further explore the complicated interplay between student and teacher demographic characteristics and disciplinary decisions. Specifically, we sought to answer the following research questions:

- 1) Do teachers report different perceptions/responses to an ambiguous behavioral incident based upon the implied race and gender of the student involved?
- 2) Do teachers report different perceptions/responses to an ambiguous behavioral incident based upon various teacher demographic factors?

## **2. Method**

### *2.1 Instrument Development*

In order to answer these research questions, we devised a scenario about a child's behavior that might be encountered in a school. Four versions of the scenario were created, with each version differing only in the name of the child (i.e., distinctively African American male name, distinctively African American female name, distinctively European American male name, distinctively European American female name).

In order to select the names to be used in the scenario, we first created a list of 16 potential names (four for each ethnicity/gender combination). These names were drawn from Fryer and Levitt's (2004) research on first names, as explained in Levitt and Dubner (2005). The authors examined baby name data from the California Department of Health Services' Office of Vital Records. They determined how often African American and European American parents gave their children a first name and used these data to create lists of the top 20 "Blackest" and "Whitest" names. Some names, like Molly, were almost exclusively given by European American parents. Other names, like DeShawn and Shanice, were almost always chosen by African American parents. In an attempt to minimize bias based on perceived socioeconomic status, we chose European American names that were identified as being popular with both low- and middle-income parents (as opposed to high-income parents). In addition, we selected the names to represent African American children from the bottom half of the top 20 "Blackest" names in an attempt to not be overly prejudicial or alert participants to the true nature/purpose of the questionnaire.

Although the 16 selected names were distinctively "White" and "Black" according to the data collected by Fryer and Levitt (2004), we wanted to confirm that our study participants would in fact associate the selected names with the intended race. Consequently, we used an online questionnaire to confirm the racial typicality of each name. Thirty-five respondents drawn from a convenience sample completed the questionnaire, which provided each name and asked participants to (a) select which race they think the person with that name would be, and (b) indicate the degree of confidence in their response using a Likert scale. Six names were unanimously rated as being African American or European American: Cody (European American), Darius (African American), Emily (European American), Jada (African American), Megan (European American), and Tyler (European American). Of these unanimous ratings, the following four were most strongly (in terms of confidence) rated in the hypothesized direction: Cody, Darius, Emily, and Jada. Based on these findings, these four names were chosen for the four versions of the scenario. This procedure for name selection is similar to that used by Chan, Lam, and Covault (2009), although we obtained a participant sample over three times the size of the one used in their study.

After selecting the names, the scenario and questionnaire were created. The scenario consisted of an ambiguously defiant, non-disruptive response from the student. The questions were developed to assess how the participants would respond to the situation, the perceived inappropriateness of the behavior, what they felt the cause and likelihood of reoccurrence of the behavior were, and how they would attempt to prevent the behavior in the future. These questions were asked to determine locus of causality and the severity with which the teacher interpreted and responded to the event. The questionnaire concluded with demographic items concerning the participants' school standing, gender, ethnicity and familiarity with behavior management in the classroom. The questionnaire and scenario in the four conditions differed only by the name and gender-related pronouns used.

Following scenario and questionnaire development, an expert review process was used to determine if improvements or changes were warranted. Specifically, four reviewers were asked five questions to assess to the

scenario itself, the questionnaire items, the format of the instrument, and the demographic items. For example, they were prompted to provide comments on the clarity of the instrument, existence of leading questions, and authenticity of the situation. Based on the feedback from these reviewers – who were each experts in research design, questionnaire development, or the content of the questionnaire – several changes were made to improve the quality of the questionnaire. For example, we added an additional response option, changed the wording on multiple questions, and decided to allow participants to mark more than one response on some items.

Next, cognitive interviews (i.e., “think alouds”) were conducted with three graduate students and two educational professionals in order to better understand how participants would interpret and respond to the scenario and items. Cognitive interviewing is a tool used to (a) assess the degree to which respondents understand what is being sought in questionnaire items and respond in a manner that the researcher intended, and (b) capture respondents’ thought processes as they progress through the items on the questionnaire (Beatty & Willis, 2007; Collins, 2003). Cognitive interviews can reveal potential problems with the questionnaire that can be addressed before it is utilized with a broader population. Based on questions raised during the cognitive interviews, we changed several of the response options to enhance the clarity of the items. In addition, the cognitive interviewees overwhelmingly reported a need for information about the age of the child. In an attempt to minimize the likelihood that participants would attribute vastly different ages to the child, and consequently respond differently, we decided to add to the scenario that the child was in the 7th grade. There was general consensus between three of the researchers and several cognitive interviewees that the scenario was realistic for that grade level (the fourth and fifth researchers were not involved until after the survey development).

After the scenario and survey were adapted based on the cognitive interviewee feedback, we conducted a pilot study to further explore and refine the instrument. The scenario and survey were administered to 135 pre-service teachers at a university in the United States. Several limitations with the survey design emerged as a result of this pilot study, and consequently several additional changes were made to the survey items. First, it became evident that the four point Likert scales originally used on questions two (assessing appropriateness of behavior) and four (assessing the likelihood of the behavior reoccurring) were not sensitive enough to adequately discern differences between conditions. Most participants selected the same response option, with a few selecting a second response option and none selecting the other two options. Thus, the Likert scales on questions two and four were extended from four points to eight points to provide more sensitivity. After analyzing the pilot data, we also decided to choose an open response format to question five, which asks the participant to describe what strategy they would use in order to reduce the likelihood of this situation reoccurring. This was done because the results of the pilot survey (which provided response options to select from) suggested that participants overwhelmingly selected the most socially desirable response. It was anticipated that the open-ended nature of this item would help minimize the chances of the participant choosing an answer based on its social desirability and more closely approximate what they would actually do in practice.

See Appendix A for the final scenario and questionnaire items for Jada. The other scenario versions were identical except the names Cody, Darius, and Emily were, used and the gender-related pronouns were changed accordingly.

## 2.2 Participants

Table 1. Summary of participant demographics

		N	%
Gender	Male	8	14.0
	Female	49	86.0
Race	Asian	1	1.8
	African American/Black	2	3.6
	White	52	94.5
Teaching Experience	2-5	5	8.8
	6-10	9	15.8
	11-15	14	24.6
	16-20	12	21.1
	21+	17	29.8

Participants were 57 teachers employed in four suburban schools located in a state in the Midwest portion of the

U.S.A. Three of the four schools were elementary schools (two were grades k-4 and one was grades 2-5), and the fourth school was a middle school (grades 6-8). The majority of the participants were White, female teachers. See Table 1 for participant demographic information.

### *2.3 Procedures*

One hundred and thirty two teachers in the four participating schools were invited to participate in the study via email. Those who chose to participate followed a link in the email that took them to one of the four survey versions: distinctively African American male name, distinctively African American female name, distinctively European American male name, or distinctively European American female name. The survey version assigned to each participant was randomly selected using a random numbers generator, and it is assumed that participants were unaware that other survey versions existed. Fifty seven teachers completed the survey for an overall response rate of 43.18%.

The study was approved by an Institutional Review Board before any procedures were carried out. This project involved a lack of full disclosure regarding the purpose of the study. As noted above, bias is often subtle and covert. Consequently, if the full purpose of the study were revealed, it is likely that participants may have responded in a more positive manner than they would respond in the actual situation. Because these were fictitious scenarios – rather than true interactions with actual children – the use of deception was not deemed to pose a significant threat. Consequently, participants were told that the purpose of the study was to examine decision-making regarding children's behaviors in school, with no mention of race or gender issues. We did offer participants the opportunity to be debriefed about the true purpose of the study, although no one utilized that opportunity.

## **3. Data Analysis**

### *3.1 Quantitative*

All quantitative questionnaire data were entered into SPSS. Data entry was spot checked for accuracy. The quantitative data were analyzed in two ways. First, descriptive statistics were run to determine the frequency and percentage of responses for the questionnaire items and demographic items. Also, to answer the research question, a chi-square test of independence was conducted on each item to see if the observed values differed significantly from the expected values based on the implied gender, ethnicity, and gender/ethnicity combination of the student described in the questionnaire. Chi-square tests were used rather than ANOVAs or t-tests due to the nominal and ordinal nature of the data. For some items, data from two or more response categories was merged together because there were so few responses in one category that it violated the general rule that chi-square analyses should only be conducted when the number of responses in each category on a question is within 1.5 times the number of responses in other categories.

### *3.2 Qualitative*

Qualitative data were also collected and analyzed. The qualitative responses were first typed into a word processor. A preliminary examination of the responses for the first part of question five and question six suggested that the strategies to reduce the behavior as well as the proposed reasons for the child's behavior fell under several different "themes." Consequently, we used a "long table" approach (Krueger & Casey, 2000) to code these data into thematic categories. In this approach, responses to each question were read aloud and assigned a categorical descriptor, which was arrived at through discussion and consensus by two of us. As statements were coded, they were cut from the written transcripts and sorted into labeled categories underneath that question on a long table. If a response to question five or six fit into more than one category, it was included in both.

The comments section of question one and the second part of question five asked the participants to describe the reasons behind their chosen response to the behavior as well as their strategy for preventing the behavior in the future (see Appendix A). We worked collaboratively to find themes for these two items, but no common trends emerged.

Following these initial analyses, a typed document was prepared that provided a written record of the categories and their associated comments for each question. These were again reviewed and any discrepancies reconciled until 100% consensus was achieved.

## **4. Results**

### *4.1 Quantitative Results*

No statistically significant differences in responding to any of the items were found based on the survey version used (i.e., Cody, Darius, Emily, Jada). The responses to most items followed a similar pattern. For example,

across all survey versions respondents were more likely to report that they would address the behavior, would not involve others, and viewed ignoring the teacher as the most inappropriate aspect of the behavior. However, a few interesting findings did emerge that suggested differential responding based on the survey version. For example, although the results were not statistically significant, respondents more frequently rated the behavior as less inappropriate for Jada and Darius; however, they more frequently rated the behavior as more inappropriate for Emily and Cody (see Figure 1).

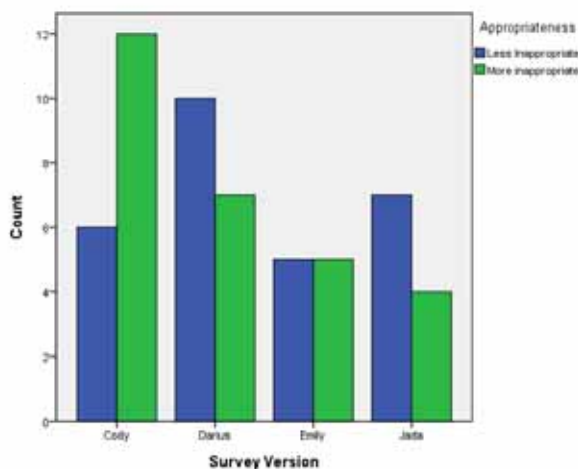


Figure 1. Perceived appropriateness of the behavior based on survey version

Analyses of differences in responding based on teacher characteristics yielded more fruitful findings. Years of teaching experience was an important factor influencing teacher perceptions and anticipated actions regarding the behavioral situation. For example, respondents with fewer years of teaching experience were significantly less likely to ignore the behaviors than were respondents with more years experience,  $F(3) = 7.811, p = .050$  (see Figure 2). Consistent with this finding, these less experienced teachers were significantly more likely to report that they would address the behaviors directly than were their more seasoned counterparts,  $F(3) = 13.775, p = .003$  (see Figure 3).

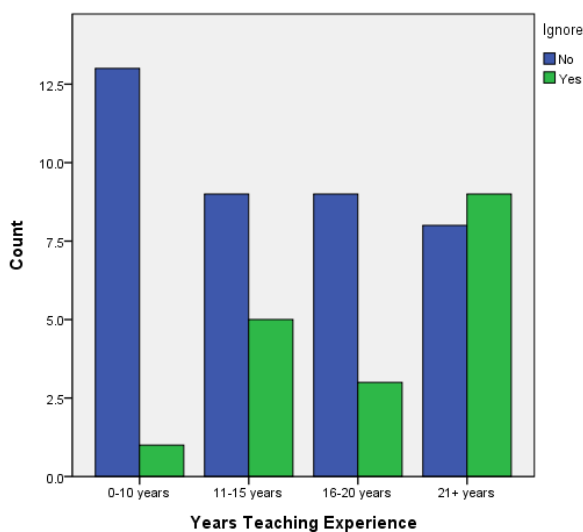


Figure 2. Number of respondents reporting that they would ignore the behavior based on number of years experience

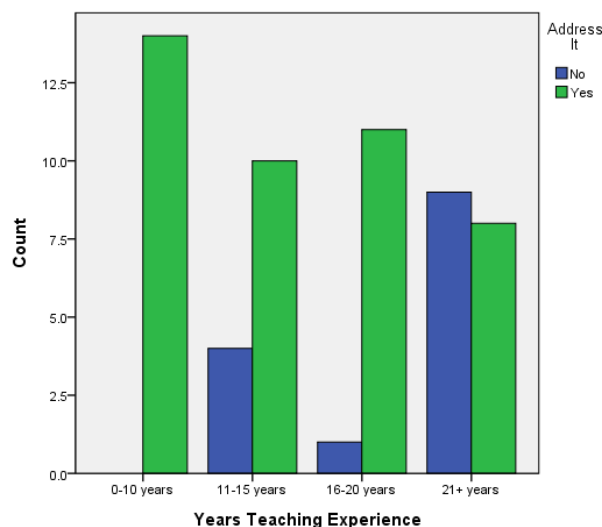


Figure 3. Number of respondents reporting that they would address the behavior based on number of years experience

In addition to years of experience, other factors may be related to why teachers responded as they did. For example, although not statistically significant, teachers who reported lower confidence levels with regard to their behavior management skills more often described the behaviors as being less appropriate than teachers who reported higher confidence levels (see Figure 4).

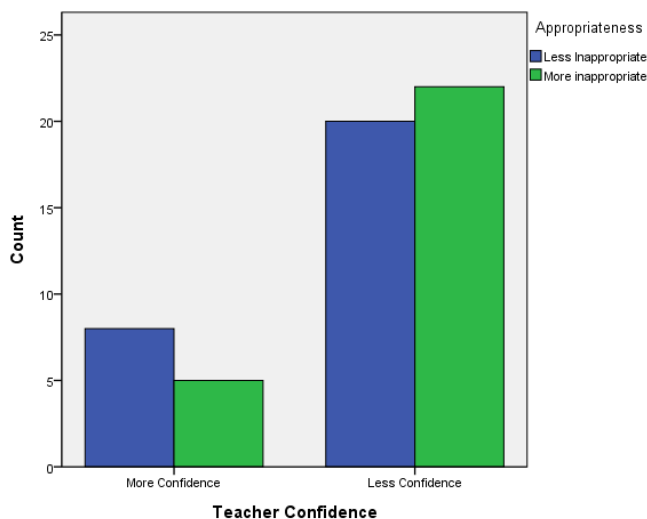


Figure 4. Perceived appropriateness of the behavior based on teacher confidence

#### 4.2 Qualitative Results

The qualitative analyses also revealed several interesting findings. Overall, results suggested more trends based on gender than based on ethnicity. While around 20 percent of teachers reading either male version attributed the behavior to home life (e.g., "...issues at home"), none did so for either female version. In addition, teachers were more likely to attribute the behaviors to age for the female scenario versions (e.g., "She is a typical teenage girl"). Respondents also only mentioned "having a bad day" as an explanation for the behavior with the two male



students (e.g., “He could have had a rough morning”). Finally, respondents only mentioned somatic reasons for the behavior in the Emily and Jada versions of the scenario (e.g., “perhaps she has a hearing problem, or an attentional issue”).

Interestingly, in terms of how they would respond to the behavior, teachers were more likely to mention the use of discipline or punishment when responding to the female survey versions (in fact, they were 6 times more likely to mention discipline in the Emily condition than in the Cody condition). In contrast, they were more likely to mention the preventative strategy of increasing teacher presence in the hallway when responding to the male survey versions. See figures 5 and 6 for a full breakdown of qualitative results.

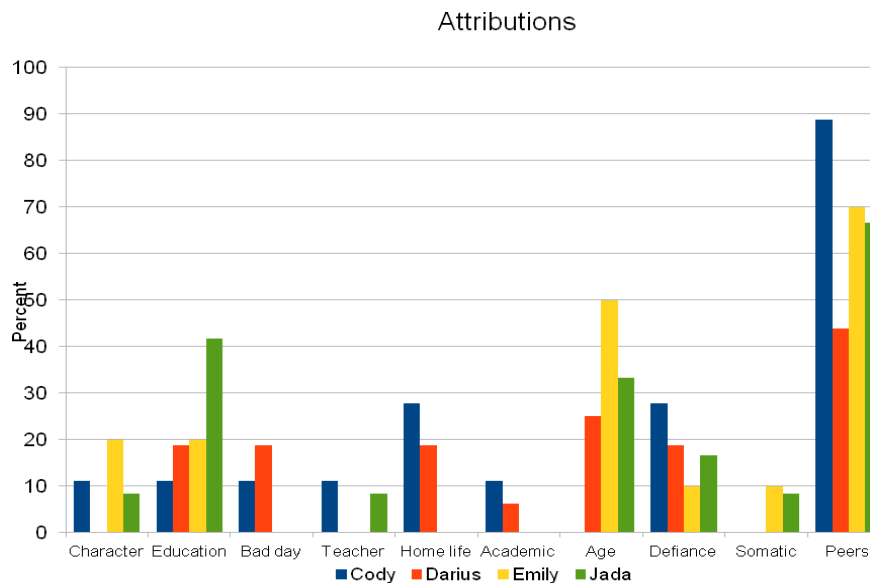


Figure 5. Perceived causes of behavior based on survey version

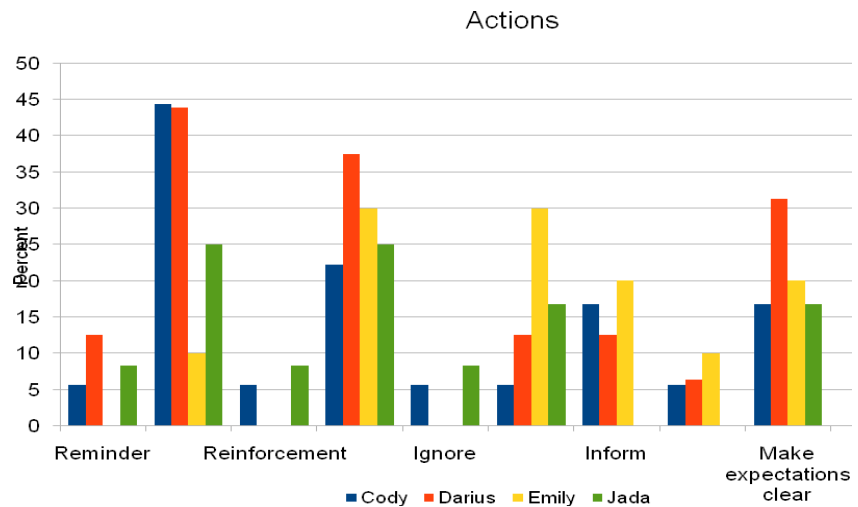


Figure 6. Reported response to behavior based on survey version

**5. Discussion**

*5.1 Summary*

This study was designed to identify the degree to which teacher and student characteristics affect disciplinary perceptions and decision making by practicing teachers. This topic deserves attention given the issue of the “discipline gap” and the overwhelmingly European American female teaching force (National Education Association, 2010) in the U.S.A. Several interesting trends in the data emerged. First, there were no statistically

significant differences in teacher responding based on the implied gender and race of the student involved in the behavioral incident. This is consistent with findings from Chan et al. (2009) suggesting a lack of racial bias, as well as findings from Green et al. (2008) suggesting a lack of gender bias when using vignettes. This may indicate that bias based on racial and gender stereotypes alone may not be influencing teachers' decision making, implicating cultural mismatch or other factors instead. Alternatively, it could suggest that practicing teachers were sensitive to being perceived as biased. Respondents may have been aware of the discipline gap between African-American and European American students or between male and female students, and responded out of concern for the way their answers would be interpreted. On the other hand, we minimized the likelihood of responding in socially desirable manner by using deception (i.e., participants were not told we were interested in these factors).

Although the findings should be interpreted cautiously due to non-significance, the teachers in the study did more frequently rate the behavior of European American students as more inappropriate than that of African American students. The finding also raises the possibility that inappropriate behavior may be viewed as more normative (i.e., normal or typical) for African American students and non-normative for European American students. Perhaps teachers held the notion that African American students should or do act a certain way, and consequently were more likely to view the behavior as "normal." This response trend may also suggest that some practicing teachers have higher expectations for European American students than African American students, a finding consistent with previous work (e.g., Shepherd, 2011). This could be concerning given findings that African American students perform best with teachers who hold high expectations (e.g., Ware, 2006).

Results also indicated that several teacher variables significantly contribute to teachers' responses to – and perceptions of – behavior. For instance, practicing teachers with fewer years of teaching experience were significantly less likely to ignore behaviors than respondents with more years of experience. Further, less experienced teachers were significantly more likely to indicate that they would address behavior directly, compared to more experienced teachers. These findings indicate that novice teachers may be more reactive in response to discipline issues; in contrast, experienced teachers may be less reactive, more patient, and potentially expend less energy in addressing discipline issues. Although older teachers are not always more experienced, this does seem aligned with previous findings that younger teachers are more likely to use harsher discipline strategies (e.g., Salvano-Pardieu et al., 2009). This adds to the extant knowledge about how teachers in different phases of their careers may interpret and approach classroom situations.

Qualitative analysis also revealed differences in behavioral attributions based on student gender. For example, the results suggested that the behaviors of male students were attributed to issues at home more often than were the behaviors of female students. This suggests that teachers may believe that the home environment impacts male students more than female students or that female students are more resilient in the face of home stressors. On the other hand, somatic complaints were only attributed to female students as a cause for behavior. This indicates that teachers may believe that female students are more affected by stable, internal issues than their male peers. It also may suggest that practicing teachers consider female students to be more sensitive to somatic complaints than male students.

Teachers were also more likely to mention the use of discipline or punishment when responding to the female survey versions. It is possible that teachers have higher behavioral expectations for female students. Moreover, this suggests that teachers may view inappropriate female behavior as more unusual or out of place. Finally, respondents were more likely to mention the preventative strategy of increasing teacher presence in the hallway when responding to the male survey versions. This indicates that teachers may be more aware of and proactive in dealing with male student behaviors. It is also possible that teachers may see and respond to more inappropriate behavior among male students because they appear to be looking for and expecting it.

### *5.2 Implications*

There are several implications of these results for educators seeking to implement fair and equitable discipline in their schools. First, the non-significant findings regarding the influence of a student's race and ethnicity on perceptions of the student's behavior provides some evidence that implicit bias may not play as large of a role in the discipline gap as previously thought. However, it remains undeniable that disproportionality does occur in classroom discipline. Given the non-significant trends in our data, the qualitative findings, and findings of similar research, it is still a possibility that implicit bias may be contribute to some degree to the discipline gap. Consequently, it is important for teachers to be aware of this phenomenon and consider its causes and consequences.

In addition to recognizing their biases, teachers should also remain careful about the expectations and

attributions they make about classroom behaviors. Expecting more from European American students may send the message that not as much can be expected from the African American students, an effect that could contribute to negative self-perceptions and ultimately may increase behavior more consistent with these perceptions. Expecting poor behavior may result in the phenomenon of the self-fulfilling prophecy, whereby behavior that didn't exist before emerges due to the expectation that it will. In other words, students who sense that teachers believe that they are poorly behaved may eventually adopt a negative identity, where one did not previously exist. In contrast, having high expectations for all students, directly teaching appropriate behavior to all students, and then equally enforcing expectations, helps students from *all cultural backgrounds*.

The findings also suggest that school administrators may want to consider the source of behavioral referrals before making disciplinary decisions. According to this study, less experienced teachers make more frequent and reactive discipline referrals. For this reason, it could be valuable to provide these teachers with mentors who are more experienced and have greater confidence in their own classroom management abilities. Doing so may result in more a more proactive and fair disciplinary environment. This is particularly relevant for building principals, who handle most office disciplinary referrals and often have the potential to implement mentoring or professional development systems to foster appropriate classroom management. Educators should also anticipate the potential for a domino effect to occur in the classrooms of inexperienced teachers with low confidence. That is, these teachers may perceive more inappropriate behavior than actually exists and then react more confrontationally to behavior. This could result in poorly managed classrooms, frequent classroom disruptions, increased office referrals, and lower academic outcomes. The potential for the exacerbation and creation of behavioral issues affirms the need to provide these teachers with intense support in order to increase confidence in their classroom management ability and to help them become less reactive.

### 5.3 Limitations

Although the sample size was adequate, the suboptimal response rate raises questions about the degree to which those who responded to the survey differed from those who did not. In addition, the study was limited by the homogeneous nature of the sample. Participants were mostly European females teaching at Midwestern schools in the U.S.A. Including more male and ethnically diverse respondents would have allowed us to gauge whether findings generalize to all teachers or whether other factors such as gender and background are more relevant. Further, given a large enough sample, we also could have gauged whether African American teachers responded differently (than their European American counterparts) to behavior exhibited by European American and African American students. Finally, due to limited access to participants, our sample was a mixture of both elementary and middle school teachers; since the student involved in the scenario was a seventh grade student, a limitation is that not all participants taught in this grade level. However, it would be anticipated that in their pre-service training, teachers do learn about developmental differences in classroom management.

Further, data collection via survey may have allowed teachers time to consider responses they perceived as pedagogically correct or even politically correct. Teachers in a real life situation would not have the same amount of time or emotional detachment from student behavior. The difference then between the responses given on surveys and actual responses that would be observed in practice could be as stark as the difference between knowing what to do and actually doing it.

It should also be noted that teachers were only asked to respond to vignettes, but did not have any background information about the child or the relationship between the student and teacher. In a real life situation, teacher responses would be tied not only to the behavior, but also to their history with a student. Without the addition of this information, the survey lacked authenticity. Finally, it is possible that participants, in spite of pretesting, did not associate the given name with the intended ethnicity.

### 5.4 Future Directions

Due to homogeneity of the participants in this study, future studies could explore differences in how respondents of different ethnicities, genders, grade-levels, and school compositions (e.g., racially diverse or not racially diverse) respond to behavioral concerns. Further inquiry might also include a scale to measure the degree of intervention (e.g., amount of punishment or reinforcement) that could uncover more significant variations in responses to behaviors. Future studies might also consider adding a second version of the questionnaire to be administered to *students* to explore how they believe a teacher should respond to this situation. This would allow for direct comparison of the impressions of students and those of teachers. Student input would allow investigators to measure such factors as the perceived justice of teacher behavior, which types of teacher responses they understand better, and which types of teachers and responses they believe are more conducive to learning. Students also could be asked how their teachers actually *do* respond to inappropriate behaviors, which

could allow researchers to see the degree of match between what teachers report they *would* do in these circumstances and what students report their teachers actually *do*.

One finding of this study suggested that teachers with fewer years of experience were less likely to ignore behaviors and more likely than their more seasoned colleagues to address them directly. Future studies need to explore whether less experienced teachers are reacting in a way that reduces or exacerbates such behaviors. Future investigation might focus on whether less experienced teachers eliminate inappropriate behaviors through direct action or whether more experienced teachers understand behavior management better and are able to control behavior more effectively and with less energy.

Finally, it would be important for future research to examine the role social desirability plays, if any, in responding. This could be achieved by adding additional scenarios that outright specify the race and gender of the child, in addition to those that merely imply it by name. If participants respond significantly more positively to minority students or male students in those that make these features more salient, it may indicate an influence of wanting to respond in a socially desirable way (e.g., not wanting to appear biased).

Future research in the aforementioned directions could add to our understanding of the factors that contribute to the discipline gap. This is important in order to design equitable learning environments where all students have access to the curriculum and can learn and develop in meaningful ways.

### Acknowledgements

We would like to acknowledge the assistance of the expert reviewers, cognitive interviewees, and participants involved in this research. In addition, we thank Jillian Black, graduate student at Miami University, for her help proofreading and formatting this article.

### References

- Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, 71, 287-311. <http://dx.doi.org/10.1093/poq/nfm006>
- Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *The American Economic Review*, 94(4), 991-1013. <http://dx.doi.org/10.2139/ssrn.422902>
- Chan, K. C., Lam, S., & Covault, J. M. (2009). White American pre-service teachers' judgments of Anglo and Hispanic student behaviors. *Intercultural Education*, 20(1), 61-70. <http://dx.doi.org/10.1080/14675980802700821>
- Chang, D. F., & Sue, S. (2003). The effects of race and problem type on teachers' assessments of student behavior. *Journal of Consulting and Clinical Psychology*, 71, 235-242. <http://dx.doi.org/10.1037/0022-006X.71.2.235>
- Children's Defense Fund. (1975). *School suspensions: Are they helping children?* Cambridge, MA: Washington Research Project.
- Collins, C. (2003). Pre-testing survey instruments: An overview of cognitive methods. *Quality of Life Research*, 12, 229-238. <http://dx.doi.org/10.1023/A:1023254226592>
- Costenbader, V., & Markson, S. (1998). School suspension: A study with secondary school students. *Journal of School Psychology*, 36(1), 59-82. [http://dx.doi.org/10.1016/S0022-4405\(97\)00050-2](http://dx.doi.org/10.1016/S0022-4405(97)00050-2)
- Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *The American Economic Review*, 95(2), 158-165. <http://dx.doi.org/10.1257/000282805774670446>
- Dovidio, J. F., Gaertner, S. L., Kawakami, K., & Hodson, G. (2002). Why can't we just get along? Interpersonal biases and interracial distrust. *Cultural Diversity and Ethnic Minority Psychology*, 8(2), 88-102. <http://dx.doi.org/10.1037/1099-9809.8.2.88>
- Downey, D. B., & Pribesh, S. (2004). When race matters: Teachers' evaluations of students' classroom behavior. *Sociology of Education*, 77, 267-282. <http://dx.doi.org/10.1177/003804070407700401>
- Fisher, C.B., Wallace, S.A., & Fenton, R.E. (2000). Discrimination distress during adolescence. *Journal of Youth and Adolescence*, 29, 679-695. <http://dx.doi.org/10.1023/A:1026455906512>
- Florida State Department of Education. (1995). *Florida School Discipline Study: 1994 Juvenile Justice Reform Act*. Tallahassee, FL: Florida State Department of Education. (ERIC Document Reproduction Service No. ED384981).

- Fryer, R. G., Jr., & Levitt, S. D. (2004). The causes and consequences of distinctively black names. *Quarterly Journal of Economics*, 119(3), 767-805. <http://dx.doi.org/10.1162/0033553041502180>.
- Green, S.P., Shriberg, D., & Farber, S.L. (2008). What's gender got to do with it? Teachers' perceptions of situation severity and requests for assistance. *Journal of Educational and Psychological Consultation*, 18, 346-373. <http://dx.doi.org/10.1080/10474410802463288>
- Gregory, J. F. (1996). The crime of punishment: Racial and gender disparities in the use of corporal punishment in the U.S. Public Schools. *Journal of Negro Education*, 64, 454-462. <http://dx.doi.org/10.2307/2967267>
- Imich, A. J. (1994). Exclusions from school: Current trends and issues. *Educational Research*, 36(1), 3-11. <http://dx.doi.org/10.1080/0013188940360101>
- Kaiser, C. R., & Pratt-Hyatt, J. S. (2009). Distributing prejudice unequally: Do Whites direct their prejudice toward strongly identified minorities? *Journal of Personality and Social Psychology*, 96, 432-445. <http://dx.doi.org/10.1037/a0012877>
- Krueger, R. A., & Casey, M. A. (2000). *Focus groups: A practical guide for applied research* (3rd ed). Thousand Oaks, CA: SAGE.
- Kukla-Acevedo, S. (2009). Do teacher characteristics matter? New results on the effects of teacher preparation on student achievement. *Economics of Education Review*, 28, 49-57. <http://dx.doi.org/10.1016/j.econedurev.2007.10.007>
- Levitt, S. D., & Dubner, S. J. (2005). *Freakonomics: A rogue economist explores the hidden side of everything*. New York, NY: HarperCollins.
- Losen, D.J., & Skiba, R.J. (2010). *Suspended Education: Urban Middle Schools in Crisis*. Retrieved from [http://civilrightsproject.ucla.edu/research/k-12-education/school-discipline/suspended-education-urban-middle-schools-in-crisis/Suspended-Education\\_FINAL-2.pdf](http://civilrightsproject.ucla.edu/research/k-12-education/school-discipline/suspended-education-urban-middle-schools-in-crisis/Suspended-Education_FINAL-2.pdf) (May 30, 2011)
- Mendez, L.M.R., & Knoff, H.M. (2003). Who gets suspended from school and why: A demographic analysis of schools and disciplinary infractions in a large school district. *Education and Treatment of Children*, 26(1), 30-51.
- Moule, J. (2009). Understanding unconscious bias and unintentional racism. *Phi Delta Kappan*, 90, 320-326. Retrieved from <http://people.uncw.edu/browna/documents/UnderstandingUnconsciousBiasUnintentionalRacism.pdf>
- Mullola, S., Jokela, M., Ravaja, N., Lipsanen, J., Hintsanen, M., Alatupa, S., & Keltikangas-Järvinen, L. (2011). Associations of student temperament and educational competence with academic achievement: The role of teacher age and teacher and student gender. *Teaching and Teacher Education*, 27, 942-951. <http://dx.doi.org/10.1016/j.tate.2011.03.005>
- National Education Association. (2010). *Status of the American public school teacher 2005-2006* (Item No. 3259-200). Retrieved from <http://www.nea.org/assets/docs/HE/2005-06StatusTextandAppendixA.pdf> (May 26, 2011)
- Noltemeyer, A. L., & Mcloughlin, C. S. (2010a). Changes in exclusionary discipline rates and disciplinary disproportionality over time. *International Journal of Special Education*, 25(1), 59-70. Retrieved from <http://www.eric.ed.gov/PDFS/EJ890566.pdf>
- Oren, M., & Jones, I. (2009). The Relationships between Child Temperament, Teacher-Child Relationships, and Teacher-Child Interactions. *International Education Studies*, 2, 122-133. Retrieved from <http://www.ccsenet.org/journal/index.php/ies/article/view/3863>
- Rausch, M. K., & Skiba, R. (2004). Disproportionality in School Discipline among Minority Students in Indiana: Description and Analysis. Children Left Behind Policy Briefs. Supplementary Analysis 2-A. Bloomington, IN: Center for Evaluation and Education Policy. (ERIC Document Reproduction Service No. ED488897).
- Safer, D. J. (1986). Nonpromotion correlates and outcomes at different grade levels. *Journal of Learning Disabilities*, 19, 500-503. <http://dx.doi.org/10.1177/002221948601900812>
- Salvano-Pardieu, V., Fontaine, R., Bouazzaoui, B., & Florer, F. (2009). Teachers' sanction in the classroom: Effect of age, experience, gender and academic context. *Teacher and Teacher Education*, 25, 1-11.
- Shepherd, M. A. (2011). Effects of ethnicity and gender on teachers' evaluations of students' spoken responses. *Urban Education*, 46(5), 1011-1028. <http://dx.doi.org/10.1177/0042085911400325>

- Skiba, R. J., Horner, R. H., Chung, C.G., Rausch, M. K., May, S. L., & Tobin, T. (2011). Race is not neutral: A national investigation of African American and Latino disproportionality in school discipline. *School Psychology Review*, 40(1), 85-107.
- Skiba, R. J., Peterson, R. L., & Williams, T. (1997). Office referrals and suspension: Disciplinary intervention in middle schools. *Education and Treatment of Children*, 20, 295-316. Retrieved from www.ebscohost.com
- Skiba, R.J., Michael, R.S., Nardo, A., & Peterson, R.L. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *Urban Review*, 34(4), 317-342. <http://dx.doi.org/10.1023/A:1021320817372>
- Tyler, K. M., Boykin, A. W., & Walton, T. R. (2006). Cultural considerations in teachers' perceptions of student classroom behavior and achievement. *Teaching and Teacher Education*, 22, 998-1005. <http://dx.doi.org/10.1016/j.tate.2006.04.017>
- Vavrus, F., & Cole, K. (2002). "I didn't do nothing": The discursive construction of suspension. *The Urban Review*, 34, 87-111. <http://dx.doi.org/10.1023/A:1015375215801>
- Wallace, J. M., Jr, 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. *The Negro Educational Review*, 59, 47-62. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2678799/pdf/nihms50094.pdf>
- Ware, F. (2006). Warm demander pedagogy: Culturally responsive teaching that supports a culture of achievement for African American students. *Urban Education*, 41, 427-256. <http://dx.doi.org/10.1177/0042085906289710>
- Wentzel, K. R. (2002). Are effective teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development*, 73(1), 287-301. <http://dx.doi.org/10.1111/1467-8624.00406>

## Appendix A

DIRECTIONS: Please read the situation described in the following paragraph.

The morning bell just rang, indicating all students should be in their classrooms. You look around and notice that Jada, a 7<sup>th</sup> grader, is the only student not in your class. As you walk over to shut the door and begin class, you notice Jada and several other students laughing and talking in the hallway. You remind Jada that the bell has rung and she should be in class. She turns to her friends, rolls her eyes, and shakes her head back and forth. Then, as Jada walks very slowly toward the door, you hear her mumble something in a negative tone of voice. When you ask Jada what she said, she continues to walk to the door without responding.

DIRECTIONS: Assume that you are the teacher in the situation you just read. With the information you have, please answer questions 1-7 below. You may decline to answer any question.

1. Which of the following choices describe how you would most likely respond to this situation? (Select all that apply)

- Ignore it (e.g., begin class).
- Address it with Jada directly (e.g., talk with her immediately or sit down with her after class).
- Involve other individuals (e.g., send Jada to the office, write an office referral, or notify Jada's parents).

Please describe why you would be most likely to respond in this way:

2. How would you describe Jada's behavior? (Select one response only)

- Extremely appropriate
- Very appropriate
- Somewhat appropriate
- Slightly appropriate
- Slightly inappropriate
- Somewhat inappropriate

- Very inappropriate
- Extremely inappropriate

3. What do you perceive as the most inappropriate aspect of Jada's behavior? (Select one response only)

- Being tardy to class
- Rolling eyes/shaking head/walking slowly
- Mumbling in a negative tone of voice
- Ignoring you when you asked what he said
- Nothing was inappropriate

4. How likely do you think Jada is to engage in this behavior (or a very similar behavior) in the future? (Select one response only)

- Extremely unlikely
- Very unlikely
- Somewhat unlikely
- Slightly unlikely
- Slightly likely
- Somewhat likely
- Very likely
- Extremely likely

5. What strategy would you most likely use to reduce the likelihood of this situation occurring again the next morning?

Please describe why you would use this strategy:

6. What do you think are the primary causes of Jada's behavior?

7. How often do you see this type of behavior in your classroom?

- Very rarely
- Rarely
- Sometimes
- Often
- Very often

### SECTION 3

DIRECTIONS: Please answer the following four demographic questions about yourself.

What is your gender?

- Female
- Male

How many years have you been teaching?

- 0-1

- 2-5
- 6-10
- 11-15
- 16-20
- 21+

What best describes your race/ethnicity? (Select all that apply)

- American Indian or Alaska Native
- Asian
- Black or African American
- Hispanic, Latino or Spanish origin
- Native Hawaiian or Other Pacific Islander
- White
- Other

How confident are you in your behavior management skills?

- Very confident
- Confident
- Not confident
- Not at all confident

How satisfied are you with your teaching experience?

- Very satisfied
- Satisfied
- Dissatisfied
- Very dissatisfied