

The Relationship Between Skin Tone and School Suspension for African Americans

Lance Hannon
Villanova University
Robert DeFina
Villanova University
Sarah Bruch
University of Iowa

June 2013

*The authors are listed in reverse-alphabetical order. Address correspondence to Lance Hannon, Department of Sociology and Criminology, SAC #204, 800 Lancaster Ave., Villanova University, Villanova, PA 19085.

This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (<http://www.cpc.unc.edu/addhealth>). No direct support was received from grant P01-HD31921 for this analysis.

ABSTRACT

This study contributes to the research literature on colorism--discrimination based on skin tone--by examining whether skin darkness affects the likelihood that African Americans will experience school suspension. Using data from The National Longitudinal Survey of Youth, logistic regression analyses indicated that darker skin tone significantly increased the odds of suspension for African American adolescents. Closer inspection of the data revealed that this overall result was disproportionately driven by the experiences of African American females. The odds of suspension were about 3 times greater for young African American women with the darkest skin tone compared to those with the lightest skin. This finding was robust to the inclusion of controls for parental SES, delinquent behavior, academic performance, and several other variables. Furthermore, this finding was replicated using similar measures in a different sample of African Americans from the National Longitudinal Study of Adolescent Health. The results suggest that discrimination in school discipline goes beyond broad categories of race to include additional distinctions in skin tone.

INTRODUCTION

A burgeoning body of "colorism" research has found that variation in skin tone within racial/ethnic groups in the United States is related to a wide range of economic, political, and social outcomes such as income (Hunter 2002), wages (Goldsmith, Hamilton, and Darity 2006), educational attainment (Loury 2009), self-esteem (Keith 2009), the likelihood of getting married (Hamilton, Goldsmith and Darity 2009), political popularity (Caruso et al. 2009), the probability of receiving the death penalty (Eberhardt et al. 2006), and a host of other consequences (Hersch 2011; Davila, Mora, and Stockly 2011; Vedantam 2010; Nakano Glenn 2009; Hall 2008; Hochschild and Weaver 2007; Maddox and Gray 2002; Hughes and Hertel 1990; Hunter 1998; Krieger 1998; Keith and Herring 1991; Ryabov 2013). Overwhelmingly, these studies have found that higher degrees of "whiteness" in skin tone are associated with better outcomes for minority populations.¹

Perhaps reflecting the growing social science evidence regarding colorism, government agencies have also been taking the issue of color discrimination more seriously. For example, in 2007 the U.S. Equal Employment Opportunity Commission (EEOC) created a new taskforce, E-RACE (Eradicating Racism And Colorism from Employment), to "highlight new and emerging discrimination issues." In the EEOC's (2007) press release announcing the initiative, they noted that while race-based

¹ However, it is worth noting that skin lightness can also be a disadvantage in certain contexts, particularly as it relates to ethnic identity and integration into ethnic groups (Hunter 2007).

discrimination cases continue to make up the bulk of filings, the number of discrimination claims based on skin tone has almost tripled in recent years.²

The present study contributes to the growing body of social scientific findings showing the importance of skin tone by examining whether skin tone affects the likelihood of suspension in school. While a large body of research has documented the existence of racial disparities in the use of school discipline (Lewin 2012; Blake, Butler, Lewis, and Darenbourg 2011; Eitle and Eitle 2004; Taylor and Foster 1986; Raffaele, Mendez, and Knoff 2003; Skiba et al. 2002; Wu et al. 1982), whereby African American youth are much more likely to be suspended at school than other racial groups, it is unknown whether there are additional disparities by skin tone. Indeed, a recent article in the American Educational Research Association's *Educational Researcher* focused on how "insights about color bias are severely wanting" and pointed out that "interrogating colorism as related to behavioral perceptions and disciplinary moments will help clarify connections between race, gender, and discipline" (Monroe 2013:15).

Understanding the various ways in which African Americans are disadvantaged in the use of school disciplinary sanctioning is increasingly important as this type of punishment is used with greater frequency by schools (Way 2011; Kupchik 2010; Hirschfield 2008). In addition, it is important to recognize that disproportionately high rates of school suspension represent more than just a temporary inconvenience for young African American men and women. Existing research in education has firmly established that racial discrepancies in school punishment meaningfully influence life trajectories (Skiba et al. 2011). Suspensions can, for example, significantly affect both the likelihood

² Of course, as Jones (2010) has eloquently argued, an increasing recognition of the legal basis for colorism claims does not mean that such cases have a good chance of winning. As she (2010:661) puts it: "...it is one thing to be able to assert a right to relief. It is another to convince a fact-finder to grant that relief."

of dropping out and going to prison, a chain of events commonly known as the “school to prison pipeline” (Gregory, et al. 2011; Christensen 2012).

Our expectation that colorism may be a problem in school suspension, in addition to the already documented racial disparities in disciplinary outcomes, can be traced in part to the recent findings in work examining punishment outcomes in criminal justice. While it is well-known that African Americans experience harsher sentences and punishments than whites within the criminal justice system (Baldus, et al. 1998; Spohn 2000), a number of recent studies also find intra-racial group disparities in punishment related to physical characteristics such as skin color and facial features (Blair et al. 2004; Pizzi et al. 2005; Eberhardt et al. 2006; Viglione et al. 2011). This work reports that African American individuals with darker skin receive more severe punishments in the criminal justice system (Pizzi et al. 2005; Gyimpah-Brempong and Price 2006; Viglione et al. 2011), and that being perceived as having more Afrocentric features is also associated with longer and harsher prison sentences (Blair et al. 2004; Eberhardt et al. 2006).

In the present paper, we argue that just as conscious or unconscious prejudice based on racial categories can create disciplinary disparities, so too can prejudice based on the degree to which an individual’s skin color matches that of the dominant group. Although not as well studied as other phenomenon, we argue that colorism may constitute a type of durable inequality (Tilly 1998) that affects the life chances of African Americans in a variety of ways. In addition, we argue that in order to understand how skin tone may affect suspension, it is necessary to take an intersectional approach that

distinguishes between the experience of young African American males and females (Choo and Ferree 2010; Collins 2000; Crenshaw 1991).

Our results, from two separate national samples of African Americans, suggest that darker skin tone increases the likelihood that youth will be suspended from school, particularly for African American females. This finding is robust to the addition of other factors thought to influence school punishment, including age, parental SES, delinquent behavior, academic performance, and urban environment. The negative association of skin tone occurs in addition to the known discrepancy in suspensions between African Americans and other racial groups. Thus, a full accounting of how African Americans are disadvantaged by school punishment needs to address both inter- and intra-racial variation.

UNDERSTANDING COLORISM IN THE U.S.

One of the earliest uses of the term “colorism” in American popular culture was by Alice Walker, the Pulitzer Prize winning author of *The Color Purple*, who described it in 1983 as “prejudicial or preferential treatment of same-race people based solely on their color” (1983:290). While the degree of attention this phenomena has received in U.S.-based empirical social science is relatively new, the history of colorism runs as deep as the history of racism in U.S. society.

The idea of colorism suggests that, in addition to a racial hierarchy, there is a skin color hierarchy where it is beneficial for minorities to have skin tones that are closer to that of the dominant group. This skin color hierarchy in the U.S. has its roots in the slavery regime, where white owners gave certain privileges to slaves with more

Eurocentric features, especially those with known white heritage. Citing the work of two prominent scholars of early race relations in the United States, Keith (2009:27) aptly describes this historical context:

Both Fraser and Myrdal argued that the superiority of mixed-race blacks was widely accepted in the slave population as a whole as a result of the status advantages and similarities between whites and mulattos in physical appearance, speech, dress, and behavior. Through this structure of privilege, buttressed by ideological underpinnings of the superiority of white blood, colorism took hold in the African-American community.³

Thus, while legislators tried to formally establish a white-black dichotomy with the idea that a single drop of black blood made one black (the “one-drop rule”), the different lived experiences of lighter-skinned blacks was an important social reality that could filter into the consciousness of both African Americans and whites.

Within the African American community, activities such as marriage, family formation and membership in organizations were influenced by the shade of one’s skin (Keith 2006; Wilder 2009). The notion of light-skinned superiority increasingly became commonplace wisdom, reflected in phrases such as, “If you’re black, get back. If you’re brown, stick around. If you’re light, you’re alright” (Hunter 2002; Inniss 2010). Indeed, both classic survey and recent ethnographic research of African Americans has revealed a variety of terms and traits used to describe variation in skin tone within the African American community that privileges lighter skin (Parrish 1946; Wilder 2010).

It is important to remember, however, that colorism is not simply “black-on-black discrimination.” Colorism is a broad phenomenon where, for example, continuous

³ King and DaCosta (1996) explain that not all those of mixed race enjoyed higher status. Those in the upper south were generally children of white servants and so were not considered superior to blacks. Those in the lower south were often children of wealthy planters and did gain advantage.

variation in skin tone affects the actions of privileged authorities, who tend to be white. Colorism is intrinsically tied to racism in that white privilege is central to both.

While broad distinctions based on racial classification have been shown to be extremely consequential in the U.S., additional divisions by skin tone within racial categories are now being recognized as having consequences as well. As Inniss (2010) succinctly put it, “Given the importance of race--skin color--in the larger society, why would gradations of color *not* be important?”

Indeed, there are even some prominent scholars who suggest that the concept of race would be better measured with a continuous measure of skin color than broad census classifications (Banton 2012). Still, most social scientists researching colorism simply suggest that taking into account variation in skin tone expands the scope in which white racism is seen (Hunter 2007). This means that skin color and designated racial category overlap, but each offers potentially unique insights into the nature of racism, depending on the historical context and current social circumstances. Along these lines, Hochschild (2012), an important contributor to the colorism literature, acknowledged that often times splitting groups up in the name of specificity limits the ability to communicate crucial information that is much more easily seen when people are lumped together (despite meaningful differences among them). Ultimately, Hochschild (2012:4) concluded that for most of U.S. history, traditional broad operationalizations of race have led to appropriate generalizations. However, in the 21st century there is a good case to be made for greater specificity and the inclusion of color scales.

DISPROPORTIONATE SCHOOL SUSPENSION RATES FOR AFRICAN AMERICANS

Explanations for disproportionate school discipline for African Americans have focused on demographic characteristics known to be empirically associated with race (e.g. household income and neighborhood environment) as well as theories of racial stereotyping and theories implying racial differences in delinquent behavior.⁴ Interestingly, one of the most frequently cited potential explanations, racial differences in parental education and income, has been shown to make very little contribution to accounting for the racial disparity (McCarthy & Hoge 1987; Skiba et al., 2002; Wallace et al. 2008).

Research has also examined the extent to which academic achievement mediates the relationship between race and school suspension (Gregory, Skiba and Noguera 2010). Like measures of family SES, while academic performance affects disciplinary outcomes, it does not explain away the association between being African American and the probability of being suspended. Moreover, even if it did account for the racial disparity, it would be difficult to interpret the effect, since racial discrimination within the context of education may produce both lower levels of academic achievement and higher levels of school suspension for African American students (Lewis et al. 2010).

Concerning potential racial differences in student misbehavior, existing research strongly suggests that such differences have little relevance for the disparity in punishment (McCarthy and Hoge 1987). This appears to be the case regardless of

⁴ A complete review of the racial disparity literature is beyond the scope of the current paper (see Losen and Skiba 2010 and Skiba et al. 2002 for excellent summaries). However, we use the literature on *inter-racial* differences in school discipline to inform our selection of core control variables for, what is to our knowledge, the first *intra-racial* analysis of the relationship between skin tone and school suspension.

whether behavior is measured with youth self-reported survey items or other means (Gregory, Skiba and Noguera 2010). For example, looking within reported offense types, Skiba et al. (2011) reported milder school disciplinary experiences for white adolescents. In sum, recent work on this issue suggests that racial differences in, for example, student behavior, academic performance, and socio-economic background explain only a modest portion of the racial disparity, and that discrimination likely plays an important role (Wallace et al. 2008; Skiba et al. 2011).

GENDER, RACE AND PUNISHMENT

Relative to African American males, disproportionate school suspension rates for African American females have received less attention in both academic studies and the popular press. The tendency to focus on the disciplinary experiences of African American males may simply reflect the fact that their suspension rate is so much higher than that of African American females, as well as a societal preoccupation with the imagined crimes of African American men (Russell-Brown 1998; Wald and Losen 2003, 2007). Regardless, similar to their African American male counterparts, recent data from the Department of Education point to significant racial disparities in the application of school punishment for African American females (Lewin 2012). For example, they are about three times more likely to be suspended than white females (Blake, Butler, Lewis, and Darenbourg 2011; Taylor and Foster 1986; Raffaele, Mendez, and Knoff 2003).

Relative to white adolescents, African American youth are more frequently disciplined for offenses that involve a high degree of subjectivity. This seems to be especially the case for African American females (Skiba et al. 2002; Costenbader and

Markson 1998; Vavarus and Cole 2002). In one of the few existing studies focused on the school disciplinary experiences of African American females, Blake et al. (2011) reported that cases of “disobedience,” “defiance” and “improper dress” made up the majority of disciplinary infractions. Blake et al. noted that their results were consistent with earlier qualitative research on how teachers reacted to African American females who were assertive (Evans 1988; Morris 2007; Ladner 1971). For example, Morris (2007) found that teachers were much more likely to cite African American female students for “unladylike” behavior, particularly being loud. Ultimately, Blake et al. (2011) concluded that similar to African American males, African American females face stereotypes that can subconsciously influence a teacher’s assessment of appropriate disciplinary action. However, for females, “these stereotypes surround Black girls’ femininity rather than their potential for violence” (2011:93).

Viglione et al. (2011) make a similar point regarding the intersection of race and gender stereotypes in their study of African American female inmates. They (2011:252) suggest that African American men may receive longer prison sentences than their white counterparts because they are “stereotyped as dangerous,” while African American women may receive longer sentences because their dark skin is seen as less feminine under the dominant white ideal of beauty. However, Viglione et al.’s analysis went further than just suggesting an advantage for white women over African American women in the ability to draw sympathy from judges and juries. Their data for over 12,000 African American women incarcerated in North Carolina showed that, even within the African American population, variation in skin darkness mattered for sentence length. That is, lighter-skinned African American women received lesser punishments

and served less time behind bars than darker-skinned African American women who were convicted of similar offenses.

METHODOLOGY

Consistent with the increased social science and legal attention to this issue, the U.S. Bureau of Labor Statistics announced in 2010 that its widely used National Longitudinal Survey of Youth (NLSY97) now includes a measure of interviewer-assessed skin tone. We make use of this important national dataset in our analyses of the relationship between skin tone and the likelihood of experiencing school suspension for African American adolescents.

The empirical analysis of school suspension uses a sample of young African American women and men from the NLSY97. In total, 1,797 African American youth were interviewed in-person and had valid records for the present study's key variable, skin darkness. Missing data for the other variables reduced the sample to 1,175 (all observations with missing data are casewise deleted). Although the NLSY97 contains the most complete data for our purposes, we also employ the National Longitudinal Study of Adolescent Health (Add Health) to examine the robustness of the findings. The Add Health provides a larger sample of 2,621 African American youth.

The basic methodology models the binary outcome of whether an African American adolescent was ever suspended using a logistic regression. The existence of a suspension is coded as "1" and the absence as "0". The logistic regression indicates how a unit change in each explanatory variable affects the log of the odds of a suspension

occurring.⁵ The models include controls for several other factors that have been shown to affect the likelihood of suspension. The model's explanatory variables include an adolescent's skin darkness, academic achievement as gauged by test scores, the frequency with which he or she has engaged in a variety of delinquent activities, the family's socio-economic status, the adolescent's age and the urbanicity of the adolescent's residence.

Additional robustness checks are also performed incorporating alternative measures of race, family composition, contextual variables that account for possible interviewer and community influences and additional individual-level influences. A more detailed description of the primary data set and the model variables follows.

The 1997 National Longitudinal Survey of Youth data and variables. The NLSY97 is a sample of about 9,000 youths who were 12 to 16 years old as of December 31, 1996. Stratified, multistage probability sampling was used to generate two independent samples: the main, national sample and an oversample of African Americans and Hispanics. The NLSY97 cohort was selected in two phases. In the first phase, a list of housing units for the cross-sectional sample and the oversample was derived from two independently selected, stratified multistage area probability samples. This ensured an accurate representation of different sections of the population defined by race, income,

⁵ The odds of an event occurring is simply the ratio of the probability of success to that of failure. So, if the probability of a suspension is .8 and, hence, that of no suspension is .2, then the odds of a suspension is $.8/.2$, or 4 to 1. Results from logistic regressions can also be expressed as an odds *ratio*. In that case, one calculates the odds under different values of an explanatory variable and takes the ratio of the different odds. In this study, skin darkness is of primary interest. Thus, one can calculate the odds of suspension given that skin tone is light and the odds of suspension given that skin tone is dark. The odds ratio of dark to light indicates how much more likely an adolescent is to be suspended if she/he has dark skin compared to light skin. For example, if the estimated dark to light odds ratio coefficient is 2, one would conclude that the odds of suspension increase by a factor of two when we compare dark skinned adolescents to light skinned ones. The skin tone measure in the present study has several darkness categories. Continuing to assume an estimated coefficient of 2, one would calculate the change in the odds ratio of increasing the darkness of an adolescent's skin by n categories as 2^n .

region, and other factors. In the second phase, subsamples of the eligible persons identified in the first phase were selected for interview. Round 1 of the survey occurred in 1997, and the youths continue to be interviewed on an annual basis.

The NLSY97 data provide three assessments of an adolescent's race. These include a self-reported race, a parent-reported race, and an interviewer-reported race. The baseline estimates use self-reported race, although the other two classifications are employed in robustness tests.

The NLSY97 suspension measure is based on a question asked of youth in 1997, "Have you ever been suspended from school?" The response is coded "1" for a positive response and "0" for a negative response.

Skin darkness measurement in the NLSY97 uses a procedure designed by Douglas Massey and Jennifer Martin, and is based on a scale developed originally by Massey, Charles, Lundy, and Fischer (2003).⁶ The measure is an 11-point scale, ranging from zero to 10, with zero representing the total absence of color, and 10 representing the darkest skin (no African Americans in the NLSY97 data had a skin color value of zero). The ten shades of skin color corresponding to the points 1 to 10 on the Massey and Martin Skin Color Scale are depicted in a chart, with each point represented by a hand differing only in color. The scale was constructed with assistance from a professional graphic designer to help capture slight but discernible differences in skin tone. An adolescent's skin darkness is assessed by an interviewer who is trained in the scale's specific use and who memorizes it so that in the interview she or he can code the skin color of the respondent without the instrument.

⁶ The following description comes from the NLSY97 documentation.

Working with the Massey and Martin Scale in another dataset, Hersch (2008:356) described it as a “unique...more continuous and specific measure of skin color relative to all other surveys.” Arguing for the validity of the scale, Hersch further noted that several demographic variables exhibited similar correlations with the Massey and Martin Scale as they did with a measure of skin tone that was objectively assessed using a reflectance spectrophotometer.

The NLSY97 data also contains a pre-constructed scale that measures self-reported delinquency. The pre-constructed variable is derived from responses about whether an adolescent has committed each of ten different types of delinquent acts (yes is coded as “1”, no “0”).⁷ The NLSY97 delinquency index score is created by summing the responses. Consequently, the resulting range for the variable is 0 to 10.

The socio-economic status (SES) measure is constructed using parents’ average level of education in years and household income. Both variables are standardized (0, 1) and added to form the SES measure.⁸ The test score variable is pre-constructed in the NLSY97, and is an index based on an adolescent’s percentile performance on math knowledge, arithmetic reasoning, word knowledge and paragraph comprehension. The

⁷ Respondents are asked the following ten questions: 1. Have you ever run away, that is, left home and stayed away at least overnight without your parent's prior knowledge or permission; 2. Have you ever carried a hand gun? When we say hand gun, we mean any fire arm other than a rifle or shotgun.; 3. Have you ever belonged to a gang?; 4. Have you ever purposely damaged or destroyed property that did not belong to you?; 5. Have you ever stolen something from a store or something that did not belong to you worth less than 50 dollars?; 6. Have you ever stolen something from a store, person or house, or something that did not belong to you worth 50 dollars or more including stealing a car?; 7. Have you ever committed other property crimes such as fencing, receiving, possessing or selling stolen property, or cheated someone by selling them something that was worthless or worth much less than what you said it was?; 8. Have you ever attacked someone with the idea of seriously hurting them or have a situation end up in a serious fight or assault of some kind?; 9. Have you ever sold or helped sell marijuana (pot, grass), hashish (hash) or other hard drugs such as heroin, cocaine or LSD?; and, 10. Have you ever been arrested by the police or taken into custody for an illegal or delinquent offense (do not include arrests for minor traffic violations)? A positive response to a question is coded as “1” and a negative “0”.

⁸ If either father’s or mother’s education is missing, the missing parent is assigned the non-missing parent’s education level. If one or two of the elements of SES (parental education and household income) are missing, the calculation of SES is based on the available data.

values range from 0 to 100, with higher values indicating better performance. NLSY97 designates an adolescent as an urban resident if she/he lives in a central city, as defined by the Census Bureau. The urbanicity variable is coded “1” for urban and “0” otherwise. The age of an adolescent is measured in years.

Table 1 contains the summary statistics for the variables. Statistics are given for the total sample of African American youth and for female and male sub-samples. The data indicate that 45 percent of African American youth overall have been suspended. The experience varies considerably by sex, however, with 53 percent of males having been suspended compared to 38 percent for females. This is consistent with previous research findings. For example, using a nationally representative sample of 10th graders, Wallace et al. (2008) found that almost 50% of African American students reported being punitively sanctioned (suspended or expelled). Also consistent with past research, African American females have somewhat higher test scores. The average delinquency score for the total sample is 1.21, meaning that youths in the sample were on average involved in 1.21 of the possible 10 delinquent acts that comprise the score. Once again, there is clear variation by sex, with the average male value at 1.59 and that of females less than 1. The average ages for males and females are virtually identical (15 years).

The explanatory variable of primary interest is the skin darkness score. The total sample average is 6.2, which is darker than the midpoint of the range and unsurprising given the sample is restricted to African Americans. Sample males have somewhat darker reported skin tone than females. This finding is consistent with other results where a scientific instrument (a light reflectance meter) was used to measure skin tone differences between African American men and women (Borrell et al. 2006). Thus, it

appears that the modest difference in skin darkness between male and female African Americans is not simply a product of interviewer bias, considering that more objective methods for assessing skin tone produce similar results. Figure 1 provides the percent distribution for the male and female samples across the full range of possible skin tone values. As can be seen by comparing Panels A and B, male values are skewed a bit more to the right than females, although the general profiles of the distributions are similar. The mass of the distributions is located between values of 4 and 9.

Table 2 presents bivariate correlations. The correlations among the independent variables are modest and so multi-collinearity is not a problem. The bivariate correlations between the explanatory variables and occurrence of suspension have the anticipated signs. Skin darkness and the occurrence of suspension are positively correlated, and are statistically significant at the 5 percent level for the overall and females samples. It is positive but not significant for the male sample. Suspension and the delinquency scale and age are positively correlated as well, while suspension and test scores and SES are negatively correlated. We now turn to whether the positive relationship between suspension and skin darkness remains in a series of logistic regression analyses.

EMPIRICAL RESULTS

Baseline logistic regression estimates. We begin with bi-variate logistic regression estimates to establish simple correlations between the probability of suspension and skin darkness. We then turn to the more important estimates from the multi-variate models, which will serve as our baseline through the rest of the paper. To facilitate comparison,

we estimate the bivariate models using only those observations that are also available for the corresponding multivariate estimations. However, we reach the same conclusions for the bivariate estimates if the largest possible samples are used.⁹

Three sets of bi-variate associations were estimated: one for the overall African American sample; one for African American male adolescents; and, one for African American female adolescents. The estimated odds ratio for the overall sample is greater than one (1.127) and statistically significant ($p=.000$), indicating that darker skin is associated with a greater chance of suspension. Furthermore, the estimated coefficient is significant for both the male and female sub-samples. The estimated odds ratio for females is greater than one (1.144) and statistically significant ($p=.001$) while that for the male sample is 1.107 ($p=.028$). Recalling that the Massey and Martin skin tone measure ranges over ten possible values (1 to 10), the impact on the odds ratio of moving from the lightest to the darkest skin tone (a change of nine categories) is 1.144^9 , or a factor of 3.4 for the female sample and 1.107^9 , or a factor of 2.5, for the male sample. That is, it is about 3.4 times as likely for a young African American woman with the darkest skin tone to be suspended compared to one with the lightest skin, and 2.5 times as likely for a young African American male.

The multi-variate model estimates are presented in Table 3 and include five additional factors predicting suspension. In the model using the total sample, we add a control for gender, given that the suspensions are more common for males. The addition of these other measures to the regression leaves the estimated odds ratio in the skin tone measure basically unchanged relative to the bivariate models. The odds ratios for each

⁹ Models were estimated both with and without the general sampling weights. The basic conclusions regarding the impact of skin tone on the likelihood of suspension are unaffected. The estimates discussed in the text do not use the weights.

sample remains greater than one and significant. For the female sample, the odds of a suspension rise by a factor of 3.6 (1.152⁹) moving from the lightest to the darkest skin tone, and the odds rise by a factor of 2.5 for the male sample.

Figure 2 displays the predicted probabilities of suspension for the male and female sub-samples, based on the estimates from the multi-variate models. The probabilities are calculated setting all model variables other than skin darkness at mean values (Table 1). As can be seen, the predicted probabilities for the male sample lie everywhere above those for the female sample, consistent with the higher suspension rates experienced by male African American students. Indeed, the probability of suspension for males exceeds 50 percent at the mean values of skin darkness and other variables. Note, however, that the probability of suspension grows faster for African American female students as skin darkens (consistent with the odds ratio estimates). The probability that a male African American student will be suspended rises from .48 when skin is lightest to about .69 when skin is darkest, a 45 percent rise. For females, the probability rises from .28 to .58, a 107 percent increase. Thus, there appears to be a gender difference in the impact of skin darkness, although both effects are significant.

Concerning the other factors that also contribute to the likelihood of being suspended, a record of delinquent behavior significantly increases the odds ($p < .01$), while higher tests scores decrease the odds ($p < .01$). Increases in an adolescent's SES lower the odds of suspension in the total and female samples, and urbanicity had a significant or marginally significant positive impact in the total and males samples. These findings correspond with those in the literature on race and school suspension. Age was found to have an insignificant effect, and not surprisingly, the gender indicator variable in the total

sample equation is large and significant estimating that males are about 1.4 times as likely as females to experience a suspension.

Robustness checks with the NLSY97. The baseline estimates indicate that having darker skin increases an African American adolescent's odds of being suspended, other things equal. In this section, we address several additional concerns about other factors that may influence the primary findings.

One issue that arises concerns the race of the interviewer. Some research has found, for example, that the race of the interviewer influences his or her perception of skin darkness. Hill (2002) concluded that African American interviewers classified white respondents' skin tones as significantly lighter than white interviewers did, and that white interviewers classified African American respondents' skin tones significantly darker than African American interviewers did. Interviewers also reported greater skin darkness variation among same-race respondents than other-race respondents. Depending on the distribution of interviewer race in relation to respondent race, some bias could conceivably be introduced. For instance, if African American respondents were matched mainly with white interviewers, observed skin darkness variation could be limited and the estimated impact of skin darkness on suspension attenuated. The idea here is that the potentially greater variation in measured skin darkness for African Americans associated with African American interviewers could strengthen the estimated relationship between suspension and skin darkness.

Although the evidence concerning possible bias due to interviewer race relies on different data sets using less rigorous skin tone measures, we explore its potential relevance for our results. NLSY97 classifies the interviewer's race in one of six

categories – white; black or African American; American Indian, Eskimo or Aleut; Asian or Pacific Islander; other race; and, Multiracial. We coded an interviewer’s race as “1” if black or African American and “0” otherwise. We then created an interaction variable, equal to the interviewer race indicator times the skin darkness measure, and included both the interaction and the levels of skin darkness and interviewer race in the baseline models.¹⁰ Neither the interaction variable nor the interviewer race variable was close to significant in the three regressions. We conclude that the effect of skin tone on suspension is independent of any concerns about interviewer race in this particular dataset.¹¹

A second issue concerns the procedure by which a youth’s race is determined. As mentioned, the NLSY97 data contain three alternative classification schemes: 1) the youth’s self-reported race (used in our baseline estimates); 2) the parent’s reported race for the youth; and, 3) the interviewer’s assessment of the youth’s race. These three modes of classification need not produce the same racial groupings. As Saperstein (2012) explains, “The experience of race in the United States is shaped by both self-identification and ascription” (p. 1484). It is possible that individuals’ life experiences and their interactions with others could lead to a different, and perhaps more nuanced, understanding of their own race than either the opinions of the parent or the interviewer. For example, knowledge of ancestry and patterns of socializing could lead a light-skinned

¹⁰ Following standard procedures, we centered both the skin darkness and interviewer race variables by subtracting the mean from each before interacting them, and used the centered levels of the variables in the regression.

¹¹ We tried two additional specifications in which we included only the level of the interviewer’s race indicator without the interaction. In the first, we used only the African American/non-African American interviewer race indicator in the baseline model. We found that the skin darkness odds ratios were significant and greater than one. The odds ratio for the indicator variable was insignificant. We also created indicator variables (0/1) for each of the six interviewer race categories and then estimated the baseline model including all the interviewer race indicators apart from white, which served as the excluded category. None of the indicator variables was significant.

person to consider themselves as black, while an interviewer, judging only on the basis of facial characteristics, could classify the person as white. There could also be “racial fluidity” whereby the overall assessment and specific relevance of racial categories change depending on the social, economic or political context (Saperstein and Penner 2012; Harris et al 2011).

To explore whether the baseline results are sensitive to our use of self-identified race, we re-estimated the models using samples of African American children based on both parent- and interviewer-assessed race. For the parent-identified category, the estimated skin darkness coefficients for the total and female samples are 1.099 ($p=.004$) and 1.147 ($p=.003$), respectively. The coefficient for males is 1.077 but marginally insignificant. ($p=.123$) For the interviewer-assessed category, the coefficients for the total and female samples are 1.111 ($p=.002$) and 1.161 ($p=.002$) respectively. The coefficient for the male sample is 1.090 and marginally significant ($p=.074$). Thus, the results for the total and female samples are unaffected by the choice of racial classification method, while the results for the male sample are reduced in magnitude and statistical significance.

A third concern is whether the respondent lived in a bi- or multi-racial family. That is, the question arises as to whether any measured advantage from lighter skin actually arises because one or more family member is non-black, which might confer higher status to the youth in the eyes of teachers and administrators. The higher status could explain the reduced likelihood of suspension instead of the adolescent’s skin darkness *per se*. Another way this could matter is that respondents with bi- or multi-racial

families may be assessed as racially ambiguous which has been shown to be a unique experience (Campbell 2009).

To explore this possibility, we excluded from our samples any adolescent who has a non-black family member as reported by their parent (such family members include grandparents who live with the youth as well as parents that do not). The results are very similar to the baseline. The estimated coefficient for skin darkness is 1.109 ($p=.003$) for the total sample, 1.152 ($p=.004$) for the female sub-sample and 1.096 ($p=.076$) for the male sub-sample. These results show that the advantages associated with skin lightness are not due to having a non-black family member.

Robustness check using the National Longitudinal Study of Adolescent Health

To complement the results based on the NLSY97 sample, we performed supplemental analyses using another publicly available national dataset that contains a measure of school suspension and an interviewer-assessed indicator of skin darkness: the National Longitudinal Study of Adolescent Health (Add Health). The use of the Add Health data allows us to assess the generalizability of the results using a different sample of African American adolescents, as well as explore additional relevant individual controls and school-level contextual variables that have been found to be important factors to understand the race-suspension link that are not available in the NLSY97.

Although some of the baseline model variables are measured in slightly different ways in the Add Health data set, they are sufficiently comparable as to provide useful robustness checks. Unlike the NLSY97, in the Add Health the school suspension variable refers specifically to only out-of-school suspensions. Also, different from the NLSY97, skin tone is assessed with just five categorical distinctions (white, light brown, medium

brown, dark brown, and black), and the interviewers were given no specific training on making skin color assessments.^{12,13}

Table 4 presents the bivariate and multivariate logistic regression model results using the Add Health sample. The key result presented in this table is that the finding of an association between skin color and suspension is replicated in the Add Health data.

To begin, the bivariate relationship between skin darkness and suspension is similar between the two samples. In the Add Health sample, the association is positive and statistically significant at the .05 level for the self-reported African American female sample (1.213; n=1481), but not statistically significant for self-reported African American males (1.099; n=1140) (Table 4, columns 1 and 4). The female sample bivariate estimate of the relationship between skin tone and suspension in the Add Health data suggests that a young African American woman with the darkest skin is about 2.2 times (1.213⁴) as likely to be suspended as one with the lightest skin (an increase of four skin color categories). This bivariate estimate is a bit less than that based on the NLSY97 (around 3), but still sizable.

We next incorporate control variables that are analogous to those used in the baseline estimates for the NLSY97: SES, delinquency, test scores, urbanicity and age

¹² We estimate the new models with specifications indicating the complex survey structure of the Add Health data following the recommendations found in the Add Health data documentation and perform design-based analyses which specify the clustered nature of the data and the stratification used in the sampling (Chantala and Tabor 1999). Doing so accounts for the potential non-independence of the observations and produces correct standard errors.

¹³ Specifically, the measures of delinquency, urbanicity and test score are constructed somewhat differently. The delinquency variable is constructed using the answers to questions about whether a child committed any of nine different offenses. If a child committed a covered offense, a value of “1” was assigned for that type of offense; if not, a value of “0” was assigned. The delinquency variable is equal to the sum of the different types of offenses committed by the child. Thus, the variable’s values range from 0 to 9. The urbanicity variable takes a value of “1” if the child lives in a “completely urban area” as defined by the Census Bureau, and a “0” if not. The test score variable is the score a child received on the Peabody vocabulary picture test.

(Table 4, columns 2 and 5). Most importantly, the estimated odds ratio for skin darkness in the female sample continued to be greater than one and significant (1.190; $p=.022$), consistent with the NLSY97 findings. The estimated odds ratio for skin darkness in the male sample remained statistically insignificant after the inclusion of the control variables (1.079; $p=.512$).

A third set of estimates expands the model to include three school-level variables that have been found to significantly affect school suspension rates (e.g., Gregory, Cornell and Fan 2011), along with two additional individual-level controls. The school-level measures capture the racial composition, level of racial inequality, and overall school climate. The racial composition is measured as the percent of the student body that is African American. The racial inequality measure is the ratio of the average parent education of African American to white students in the school. And the school climate measure combines eight student-reported items measuring the degree to which students report: having trouble getting along with teachers; thinking teachers treat kids fairly; having trouble getting along with other students; feeling close to people at school; feeling like part of school; thinking students at this school are prejudiced; feeling happy to be at this school; and feeling safe at school.

In addition to accounting for differences across school contexts, the Add Health data also allow us to examine two other factors that may affect the association between skin color and suspension. First, we include a measure of interviewer-assessed physical attractiveness (a simple Likert scale). This measure is included because Hill (2002) notes that skin tone can influence perceptions of attractiveness, which in turn can change the likelihood that a given behavior will be worthy of suspension. Second, we include a

measure of self-esteem because previous work has shown that skin tone can affect one's self-esteem, given an aesthetic norm that prefers skin lightness to darkness (Thompson and Keith 2001). In this instance, darker skin can reduce self-esteem and lead to actions that leave one vulnerable to suspension.

The results are displayed in columns 3 and 6 of Table 4. Among the new variables, the school climate measure demonstrated the clearest relationship with suspension. For both the female and male samples the association was negative and significant, indicating that a better school climate is associated with fewer suspensions. The interviewer's perception of respondent attractiveness was not significantly associated with suspension in either sample. Interestingly, self-esteem was negatively associated with suspension for African American females ($p=.070$), but positively associated for African American males ($p=.074$). This result underscores the importance of recognizing potentially different racial scripts for African American males and females, and thus the need for gender-specific analyses. Most important for the current study, as with the other Add Health results, the estimated coefficient for skin darkness remains significant for the female sample (1.184; $p=.044$), but continues to be insignificant for the male sample (.957; $p=.746$).

In sum, estimates from comparable models using an entirely different sample lead to the same conclusion. That is, skin tone is found to have significant impacts on the likelihood of school suspension for young African American females, both in a statistical and practical sense. The inclusion of additional individual-level and school-level controls does not diminish this result. The lack of significance found for males in the Add Health

data corresponds to the less consistent evidence of a significant effect for males in the NLSY97.

EXPLAINING THE GENDER-SPECIFIC IMPACT OF SKIN TONE

Although further research is needed to corroborate the evidence presented here, the findings of the present analysis suggest that skin shade matters for the likelihood of suspension, with the impact being more discernible for African American females than African American males. There are several possible explanations for the gender-specific nature of the results that rely on important intersections of race, skin tone and gender (Choo and Ferree 2010; Collins 2001; Baca Zinn and Thornton Dill 1996). That is, rather than interpreting skin darkness as an isolated disadvantage with a unique and unchanging effect, it is more useful to view it as part of a matrix of domination in which it converges with other modes of oppression in either mutually reinforcing or limiting ways. Just as feminism must be comprehended in a multi-racial context (Baca Zinn and Thornton Dill 1996), so must the effects of skin darkness be thought of as contingent upon race and gender. Furthermore, historical, social, political and spatial contexts can alter the forms and extents of oppression and, hence, their interactions.

There is evidence, for example, that black women tend to describe those with lighter skin using generally positive terms; alternatively, they apply negative descriptors to those with dark skin (Wilder 2011). The evidence also suggests that black women incorporate the messages carried by the names, both positive and negative, forming a “color habitus.” These internalized scripts can guide attitudes and behaviors, and shape the meanings that young black women attach to the words and actions of others.

Consequently, particular responses to the same stimuli by dark skinned girls might be more aggressive and assertive than those of lighter skinned girls, given their personal histories of inferior treatment and denigration. Those responses, coupled with dominant beliefs about femininity, can lead to disproportionate punishment.

Indeed, Joyce Ladner's classic study of poor black girls in urban St. Louis describes how daily encounters with colorism outside of the school system can contribute to survival strategies that, while functional in the community, can be self-defeating in other settings like the workplace or school (Ladner 1971). If, as some research suggests (Blake et al. 2011), African American females are more likely to be sanctioned for offenses involving a high degree of subjective discretion (such as talking loudly), then the potential for more nuanced types of unconscious discrimination will be greater for darker skinned girls who may more frequently exhibit these adaptive behaviors. Thus, race, gender and skin color combine with dominant norms in ways that reinforce and reproduce the inequality and oppression faced by females with darker skin.

Intersections of gender and skin tone also become relevant with regard to images of feminine beauty. For example, Keith (2009:26) has argued that, "The gendered nature of colorism stems from the close link between skin tone and perceptions of physical attractiveness, and from a double standard that applies expectations of attractiveness more rigidly to women." Hunter (2002) identified what she called a skin-tone "beauty queue," whereby lighter skinned women are considered more beautiful than darker skinned women. The importance of such perceptions here is that attractiveness brings with it advantages in a variety of realms, such as in the job and marriage market, and in education (Hunter 2005). Concerning suspension, a presumption of superiority can lead

to “second chances” or the “benefit of the doubt” not available to darker skinned girls. The end result could be disproportionate school punishment for those with darker skin.

The weaker and less consistent correspondence between skin tone and suspension for male African American students might also reflect significant interactions between race, gender and skin tone. Unlike black women, the controlling image of black men in the United States is of a dangerous, criminal predator. The image has been socially constructed in a variety of ways, perhaps none so important as the War on Drugs and a policy of mass incarceration that has targeted black communities and especially black men.

It is possible that the effect of skin darkness on punishment for men may operate through the “Criminal Black Man” stereotype (Russell-Brown 1998), in that darker skin color strengthens the association of young African American men with that stereotype. As such, the disadvantage for darker-skinned young men derives from their link to a disfavored group (criminals) in an analogous way that the disadvantage for young women does (unfeminine and unattractive). The less pronounced effect of skin tone for men could reflect the pervasiveness of the “Criminal Black Man” stereotype. That is, the stereotype may be so strong that the effect of categorical racial membership for African American males makes any finer distinctions based on skin tone of secondary importance. Thus, the different meanings of race for black men and women and their intersections with skin darkness could produce alternative outcomes with regard to suspensions.¹⁴

¹⁴ The null findings for men in the Add Health sample might also arise due to the crudeness of the Add Health skin darkness measure. The noisiness of the measure could result in the well-known errors-in-variables problem, such that the skin darkness coefficient is biased toward zero. Indeed, the size of the estimated effects based on the Add Health data are smaller for both males and females compared to those

CONCLUSION

Using two national samples, multivariate logistic regression analyses revealed that skin tone matters for the likelihood that African Americans will experience school suspension. In particular, the odds of a suspension were about three times greater if a young African American woman has the darkest skin tone compared to the lightest. The importance of skin darkness for female African American adolescents appeared in both datasets (the NLSY97 and the Add Health) and was robust to the inclusion of a variety of other factors that could explain the link between race and suspension, including a combined index of parental education and income, a measure of student academic performance, and an index of delinquent behavior. This result was also independent of interviewer race, urban environment, as well as school-level characteristics, perceived physical attractiveness and reported self-esteem. Although less consistent, there was also evidence that skin darkness affects the likelihood of suspension for African American males.

The primary objective of the current analysis was to take a first step toward documenting any *intra*-racial disparity in school suspension by skin tone for young African Americans. Following the path taken in the large research literature documenting an *inter*-racial disparity, we have sought to first establish the basic relationship between skin tone and suspension. Our hope is that, given the basic finding that there is disproportionate punishment for darker skinned versus lighter skinned African Americans, future qualitative and quantitative research efforts will turn to

based on the NLSY97 data. The significant findings for the female sample could simply indicate that the effect is stronger for women and can be seen despite shortcomings in the skin darkness measure.

empirically identifying what is and what is not a likely cause of the disproportionality. Based on the now sizable empirical evidence on the unequal use of discipline between racial groups, we agree with Skiba et al. (2011:104) that,

The fact of racial/ethnic disproportionality in school discipline has been widely and, we would argue, conclusively demonstrated... These differences do not appear to be explainable solely by the economic status of those students, nor through a higher rate of disruption for students of color... In the absence of an evidence-based rationale that could explain widespread disparities in disciplinary treatment, it must be concluded that the ubiquitous differential removal from the opportunity to learn for African American and Latino students represents a violation of the civil rights protections that have developed in this country since *Brown v. Board of Education*.

Additionally, we believe that our intra-racial skin color findings could reasonably be interpreted as offering even further evidence that the observed racial disparity is not due to differences in behavior or other factors, especially if one believes that externally-assessed skin tone and self-identified racial grouping are both distinctly useful ways of operationalizing the multidimensional concept of race (Telles 2012).

The U.S. has become increasingly diverse, especially through Hispanic immigration, and the number of people claiming a mixed-racial identity has more than doubled since the last Census. As noted earlier, the Equal Employment Opportunity Commission (EEOC) has also seen a substantial increase in the number of discrimination claims based on skin tone. In general, there appears to be growing recognition of the prevalence and harmfulness of colorism in U.S. society. Still, we argue that more attention is warranted and more research is needed, particularly on the relationship between skin tone and the unequal distribution of punishment.

If the findings of the current analysis are corroborated with other results, then relevant agencies should seriously consider expanding the monitoring of racial disparities

in school discipline to include skin tone differences. Our estimates suggest that an exclusive focus on race masks important skin tone effects and that ultimately a full accounting of the relationship between white privilege and punishment needs to address both inter- and intra-racial differences. While such an expansion would complicate the popular discourse on racial inequality, policy solutions to discrimination that are based on an oversimplification of racial dynamics may ultimately do more harm than good, especially if low levels of policy effectiveness are interpreted as proof that the problem is somehow intractable. Moreover, without attention to colorism, significant within-race variation in punishment might be misinterpreted as evidence that the problem lies with the individual and that prejudice is irrelevant.

REFERENCES

- Baca-Zinn, M. & Thornton-Hill, B. (1996). Theorizing Difference from Multiracial Feminism. *Feminist Studies*, 22, 2.
- Baldus, D.C., Woodworth, G., Zuckerman, D., Weiner, N.A., and Broffitt, B. (1998). Racial Discrimination and the death penalty in the post-Furman era: An empirical and legal overview, with recent findings from Philadelphia. *Cornell Law Review* 83: 1638-1770.
- Banton, M. (2012). The colour line and the colour scale in the twentieth century. *Ethnic and Racial Studies*, doi: 10.1080/01419870.2011.605902
- Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science*, 15(10), 674-679.
- Blake, J. J., Butler, B. A., Lewis, C. L., & Darenbourg, A. (2011). Unmasking the inequitable discipline experiences of urban Black girls: Implications for urban stakeholders. *Urban Review*, 43, 90-106.
- Borrell, L., Kiefe, C., Williams, D., Diez-Roux, A., & Gordon-Larsen, P. (2006). Self-reported health, perceived discrimination, and skin color in African Americans in the CARDIA study. *Social Science and Medicine*, 63(6), 1415–1427.
- Campbell, M. (2009). “Multiracial Groups and Educational Inequality: A Rainbow or a Divide?” *Social Problems* 56(3): 425-446.
- Caruso, E. M., Mead, N. L., & Balcetis, E. (2009). Political partisanship influences perception of biracial candidates’ skin tone. *Proceedings of the National Academy of Sciences*, 106, 20168-20173.

- Chantala, K. & Tabor, J. (1999). Strategies to perform a design-based analysis using the Add Health data.
<https://www.cpc.unc.edu/projects/addhealth/data/guides/weight1.pdf>
(accessed 2/13/2013).
- Christensen, L. (2012). "The Classroom-to-prison pipeline." *Rethinking Schools*.
26(2):24-27.
- Choo, H.Y. & Ferree, M.M. (2010). "Practicing Intersectionality in Sociological Research: A Critical Analysis of Inclusions, Interactions, and Institutions in the Study of Inequalities." *Sociological Theory* 28(2): 129-149.
- Collins, P.H. (2000). *Black Feminist Thought*. New York: Routledge.
- Costenbader, V., & Markson, S. (1998). School suspension: A study with secondary school students. *Journal of School Psychology*, 36, 59-82. Hattie, J.A.C. (2002). Classroom composition and peer effects. *International Journal of Educational Research*, 37, 449-482.
- Crenshaw, K. (1991). "Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color." *Stanford Law Review* 43:1241-99.
- Davila, A., Mora, M.T. & Stockly, S.K. (2011). Does Mestizaje Matter in the US? Economic Stratification of Mexican Immigrants. *American Economic Review* 101(3): 593-597.
- Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy: Perceived stereotypicality of black defendants predicts capital-sentencing outcomes. *Psychological Science*, 17(5), 383-386.

- EEOC (2007). EEOC takes new approach to fighting racism and colorism in the 21st century workplace. Accessed April 8, 2012 Press Release:
<http://www.eeoc.gov/initiatives/e-race/index.html>
- Eitle, T. M. N., & Eitle, D. J. (2004). Inequality, segregation, and the overrepresentation of African Americans in school suspensions. *Sociological Perspectives*, 47, 269–287.
- Evans, G. (1988). Those loud black girls. In D. Spender & E. Sarah (Eds.), *Learning to lose: Sexism and education* (2nd ed., pp. 183–190). London: Women’s Press.
- Fenning, P., & Rose, J. (2007). Overrepresentation of African American students in exclusionary discipline: The role of school policy. *Urban Education*, 42, 536–559.
- Goldsmith, A., Hamilton, D. & Darity, Jr. W. (2006). Shades of Discrimination: Skin Tone and Wages. *American Economic Review* 96(2): 242-245.
- Gregory, A., Skiba, R. J., & Noquera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Education Researcher*, 39(1), 59-68.
- Gregory, A., & Thompson, A. R. (2010). African American high school students and variability in behavior across classrooms. *Journal of Community Psychology*, 38(3), 386-402.
- , Cornell, D & Fan X. (2011). “The relationship of school structure and support to suspension rates for black and white high school students. *American Educational Research Journal*, 20(10), 1-31.

- Gregory, J. F. (1995). 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.
- Gyimah-Brempong, K., & Price, G.N. (2006). Crime and punishment: And skin hue too? *American Economic Association* 96(2): 246-250.
- Hall, Ronald E. (2008). *Racism in the 21st Century: An Empirical Analysis of Skin Color*. Springer, New York.
- Hamilton, D., Goldsmith, A. & Darity, W. (2009). Shedding 'light' on marriage: The influence of skin shade on marriage for Black females. *Journal of Economic Behavior and Organization*, 72(1):30-50
- Harris, A, Evans H. & Beckett K. (2011). Courtesy stigma and monetary sanctions: Toward a socio-cultural theory of punishment. *American Sociological Review* 76(2):234-264.
- Hersch, J. (2011). The persistence of skin color discrimination for immigrants. *Social Science Research* 40 (2011) 1337–1349.
- Hersch, J. (2008). Profiling the new immigrant worker: the effects of skin color and height. *Journal of Labor Economics* 26 (2), 345–386.
- Hersch, J. (2006). Skin tone effects among African Americans: perceptions and reality. *American Economic Review* 96 (2), 251–255.
- Hill, M.E. (2002). Skin Color and the Perception of Attractiveness Among African Americans: Does Gender Make a Difference. *Social Psychology Quarterly* 65: 77–91.

- Hirschfield, Paul J. 2008. "Preparing for Prison?: The Criminalization of School Discipline in the USA." *Theoretical Criminology* 79-101.
- Hochschild, J. (2012). Lumpers or splitters: analytic and political choices in studying colour lines and colour scales. *Ethnic and Racial Studies*, DOI:10.1080/01419870.2011.645847
- Hochschild, J. L., & Weaver, V. (2007). The skin color paradox and the American racial order. *Social Forces*, 86(2), 643-670.
- Hughes, M., & Hertel, B.R. (1990). The significance of color remains: a study of life chances, mate selection, and ethnic consciousness among black Americans. *Social Forces* 68 (4), 1105–1120.
- Hunter, M.L. (2007). The persistent problem of colorism: Skin tone, status, and inequality. *Sociology Compass* 1(1), 237-254.
- Hunter, M.L. (2005). *Race, Gender, and the Politics of Skin Tone*. New York: Routledge.
- Hunter, M.L. (2002). "If You're Light, You're Alright": Light Skin Color as Social Capital for Women of Color. *Gender & Society* 16: 175–93.
- Hunter, M. L. (1998). Colorstruck: Skin color stratification in the lives of African American women. *Sociological Inquiry*, 68(4), 517-535.
- Inniss, J.P. (2010). Colorism: The Hierarchical nature of skin tone that makes "light alright." *Everyday Sociology Blog* Accessed May 13, 2012: <http://www.everydaysociologyblog.com/2010/01>
- Jones, T. (2010). Intra-Group Preferencing: Proving Skin Color and Identity Performance Discrimination, 34 *N.Y.U. Review of Law & Social Change* 657-707.

- Keith, V. M., & Herring, C. (1991). Skin tone stratification in the black community. *The American Journal of Sociology*, 97(3), 760-778.
- Keith, V.M. (2009). A Colorstruck World: Skin Tone, Achievement, and Self- Esteem among African American Women. Pp. 25-39 in Evelyn Nakano Glenn (Ed.), *Shades of Difference: Why Skin Color Matters*. Los Altos: Stanford Univ. Press.
- King, Rebecca Chiyoko and Kimberly McClain DaCosta. (1996). “Changing Face, Changing Race: The Remaking of Race in the Japanese American and African American Communities.” In Root, Maria P.P., ed. *The Multi Racial Experience: Racial Borders as the New Frontier*, 1996 (CA: Sage Publications), 227-244.
- Krieger, N., Sidney, S. & Coakley, E. (1998). Racial discrimination and skin color in the CARDIA study: Implications for public health research. *American Journal of Public Health* 88, no. 9:1308–13.
- Kupchik, Aaron. 2010. *Homeroom Security: School Discipline in an Age of Fear*. New York: New York University Press.
- Ladner, J. (1971). *Tomorrow's Tomorrow: The Black Woman*. Garden City: Doubleday.
- Losen, D., & Skiba, R. (2010). *Suspended education: Urban middle schools in crisis*. Montgomery, AL: Southern Poverty Law Center.
- Lewin, T. (2012). Black students face more harsh discipline, data shows. *The New York Times* A11, March 6.
- Lewis, C. W., Butler, B. R., Bonner III, F. A., & Joubert, M. (2010). African American male discipline patterns and school district responses resulting impact on academic achievement: Implications for urban educators and policy makers. *Journal of African American Males in Education*, 1-19.

- Loury, L. D. (2009). Am I still too black for you? Schooling and secular change in skin tone effects. *Economics of Education Review* 28 (4), 428–433.
- Maddox, Keith B., & Stephanie, A. Gray. (2002). Cognitive Representations of Black Americans: Re-exploring the Role of Skin Tone. *Personality and Social Psychology Bulletin*, 27, 250–259.
- Massey, D. S., Charles, C.Z., Lundy, G. & Fischer, M.J. (2003). *Source of the River: The Social Origins of Freshmen at America's Selective Colleges and Universities*. Princeton, NJ: Princeton University Press.
- McFadden, A. C., Marsh, G. E., Price, B. J., & Hwang, Y. (1992). A study of race and gender bias in the punishment of school children. *Education & Treatment of Children*, 15, 140–146.
- McCarthy, J.D. & Hoge, D.R. (1987). The social construction of school punishment: Racial disadvantage out of universalistic process. *Social Forces*, 65, 1101-1120.
- Monroe, C.R. (2013). Colorizing educational research: African American life and schooling as an exemplar. *Educational Researcher*, 42(1) 9-19.
- Morris, E. W. (2007). “Ladies” or “Loudies”? Perceptions and experiences of Black girls in classrooms. *Youth and Society*, 38, 490–515.
- Nakano-Glenn, E. (2009). *Shades of Difference: Why Skin Color Matters*. Stanford University Press, Palo Alto, California.
- Parrish, C.H. (1946). Color Names and Color Notions. *The Journal of Negro Education*, 15, 1, 13-20.
- Pizzi, W. T., Blair, I. V., & Judd, C. M. (2005). Discrimination in sentencing on the basis of Afrocentric features. *Michigan Journal of Race & Law*, 10, 327-355.

- Raffaele Mendez, L. M., & Knoff, H. M. (2003). Who gets suspended from school and why? A demographic analysis of schools and disciplinary infractions in a large school district. *Education and Treatment of Children*, 26, 30–51.
- Raffaele Mendez, L.M., Knoff, H.M. & Ferron, J.F. (2002). School demographic variables and out-of-school suspension rates: A quantitative and qualitative analysis of a large, ethnically diverse school district. *Psychology in the Schools*, 39, 259-277.
- Russell-Brown, K. (1998). *The color of crime*. New York: New York University Press.
- Russell, K., Wilson, M., & Hall, R. (1992). *The color complex: The politics of skin color among African Americans*. New York: Harcourt Brace Jovanovich.
- Ryabov, I. (2013). Colorism and School-to-Work and School-to-College Transitions of African American Adolescents. *Race and Social Problems*, 5(1), 15-27.
- Saperstein, A. (2012). Capturing complexity in the United States: which aspects of race matter and when? *Ethnic and Racial Studies*, 35:8, 1484-1502.
- Saperstein, A. and Penner, A.M. (2012). Racial Fluidity and Inequality in the United States. *American Journal of Sociology*, 118 (3) , 676-727.
- Skiba, R.J., Michael, R.S., Nardo, A.C. & Peterson, R. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *Urban Review*, 34, 317-342.
- 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.

- Spohn, C.C. (2000). Thirty years of sentencing reform: The quest for a racially neutral sentencing process. In J. Horney (Ed.), *Criminal justice 2000: Vol. 3. Policies, processes, and decisions of the criminal justice system* (pp. 427-501). Washington, DC: Department of Justice, National Institute of Justice.
- Taylor, M.C., & Foster, G. A. (1986). Bad boys and school suspensions: Public policy implications for black males. *Sociological Inquiry*, 56, 498-506.
- Telles, E. (2012). The Overlapping Concepts of Race and Color in Latin America. *Ethnic and Racial Studies*, DOI:10.1080/01419870.2012.657209
- Tilly, C. (1998). *Durable Inequality* (Berkeley: University of California Press).
- Vavrus, F. & Cole, K. (2002). "I didn't do nothin'": The discursive construction of school suspension. *The Urban Review*, 34, 87-111.
- Vedantam, S. (2010). Shades of Prejudice. *The New York Times* A31, January 19.
- Viglione, J., Hannon, L. & DeFina, R. (2011). "The Impact of Light Skin on Prison Time for Black Female Offenders." *The Social Science Journal* 48: 250-258.
- Walker, Alice. (1983). *In Search of Our Mothers' Gardens*. San Diego: Harcourt Brace Jovanovich, Publishers.
- Wallace, J. M., Jr., Goodkind, S., Wallace, C., & 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.
- Wald, J., & Losen, D.J. (2003). Defining and redirecting a school-to-prison pipeline. In J. Wald & D.J. Losen (Eds.), *New directions for youth development* (no. 99; *Deconstructing the school-to-prison pipeline*) (pp. 9-15). San Francisco: Jossey-Bass.

- Wald, J., & Losen, D. J. (2007). Out of sight: The journey through the school-to-prison pipeline. In S. Books (Ed.), *Invisible children in the society and its schools* (3rd ed.), (pp. 23–27). Mahwah, NJ: Lawrence Erlbaum Associates.
- Way, Sandra. 2011. “School Discipline and Disruptive Classroom Behavior: The Moderating Effects of Student Perceptions.” *Sociological Quarterly* 52: 346-375.
- Wilder, J. (2010). Revisiting “Color Names and Color Notions” A Contemporary examination of the language and attitudes of skin color among young black women. *Journal of Black Studies*, 41, 184-206.
- Wu, S.C., Pink, W.T., Crain, R.L., & Moles, O. (1982). Student suspension: A critical reappraisal. *The Urban Review*, 14, 245-303.

Table 1: Summary Statistics for African American Youth, by NLSY97 Sample

Panel A: Total

<u>Variable</u>	<u>Obs</u>	<u>Mean</u>	<u>Std. Dev.</u>	<u>Min</u>	<u>Max</u>
Suspend	1175	0.45	0.50	0	1
Skin darkness	1175	6.20	1.92	1	10
SES	1175	-0.20	1.21	-4.68	5.19
Test score	1175	29.57	24.70	0	99.66
Delinquency	1175	1.21	1.59	0	9
Urban	1175	0.50	0.50	0	1
Age	1175	14.94	1.40	13	17

Panel B: Females

<u>Variable</u>	<u>Obs</u>	<u>Mean</u>	<u>Std. Dev.</u>	<u>Min</u>	<u>Max</u>
Suspend	649	0.38	0.49	0	1
Skin darkness	649	6.06	1.93	1	10
SES	649	-0.22	1.17	-4.59	3.46
Test score	649	31.55	24.95	0	99.66
Delinquency	649	0.90	1.33	0	9
Urban	649	0.52	0.50	0	1
Age	649	14.96	1.43	13	17

Panel C: Males

<u>Variable</u>	<u>Obs</u>	<u>Mean</u>	<u>Std. Dev.</u>	<u>Min</u>	<u>Max</u>
Suspend	526	0.53	0.50	0	1
Skin darkness	526	6.37	1.90	1	10
SES	526	-0.17	1.26	-4.68	5.19
Test score	526	27.12	24.19	0	99.02
Delinquency	526	1.59	1.80	0	9
Urban	526	0.47	0.50	0	1
Age	526	14.92	1.36	13	17

Table 2: Bivariate Correlations for NLSY97

Panel A: All African American Youth

	Suspend	Skin darkness	SES	Test score	Delinquency	Urban	Age
Suspend	1						
Skin darkness	0.1217*	1					
SES	-0.1921*	-0.0889*	1				
Test score	-0.2467*	-0.1162*	0.4431*	1			
Delinquency	0.2766*	0.0066	0.0168	-0.0137	1		
Urban	0.1053*	0.0239	0.0099	0.0012	0.0774	1	
Age	0.0361	-0.0247	-0.001	-0.0061	0.0949*	-0.0207	1

* indicates $p < .05$

Panel B: Female African American Youth

	Suspend	Skin darkness	SES	Test score	Delinquency	Urban	Age
Suspend	1						
Skin darkness	0.1235*	1					
SES	-0.2825*	-0.044	1				
Test score	-0.2974*	-0.0768	0.4711*	1			
Delinquency	0.2520*	-0.0008	-0.0126	0.0158	1		
Urban	0.0884	0.0908	0.0139	0.0089	0.1145	1	
Age	0.0803	-0.0305	-0.0112	-0.0305	0.1091	0.0241	1

* indicates $p < .05$

Panel C: Male African American Youth

	Suspend	Skin darkness	SES	Test score	Delinquency	Urban	Age
Suspend	1						
Skin darkness	0.0962	1					
SES	-0.1009	-0.1451*	1				
Test score	-0.1620*	-0.1523*	0.4185*	1			
Delinquency	0.2587*	-0.0204	0.036	-0.0042	1		
Urban	0.1449*	-0.0509	0.0071	-0.0186	0.0694	1	
Age	-0.0141	-0.0148	0.0118	0.0235	0.0943	-0.0802	1

* indicates $p < .05$

Table 3: Multivariate Logistic Regression Models of School Suspension^a
 (respondent-identified race; z-statistic using robust standard error in parentheses)

<u>Variable</u>	<u>Total</u>	<u>Male</u>	<u>Female</u>
Skin darkness	1.112** (3.10)	1.104* (1.99)	1.152** (2.93)
SES	.794** (-3.86)	.923 (-1.02)	.652** (-4.73)
Test score	.999** (-5.79)	.999** (-2.74)	.999** (-5.5325)
Delinquency	1.484** (8.31)	1.417** (5.64)	1.600** (5.91)
Urban	1.532** (3.29)	1.780** (3.04)	1.304 (1.48)
Age	1.028 (.61)	.965 (-.52)	1.078 (1.21)
Male	1.443** (-2.75)	--	--
Constant	.189* (-2.29)	.637 (-.42)	.081* (-2.50)
Pseudo R ²	.134	.093	.166
No. of obs.	1175	526	649

^a+, * and ** indicate p.<.10, p<.05 and p<.01, respectively. Z-statistics for odds ratios less than one are negative.

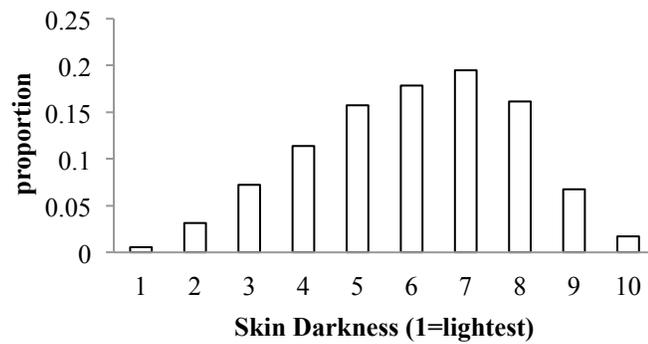
**Table 4: Multivariate Logistic Regression Models of School Suspension^a
based on the Add Health data**

Variable	Female sample			Male sample		
	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>
Skin darkness	1.213* (2.45)	1.190* (2.33)	1.184* (2.04)	1.099 (.86)	1.079 (.66)	.957 (-.33)
SES	--	.600** (-4.12)	.663* (-2.55)	--	.668** (-3.53)	.698** (-2.75)
Test score	--	.981 (-1.85)	.983* (-2.13)	--	.992 (-.89)	.989 (-1.57)
Delinquency	--	1.573** (8.17)	1.569** (7.65)	--	1.413** (5.06)	1.372** (4.93)
Urban	--	1.155 (.71)	1.143 (.67)	--	1.142 (.64)	.996 (-.02)
Age	--	1.052 (.83)	1.070 (1.39)	--	1.023 (.37)	1.094 (1.79)
Percent black	--	--	1.607 (1.07)	--	--	1.055 (.10)
School climate	--	--	.026** (-4.58)	--	--	.053** (-3.12)
Racial inequality	--	--	1.382 (.81)	--	--	.848 (-.33)
Self-esteem	--	--	.752+ (-1.83)	--	--	1.408+ (1.81)
Physical attractiveness	--	--	.942 (-.56)	--	--	.967 (-.23)
Constant	.333 (-3.44)	.281 (-.86)	2062.3** (3.20)	.776 (-.55)	.535 (-.46)	99.046 (1.43)
No. of obs.	1481	1481	1331	1140	1140	1009

^a+, * and ** indicate p.<10, p.<.05 and p.<.01, respectively. t-statistics for odds ratios are in parenthesis.

Figure 1: Distribution of Skin Tone Values, by NLSY97 Sample

Panel A: Female African American Youth



Panel B: Male African American Youth

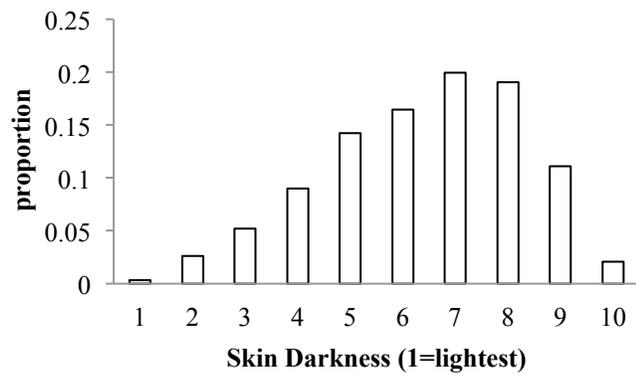
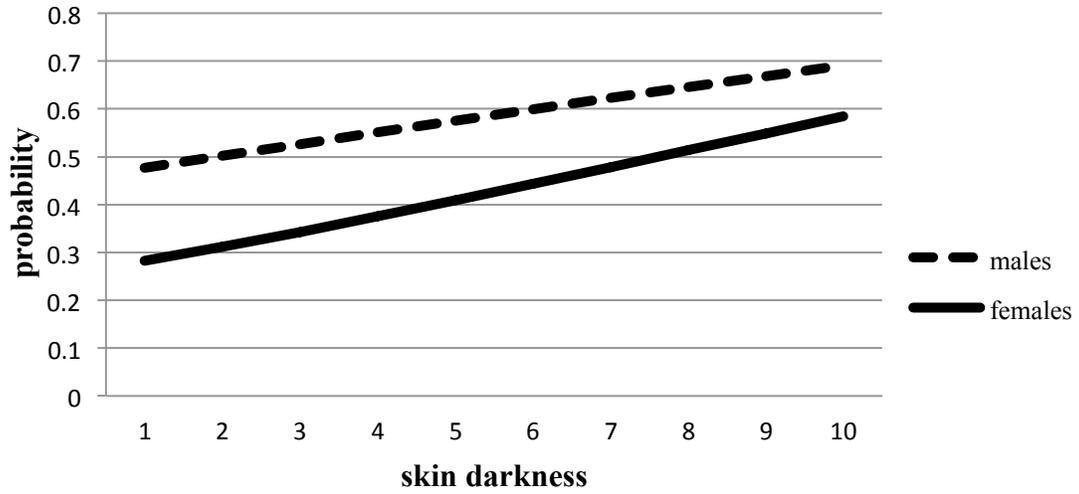


Figure 2: Probabilities of Suspension for African American Males and Females, by Skin Darkness*



*Authors' calculations based on estimates in Table 3.