Northwestern | INSTITUTE FOR POLICY RESEARCH

IPR Working Paper Series

WP-20-49

Public Perceptions of Black Girls and their Punitive Consequences

Sally Nuamah
Northwestern University and IPR

Version: March 1, 2021

DRAFT

Please do not quote or distribute without permission.

Abstract

How do race and gender affect public support for the punishment of Black girls? Across the United States, Black women are imprisoned at twice the rate of white women and Black girls represent the fastest growing juvenile justice population. Despite these jarring statistics, little research exists on the public perceptions contributing to these troubling patterns across race and gender, particularly in public opinion research. This paper uses an original survey experiment of Americans to determine the public perceptions shaping the punishment of Black girls. The analysis reveals that Black girls are seen as older, more dangerous, and more knowledgeable about sex. Further, they are viewed as deserving of harsher punishments than any other student. These findings have serious implications for the study of race, gender, criminal justice, and public opinion as they specify the potential role of the American public in contributing to the uneven punitive experiences of Black girls.

Introduction

In May 2020, Michigan judge, Mary Ellen Brenan, made headlines when it was revealed that she detained a 15-year-old Black girl, referred to as "Grace," for not completing her online homework. Grace joined nearly 25% of public-school students across the United States who had failed to complete their online homework while schools were physically closed due to the global COVID-19 pandemic (Cohen 2020). Why then did she seem to be the only one arrested and incarcerated for it?

Grace's arrest and incarceration drew national attention partly due to its framing by various media outlets as novel, but in reality, the punishment of Black girls is far from unusual. Black girls are punished—through suspension, arrests, and incarcerations—at alarmingly high rates in the United States. At school age, Black girls are suspended seven times more than white girls and more than all non-Black boys. Despite only making up 16% of the student population, Black girls represent 43% of those arrested in school incidents (U.S., Department of Education, Office of Civil Rights 2014). Once arrested, Black girls are nearly four times more likely than white girls to be incarcerated, and, as adults, Black women are imprisoned at twice the rate of white women.

Despite the jarring statistics on the punitive experiences of Black girls**,** little research exists on the public perceptions contributing to these troubling patterns across race *and* gender. The potential political consequences of their punitive experiences are even lesser known. In political science, much of the work on race, criminal justice, and political behavior has focused on Black men and/or aggregated population level data (e.g., Gilens 1996; Peffley and Huruwitz 2010; Enns 2016). Very few published studies in American political behavior have focused on the punitive experiences of Black women and girls, even as they represent a disproportionate percent of those suspended, arrested and detained across the United States.

This paper seeks to intervene on this gap. Using original experimental data, the paper investigates the public perceptions shaping the punitive experiences of Black girls, in particular. The analysis reveals that they are seen as older, more dangerous, and more knowledgeable about sex than their peers. Further, they are viewed as deserving of harsher punishments, in this case, suspension and detention, more than any other student. These findings have serious implications for research on the punitive experiences of Black girls, public opinion and its potential political consequences.

Race, Gender and Public Perceptions

Public attitudes toward marginalized groups often shape their access to, and engagement with, policies (e.g., Gilens 1996; Soss and Schram 2009). For example, Gilens (1996) classic work on welfare attitudes finds that Americans opposition to the policy is directly related to their perceptions of Black recipients – most of whom are women - as lazy and undeserving. Similarly, Soss and Scram's (2009) study of welfare violations finds that those responsible for welfare distribution are more willing to sanction Black women (and Latinx women) for committing the same behaviors as white women. They determine that the uneven distribution of sanctions for the same actions are due, in part, to racially bias perceptions of welfare recipients.

Both works have clear implications for the policy experiences of Back women. Yet, like much of American politics research, they do not center their experiences. More typically, Black women are discussed as an outcome of the analysis, rather than an integral part of the theoretical foundation. In other words, race and gender are not examined jointly, nor is intersectionality a central component of their conceptualization (Hancock 2004).

Established research on intersectionality, however, makes clear that race and gender are mutually constitutive for Black women, and therefore, neither a race nor gender lens alone can

explain their experiences (see for ex. Collins,1990; Crenshaw, 1989). Regarding the criminal justice system, specifically, Black women experience racialized forms of punishment due to their blackness, while simultaneously experiencing gendered forms of punishment, due to their womanhood (Ritchie and Jones-Brown, 2017). Accordingly, existing at the intersections of multiple axes of oppression – race and gender - should produce unique experiences within the criminal justice system for Black women.

Intersectional experiences with punishment, in particular, are rooted in racist and sexist practices beginning with slavery (e.g., George 2015). During slavery the sexual exploitation of Black women was justified through the development of stereotypes that labeled Black women as seductive, hypersexual, and immoral. Now ingrained in American culture, these stereotypes created a hierarchy of femininity, in which white women, understood to be sexually pure and moral, represented the feminist ideal, while Black women, understood to be sexually promiscuous, represented a deviation. As Black women were cast as a deviation from the norm, institutions responded by inflicting "social correction" - forms of punishment aimed to "fix" their behavior (George 2015, 102). Many of these same perceptions of Black women persist today and affect the ways in which they are disproportionately punished. In other words, sexism and racism often interact in shaping the carceral experiences of Black women. And yet, the racialized *and* gendered challenges posed by the contemporary carceral state on Black women are often rendered invisible, especially in political science.

To be sure, there is an increasing literature on the political evaluations of Black women, either as political candidates or voters (e.g., Sigelman and Welch, 1984; Philpot and Walton Jr, 2007; Lemi and Brown, 2019). This research engages in important empirical research at the intersection of race and gender, thereby constituting an analysis that is more accurately illustrative

of how Black women experience their lives. Most notably, these works reveal the dual burden of being at the axes of multiple oppressions, often finding that stereotypes of Black women (e.g., angry) shape how the public perceives them and thus lessens their likeliness to support them without a higher burden of proof (e.g., Philpot and Walton jr., 2007; Hancock 2004, 2007a, 2007b). Yet, this work rarely investigates the carceral experiences of Black women, the public perceptions shaping them, or their political impacts.

A notable exception is political scientist Ange Marie-Hancock (2004), whose classic work shows how stereotypical depictions of Black women as lazy, hyper-fertile, and irresponsible mothers - based on both race and gender - shaped the public debate surrounding welfare reforms and ultimately influenced the policies that followed. In particular, she demonstrates how African American legislators themselves voted for welfare reform (which disproportionately punishes Black women), reaffirming the same negative tropes of Black women as white legislators. This work provides an uncommon, but crucial example of how public perceptions of Black women can have serious consequences on the punitive policies that affect their lives.[1]

The Punishment of Black Girls

The same negative perceptions that affect the experiences of Black women either start with, or trickle down to, Black girls (Harris-Lacewall 2008).[2] Research in psychology, childhood studies,

[1] It is important to note that it is even less common to see research on Black women and stereotypes beyond the "welfare queen."

[2] According to Harris-Lacewall (2008), "The centrality of strength to African American women's self-concept is further reflected in the lessons that black women pass on to black girls (6)…The strong black woman is at the center of this socialization project for black girls (7).

and sociology, for example, finds that Black children are often dehumanized and thus perceived as in need of less socio-emotional support when exhibiting the same or similar behaviors as white students (e.g., Goff et al 2014; Okonofua and Eberhardt 2015). These perceptions, according to literature, are also a product of stereotypes, for example the Strong Black Women trope that makes medical physicians less likely to take Black women's pain seriously, or the Angry Black Women trope that makes police officers more likely to arraign Black women for mental breakdowns (see for example Young, 1989; Weisse 2003; Perry 2011). In sum, this literature suggests that negative stereotypes of Black women, and their negative impacts, likely affect Black girls much before they go on to become Black women, let alone political candidates.

Indeed, while Black women represent a disproportionate percent of women in prison, Black girls – starting as early as pre-school – represent a disproportionate percent of those suspended, arrested and eventually detained. Within a school context, these gendered punishments are reflected in the ways they are pushed out of school and into confinement (Morris 2016; Ritchie 2017). For example, between 2013 and 2014, 20% of female preschoolers were Black, but Black girls comprised 54% of female preschool children with one or more suspensions (U.S. Department of Education 2014). In elementary and middle schools, Black girls are five times more likely to be suspended from school than their same-gendered peers (Epstein et al., 2017). Their disproportionately punitive experiences do not stop at the schoolhouse, however.

At the courthouse, Black girls also receive harsher punishments compared to white girls. In a study conducted by the American Bar and National Bar Association, they found that seven of every 10 cases involving white girls were dismissed by prosecutors, compared with only three of

every 10 cases involving Black girls (American Bar and National Bar Association 2001). Further, Black girls often fail to receive equal opportunities for diversion - strategies that prosecutors may assign as disciplinary measures instead of formal processing. This treatment extends to the foster care system, where Black girls are three times more likely to be removed from their homes and placed in state custody than their white peers (Roberts 2009; Chesney-Lind and Jones 2010).

Across multiple systems – education, criminal justice, foster care - the unfair nature of Black girls' punishment is clear. And yet, the public's role in shaping these experiences is not. So, what are the public perceptions shaping Black girls? How do race and gender stereotypes affect public support for their punishment? The next section provides a brief review of literature on this topic before delving into the specific goals and contributions of this investigation.

Perceptions of Black Girls and their Punitive Consequences

Established research on the punitive experiences of Black girls argues that these disparities are rooted in racism and sexism (Wun 2016; Morris 2007). As stated by Priscilla Ocen, "histories of racial and gender subordination, including slavery and Jim Crow, …interacted with the category of childhood to create a liminal category of childhood that renders Black girls vulnerable to sexual exploitation and criminalization (Ocen 2015, 1600)." Since Black girls were not considered children - in the traditional sense of the term - they were not provided with the privileges that childhood affords, for example, innocence.

In the contemporary period, this historical notion of Black girls as occupying a type of liminal childhood interacts with current group perceptions of Black girls' behavior (e.g., being too talkative, loud, knowledgeable) as inconsistent with traditional feminine norms (Crenshaw, 1989; Crenshaw et al., 2014). The criminalization of Black girls, in particular, is often based on their perceived defiance of these norms (Evans-Winters and Esposito, 2010). Just as is this case with

Black women, those who do not conform with the ideals of femininity (e.g., docile, meek, polite, quiet) are targeted and punished as a mechanism of "correcting" their behavior. This practice of engaging in punishment as mechanism for correcting Black girls' behavior often begins in the classroom (Wun 2016b).

Over the past three decades, schools have mirrored aspects of the criminal justice system, particularly through the use of zero tolerance policies (Blake et al 2011; Hines-Datiri and Andrews, 2020). These zero tolerance policies – established in 1996 as mandatory punishments for perceived rule violations – have contributed to the outsourcing of disciplinary responsibilities from schools to juvenile courts and school resource officers. Similar to prisons, the students most affected by these policies are Black and brown, with Black students, for example, making close to 30% of in school-arrests, despite only making up of 16% of the school population (Nelson and Lind, 2015). The outcome of these actions is a school system more akin to the carceral system.

The carceral experiences of Black students are also gendered through the rules and regulations included in the school code of conduct (Aghasaleh 2018). Many dress code requirements, for example, only apply to girls' attires, such as skirt/shorts length and sleeve coverage (Kosciw et al., 2016). These gendered regulations are often justified with the rationale that female students' bodies can act as a distraction to their male peers and that it is "unlady-like" to dress in ways deemed inappropriate according to traditional gender norms. The consequence, however, is that girls' bodies are policed into submission as violations of them often result in their punishment.

Black girls are doubly disadvantaged by policies such as these because of *how their bodies are viewed* by the public. Historically, Black girls were accused of being sexually promiscuous or "Jezebels," suggesting that they had heightened sexual appetites (e.g., French 2013). Existing

research suggests that these perceptions continue to shape the ways that Black girls are viewed compared to white girls today, thus demonstrating "how deeply entrenched controlling narratives of Black women and girls are – no matter how young and small they are" (Ritchie 74, 2017). Returning to school policies, specifically, this means that even if Black girls are not actively violating a dress code, they may be perceived as intentionally doing so because of their perceived hypersexuality (French, 2013; Townsend et al., 2010).

The perception of Black girls as hypersexual undermines their ability to claim their childhood, and often makes them automatically more responsible for their actions (Morris 2016; Epstein el 2017). Perhaps the most notable recent example is that of a 9 year-old Black girl who was arrested and pepper sprayed by police officers after throwing a tantrum in Rochester, New York in January 2021. Following the incident, it was reported that "at one point, one officer says, 'You're acting like a child,' to which the girl can be heard responding, 'I am a child!'" (Ly and Levenson 2021). The exchange between the officer and the 9 year-old girl illustrated the ways in which Black girls are explicitly aged-up and thus held responsible for actions that would be considered normal for most children. This perception of Black girls as un-child-like, or rather, adults, likely play an important role in how they are both perceived and thus punished across the nation. Nonetheless, this has not been investigated in the race, criminal justice and political opinion literature.

Expectations: How the Adultification of Black Girls Shape their Punishment

The lion share of existing studies in political science on the relationship between public attitudes and punishment tends to focus on the perceptions of Black men (e.g., Peffley and Huruwitz 2010). This work finds that most Americans support Black men's incarceration, for example, if they perceive the justice system as fair (Enns 2016). While established research on Black men finds

that perceptions of fairness shape attitudes toward punishment, the emerging literature on Black girls (e.g., Epstein et al., 2017) suggest that perceptions of them as more adult-like (e.g., mature, knowledgeable about sex) would lead the respondents to indicate that Black girls deserve more punishment, regardless of belief in fairness.

The "adultification" of Black girls—the process by which Black girls are viewed as more adult-like and, thus, less innocent relative to white girls of the same age—likely play a significant role in how they are disciplined. A 2017 study of attitudes toward Black and white girls from infancy to 19 years of age revealed that white respondents view Black girls as adults as early as the age of five; that is five years earlier than Black boys and significantly earlier than every other demographic group studied (Epstein et al., 2017). Further, they found that Black girls were viewed as in need of less nurturing, protection, support, comfort and were perceived as more knowledgeable about sex and adult topics then white girls.

The authors conclude that these perceptions of Black girls likely shape their disproportionate experiences with punishment, **but this specific relationship – between perceptions of adultification and punishment - has not been tested.** Further, prior work only compares the experiences of Black girls to white girls, thus we know very little about how Black girls' experiences compare to white boys and Black boys.

Accordingly, this paper expands on this important work by investigating the public perceptions shaping the punitive experiences of Black girls. In particular, it examines how race, gender and age (as measured by their status as a school student)[3] shape public perceptions of Black

[3] The experiment does not specify age in order to avoid priming respondents on the basis of age as opposed to race and gender.

girls, and the extent to which they are perceived as more deserving of punishment relative to others. To investigate this, I propose and test the following hypotheses:

H1: Black girls are viewed as older than their peers

H2: Black girls are viewed as more of a threat (or a danger) than their peers

H3: Black girls are viewed as more sexual than their peers

H4: Black girls are viewed as deserving of more punishment than their peers

Overall, I expect the public to perceive Black girls as more deserving of punishment than their peers. Further I expect this relationship to be, in part, dependent upon the extent to which students' behavior is consistent with particular racial and gender stereotypes. Black girls, in particular, are viewed through a stereotype of hyper-sexualization - as described above (see also Perry 2011). The perception of Black girls as hypersexual should play a role in how responsible they are deemed for their actions. This should also shape perceptions of how much of a threat they are to their peers, in that Black girls may be viewed as able to potentially corrupt others with their "deviance."[4] Therefore, one would expect Black girls to be perceived as older, more threatening, and more knowledgeable about sex in ways that directly shape public support for their punishment. In other words, racist and sexist perceptions of Black girls should undergird public support for their disproportionate punitive experiences.

[4] This conception of Black girls as a danger or a threat is not to be confused with the idea of Black boys or men as "violent," but rather as connected to their potential deviance due to perceptions of them as "hypersexual."

Methods, Data, and Strategies for Analysis

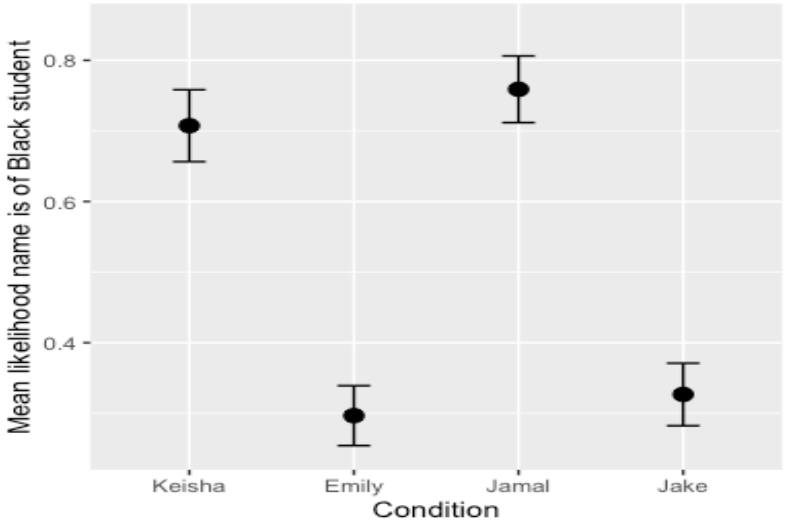
The above hypotheses were tested on one of four scenarios, containing four conditions each, presented to a representative sample of 1466 adults between March 8th and March 15th, 2020.[5] At the most general level, the scenario of focus tests whether evaluations of students' behavior and support for the punishment of students is simultaneously affected by students' race and gender. More specifically, to examine perceptions of punishment, respondents were randomly assigned a scenario regarding a dress code violation that varied by race and gender using the names Keisha (Black girl), Emily (white girl), Jamal (Black boy), and Jake (white boy). Following previous research, perceived race and gender of the individuals in the vignettes are manipulated through names only. The names were selected based on those used in previous work, or which previous

[5] Though online non-probability samples are not nationally representative and, hence, inappropriate for descriptive research or inferences (Baker et al. 2013), they are suitable for experimental survey research. Extant research shows that 1) these samples are more diverse than other types of samples commonly used in experimental research, such as college students (Berinsky, Huber, and Lenz 2012); and 2) experimental effects identified among probability-based, nationally representative samples are replicable among online, non-probability based samples, including those provided by Lucid (Coppock and McClellan 2019). We also engaged in a replication of this study in October 2020 (discussed below and in the appendix) to further test the effects.

work indicates as being disproportionately common among particular racial-gender subgroups (Bertrand and Mullainathan 2004; Okonofua and Eberhardt 2015; Gilliam, et al., 2016).

To demonstrate the effectiveness of the racialized names, Figure 1 below displays how likely it was that participants thought the name of the student they read about was the name of a Black student. Responses were rescaled to range from 0 to 1. The figure also provides bands indicating 95% confidence intervals, which enable judgments from the figures about which responses are statistically distinguishable from one another. As the figure shows, the names Jamal and Keisha were significantly more likely (M=0.76 se=0.03; M=0.71 se= 0.03) than the names Jake and Emily (M=0.33 se=0.02; M=0.29 se=0.02) to be thought of as belonging to Black students. The difference is about 40 percent between the Black and White names of students, suggesting the manipulation was effective. These differences are robust across regression analyses controlling for demographic variables.

Figure 1: Mean Likelihood That Name is of a Black Student by Experimental Condition



To test the hypotheses, using the above names, the specific scenario offered to respondents was the following:

> Consider an instance where a student named _____ is wearing shorts to school. When _____ arrives to class, the teacher tells _____ that _____ shorts violate the school's dress code policy. The teacher tells _____ to leave class and go to the office. This also isn't the first time that _____ has worn clothes that violate the school's dress code policy. For example, _____ has previously worn tank tops that are against school rules.

The content of the dress code scenario was developed based on qualitative research (e.g., Morris 2016) that demonstrates that the punitive experiences of Black girls begin at the schoolhouse and are commonly overlooked due to their racialized *and* gendered nature.[6] The intersection of both makes them difficult to identify in studies that are focused on just issues of race *or* gender. Thus, the vignette is meant to capture a scenario that could feasibly happen to both boys and girls but could still be perceived differently if viewed through a racialized and gendered lens (e.g., how clothing may be viewed on Black girls vs. white girls). Dress code violations, specifically, are commonly referenced as an example of this in the relevant literature (see Morris, 2016; Perry, 2011; Evans-Winters and Espositio, 2010).

[6] The decision not to use a scenario focused on fighting, for example, was made in hope that the analysis could avoid re-creating tests that have the potential to erase the gendered nature of punishment. Indeed, part of why girls' punitive experiences have been ignored is because of their differences from the masculinized ways that carceral experiences are typically measured (e.g., incarceration as opposed to being forced to change clothes). Accordingly, the investigation wanted to center those events that girls tend to find themselves in as opposed to study traditional scenarios typically attributed to boys.

The dress code scenario was followed by evaluations of culpability or innocence, measured by replicating three items developed by Goff et al., (2014). These items were originally designed to measure the perceived innocence of children within a criminal justice context and, in this case, have been adapted to measure the perceived innocence of children within an educational context. Adultification, specifically, is measured by replicating items used by Goff et al., (2014) and Epstein, Blake, and Gonzales (2017), who respectively developed scales measuring qualities associated with adulthood and racial and gender stereotypes that individuals might assign to, and subsequently evaluate Black children as more adult-like. In particular, participants were asked 1) if the student was acting older than their age 2) if the student posed a danger or threat to others, 3) if the student was knowledgeable about sex, and 4) if the student was provided with the appropriate level of punishment for their behavior. Since all other properties of the scenarios and questionnaire were identical, any observed differences between conditions in responses to the questions about the scenario can be attributed to the race and gender, and the behavior of the student described in the questions.[7]

Table 1. Sample Demographics

Total n	1,466
Age (mean)	41.4 years (median=39)
Female	49.1%
Male	50.1%
Transgender/Gender non-conforming	0.8%
Education	
Less than high school diploma	0.4%

[7] Given the general possibility of question order effects the order in which all covariates and individual items are measured is randomized, and any identities or demographics possibly related to relevant covariates and dependent measures are measured at the end of the surveys (e.g., racial and gender identity).

High school diploma	9.9%
Some college	28.4%
College graduate	43.3%
Graduate or professional school	17.9%
Annual Household Income	
Less than $10,000	2.9%
$10,000-$39,999	28.2%
$40,000-$69,999	33.7%
$70,000-$99,999	19.2%
$100,000 and above	15.9%
Race and ethnicity	
White	77.8%
Black	8.9%
Hispanic/Latinx	4.6%
Asian	6.3%
Other race or ethnicity	2.3%
% Living w/school-age children	34.7%
Partisanship	
Democrat/Lean Democrat	55.7%
Republican/Lean Republican	32.9%
Pure Independent	11.4%
Ideology (1-7; 7=extremely conservative)	
Mean	3.6
Median	3 (Slightly liberal)
Mode	2 (Liberal)

Table 1 provides data on the basic demographics of the final sample. The sample is primarily white and relatively highly educated, democratic, and liberal.[8] Nevertheless, the sample is more diverse than many samples commonly used in experimental survey research and consistent with research on MTurk samples (see e.g., Berinsky, Huber, and Lenz, 2012; Huff and Tingley, 2015).[9] The sample population was randomly distributed across multiple scenarios, one of which

[8] Further, the sample's liberal leaning should act as a more rigorous test of the hypotheses.

[9] One of the benefits of using Mturk is that it is much better understood than other platforms with dozens of articles written about its benefits and limitations. Many have questioned the external

was the dress code scenario described above. Those presented with the dress code scenario were then asked the following questions in Table 2 (see also table A.2 in the appendix):

Table 2. Dependent Variables

DV1 (adult2): How much do students like ___act older than their age?
DV2 (adult3): How much are students like ___ danger to others?
DV3 (adult5): How are students like ___experienced with sex?
DV4 (suspension): How harsh is suspension is for students like ___?

The next section displays experimental conditions for those variables where significant differences between the Keisha condition and other conditions exist. The figures provide 95% confidence interval bands, which enable judgments from the figures of which responses are statistically distinguishable from one another. Further, all measures are scaled 0 to 1 for ease of interpretation.[10]

All statistically significant differences were further analyzed with OLS regression that controlled for respondent age, gender, race/ethnicity, education, income, and ideology (See table A.3 in Appendix). Two types of models were run: one with dummy variables for the race (1=Black) and gender (1=female) of the student in the treatments and their interaction, and one with dummy

validity of these samples due to the type of respondents that engage with the platform, however the developments of the program have made it much easier to ensure a diverse representative sample. While there have been concerns that Mturk participants are motivated by financial incentives and thus less attentive, several studies have demonstrated that their attention is the same, if not better than survey pools that rely on college students (See Huff and Tingley, 2015).

[10] One can thus determine the percentage by multiplying 100xB (coefficient). In addition, the results for all respondents (as opposed to compliers) are reported - as a conservative measure of the findings.

variables for each experimental condition with the Black female student condition serving as the baseline. The latter allows for the identification of statistical differences between the Black female student condition and other conditions (for more details on the power analysis see appendix).[11] The next section details the findings of the experimental analysis.

Findings

Figure 1 Mean Perception of Students' Acting Older than Age by Experimental Condition

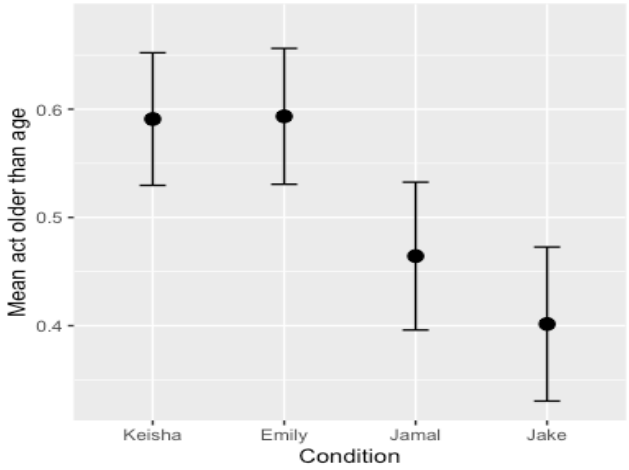
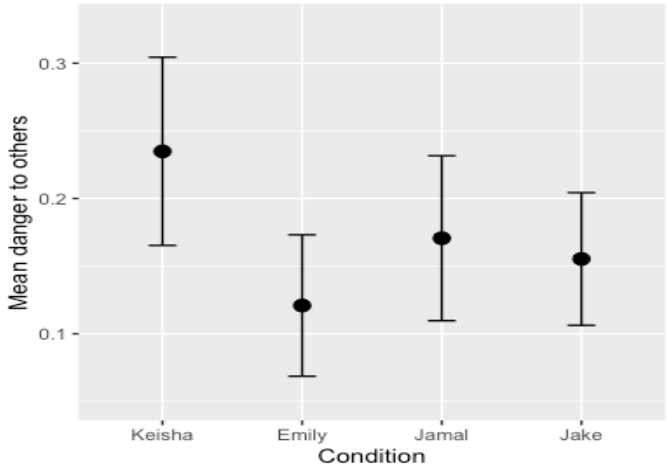


Figure 1 displays participant responses to the question of how much students like the one in the dress code scenario act older than their age. The figure provides evidence that adultification is gendered. In particular, the data reveals that students like Emily and Keisha (M= 0.59 se=0.03 for both) were both perceived to act older than students like Jake or Jamal (M=0.40 se=0.04; M=0.46 se=0.03). In fact, there was a nearly 19 percentage point difference in how Emily and Keisha were

[11] The survey had sufficient power to detect medium-sized effects (Cohen's d, or the standardized effect size of approximately 0.5). At the 95% confidence level, the pilot survey has power of about 0.90 which is far above the conventional threshold of 0.8. Under powered analysis are not included in paper but are included in appendix.

viewed as compared to Jake, and 13 percentage point difference when compared to Jamal.[12] The effect of the students' genders on participant responses is also robust in the regression analysis controlling for demographic variables.

Figure 2: Mean Perception of Danger of Students to Others by Experimental Condition



There is also evidence that perceptions of dress code violations are not only gendered but also racialized. Figure 2 shows that Keisha was thought to be significantly more of a threat than Emily (M= 0.24 se= 0.03; M= 0.12 se=0.02). In particular, respondents were twice as likely to view Keisha as a threat or danger to others, and this finding is significant at the 0.05 level. In contrast, there were no differences in how dangerous Emily, Jamal, and Jake were perceived to be (M=0.12 se= 0.02; M=0.17 se=0.03; M=0.16 se=0.02). This suggests that Black girl students are evaluated through a unique racial and gender lens.

[12] There are no significant differences between the perceptions of students like Emily and Keisha, or those like Jake and Jamal.

Figure 3: Mean Perception of Students' Knowledge of Sex by Experimental Condition

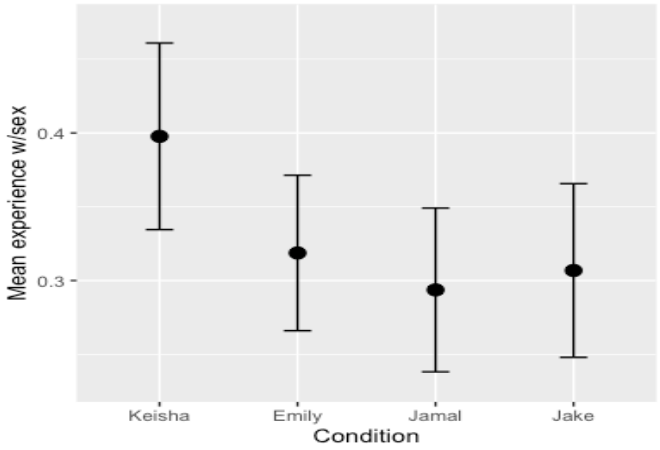
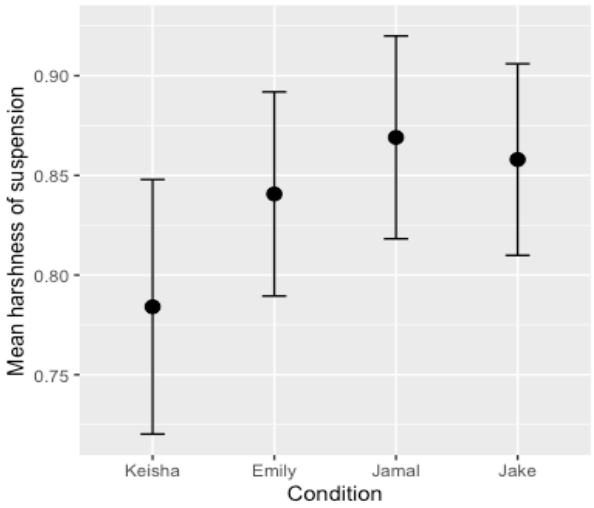


Figure 3 shows that Keisha was also thought to be nearly more knowledgeable about sex than other students (M=0.40 se=0.03), although differences in how people evaluated Keisha and Jamal (M=0.40 se=0.03; M=0.29 se=0.03) and Keisha and Jake (M=0.40 se=0.03; M=0.31 se= 0.03) are only significant at the 0.10 level. The differences between Keisha and Emily (M=0.40 se =0.03; M=0.33 se = 0.03) are significant at the 0.05 level. Across conditions, respondents rate Keisha 7 to 10 percentage points more likely to be Knowledgeable about sex.

Figure 4: Mean Harshness of Suspension for Behavior by Experimental Condition



Finally, figure 4 shows that suspension was thought to be more (unfairly) harsh for all other students than Keisha. In other words, suspension was seen as most appropriate when the student was Keisha (M=0.78 se=0.03). The nearly ten percentage point differences between the Keisha, Jamal and Jake (M=0.78 se =0.03; M= 0.87 se=0.02; M=0.86 se=0.02) conditions are significant at the 0.05 level (Cohen's D = 0.58, power = 0.97). Differences between the Keisha and Emily (M=0.78 se=0.03; M=0.84 se=0.02) condition are significant at the 0.10 level.

Discussion and Additional Considerations

Figures 2-4 are consistent with research on the intersectional nature of the punishment of Black girl students in schools (e.g., Morris 2016). They provide indirect evidence that Black girl students are evaluated and seen as more threatening due to racialized and gendered beliefs about sexuality and dress. In particular, the fact that the investigation begins to reveal differences between Emily and Keisha once the term danger or threat is included suggests that the public associates the dangerous aspects of being "older," for example the ability to corrupt, or be a negative influence, with Keisha in a way that they do not associate with Emily. These differences help to tease out how gender alone, may contribute to perceptions of young girls as "maternal" and thus older or adult-like, but when interacted with race affirm public conceptions of Black girls as "deviant." Altogether, the result of this is that, while both Emily and Keisha are viewed as older, Kiesha is also viewed as relatively more dangerous, knowledgeable about sex, and appropriately punished for violating the dress code.

The above findings are not without their limits, however. One could argue, for example, that because this experiment was embedded in a larger survey, there was a relatively low number of participants involved thus affecting the power of the results. To address this potential concern, a power analysis (included in the appendix) was conducted and revealed that in the areas most

relevant to the investigation the sample size was sufficient to support the analysis. While this smaller sample could skew the reported results, I suspect that this in fact provided a more conservative illustration of the findings. To be sure of this, however, we engaged in a replication of the study in October 2020 – nearly six months after the first experiment - on CloudResearch. Since this relatively short period was fraught with multiple major events - a global pandemic, contested election and racial uprisings - it is unclear to what extent it is a better test than the first. Nonetheless, it allowed us to double the sample population (n = 2266) and ask more questions regarding the extent to which perceptions of Black girls contribute to their punishment. The full analyses can be found in the appendix as well, but in sum it reveals that respondents (especially compliers) viewed Keisha as the most knowledgeable about sex, the most responsible for her actions, and the most likely to continue violating the school dress code. In a similar vein, more severe punishments (suspension and detention) were seen as being less harsh for Keisha than her peers – even though they all committed the same action (see figure 7-2 and 7-3 in appendix).

Finally, tables' 2A and 2B in the appendix explicitly show whether and how stereotypes surrounding Black girls (e.g., Black girls being more sexual or mature than their peers) affect people's punitive sentiments.[13] In particular, table 2A reveals that respondents who viewed Black girls as more sexual were more likely to support their detention, suspension and expulsion. Further, table 2B reveals that respondents who viewed Black girls as more mature were more likely support more severe punishments for them including changing their clothes, detention, suspension and

[13] Note that *n* is around 550 – this is because we randomly assigned one of the four question sets (i.e., regarding Black women, Black girls, white women, and white girls) for survey length reasons.

expulsion. Both tables reveal robust, significant, and positive coefficients for these stereotype variables, demonstrating that those who believe that Black girls are more sexual and mature than their peers were also more supportive of punishment, regardless of the experimental condition they were assigned. All these findings support the initial hypotheses: public perceptions of Black girls shape support for their uneven punishment.[14]

Conclusion

Over fifteen years ago, Evelyn Simien wrote, "empirical assessments of the simultaneous effects of race and gender are indeed rare" in political science (2005, 531). Today, a growing number of political scientists have taken on this work, demonstrating how categories of marginalization intersect with each other to shape Black women's political lives (e.g., Hancock 2004; Smooth 2006; Bonilla and Tillery 2020). Using the theoretical and empirical tools of intersectionality, specifically, these scholars have largely focused on Black women as political candidates, and to a

[14] I should also note that while the theoretical framework for this analysis was conceived of using intersectionality, the use of experimental data on public opinion limits the ability to conduct an investigation that fully captures what is meant by the concept. Instead, it examines race and gender categories together to determine their compounded impact on public perceptions. In so doing, it provides at best, an estimate or approximation of how the public perceives young people at the intersection of race and gender and its implications for their punishment.

lesser extent, evaluators of politics (Hancock 2007a, 2007b; Simien 2007; Brown, 2014). The carceral experiences of Black women and/or girls, however, remains underexplored. [15]

Accordingly, this investigation places the punitive experiences of Black girls at the center of research on race, gender and American public opinion. In so doing, it affirms established research on the ways that racialized and gendered stereotypes interact to yield distinct punitive experiences, particularly as it relates to Black girls (Crenshaw 1989; Collins 1990). Further, it lends evidence to research that illustrates how Black girls may be punished by policies that fail to realize the unique punishments that they experience as a result of their intersecting, marginalized identities.

As Ange-Marie Hancock explains in her classic work on intersectionality, research that has conflated "group unity" with "group uniformity" - assuming that individuals who share one marginalized identity [black *or* female] have uniform experiences of discrimination - are incomplete (Hancock 2007, 65). The policy solutions that emanate from this sort of research, Hancock explains, often benefit white women at the cost of women of color, and often exacerbate existing inequities (Hancock 2007). The limitations of these policies exist for Black girls as well, thus making them the disproportionate and regular recipients of punitive policies. This

[15] It is important to note that even research on race-class subjugated communities and the carceral state in political science tend not to take seriously the role of gender and its intersections with race.

investigation specifies the role of the American public in contributing to the uneven punitive experiences that these policies produce for Black girls. [16]

Beyond public opinion, this work raises serious questions about the consequences of Black girls' punishment for democracy at large. Indeed, what lessons might Grace draw from her current experiences with the justice system, and how might they impact her relationship with the government as an adult? Research on the punishment of Black men has found that incarceration not only lowers the political participation—voting, protesting, attending community meetings—of those who have felony convictions but also that of their families and neighbors (e.g., Burch 2013). Studies on the political effects of punishment for Black women are less common, but there may be similar negative consequences of high rates of incarceration and detention on what has been record levels of political participation (Brown 2014; Farris and Holman 2014).

Ultimately, given the superlative participation of Black women, one would expect the increasingly punitive experiences of Black girls to have lasting consequences on the future strength of American democracy – especially as they become adults. Still, before Black women become adult voters and/or candidates, they are Black girls. And as this investigation reveals, Black girls are disproportionately punished, in part, due to public attitudes toward them. If American democracy is only as strong as its participants, then the punishment of Black girls and the perceptions shaping them, must not only be understood but also dismantled. In short: the punitive experiences of Black girls can no longer be ignored.

[16] Hancock (2004) makes clear how public opinion can be especially consequential for the development of punitive policies for Black women. I suspect this to still be the case for Black girls in the contemporary period.

References

Aghasaleh, Rouhollah. 2018. "Oppressive curriculum: Sexist, racist, classist, and homophobic practice of dress codes in schooling." *Journal of African American Studies* 22, no. 1: 94-108.

American Bar Association and National Bar Association. 2001. *Justice by gender: the lack of appropriate prevention, diversion and treatment alternatives for girls in the justice system: a report.*

Baker, Reg, J. Michael Brick, Nancy A. Bates, Mike Battaglia, Mick P. Couper, Jill A. Dever, Krista J. Gile, and Roger Tourangeau. 2013. "Summary report of the AAPOR task force on non-probability sampling." *Journal of survey statistics and methodology* no.1(2): 90-143.

Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. 2012."Evaluating online labor markets for experimental research: Amazon. com's Mechanical Turk." *Political analysis* 20, no. 3: 351-368.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination." *American economic review* 94, no. 4: 991-1013.

Bonilla, Tabitha, and Alvin B. Tillery. 2020. "Which Identity Frames Boost Support for and Mobilization in the #BlackLivesMatter Movement? An Experimental Test." *American Political Science Review* 114, no.4 (2020): 947-962.

Burch, T. 2013. *Trading democracy for justice: Criminal convictions and the decline of neighborhood political participation*. University of Chicago press.

Brown, Nadia E. 2014. “Political Participation of Women of Color: An Intersectional Analysis.”

Journal of Women, Politics & Policy no. 4(35): 315–48.

Cohen, Jodi. 2020. “Judge Won’t Free Michigan Teenager Sent to Juvenile Detention After Not Doing Online School Work.” *Propublica*, Accessed, July 20, 2020 https://www.propublica.org/article/judge-wont-free-michigan-teenager-sent-to-juvenile-detention-after-not-doing-online-schoolwork

Collins, Patricia. H. 1990. *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. Routledge.

Collins, Patricia Hill. 1993. "Toward a new vision: Race, class, and gender as categories of analysis and connection." *Race, Sex & Class*, 25-45.

Coppock, Alexander, and Oliver A. McClellan. 2019. "Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents." *Research & Politics* no. 1 (6): 1-14.

Chesney-Lind, Meda, and Nikki Jones (Eds.) 2010. *Fighting for girls: New perspectives on gender and violence*. SUNY Press.

Crenshaw, Kimberly. 1989. Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *U. Chi. Legal f.*, 139 – 167.

Crenshaw, Kimberly. W. with Priscilla. Ocen and Jyoti Nanda. 2014. “Black Girls Matter: Pushed Out, Overpoliced, and Underprotected,” African American Policy Forum and Columbia Law School Center for Intersectionality and Policy Studies.

Enns, Peter. K. 2016. *Incarceration Nation: How the United States Became the Most Punitive Democracy in the World*. New York: Cambridge University Press.

Epstein, Rebecca. and Blake, Jamila. and González, Thalia 2017. Girlhood Interrupted: The

Erasure of Black Girls' Childhood. *Georgetown University Law Center on Poverty and Inequality, report*, 1-24.

Evans-Winters, Venus. E., & Esposito, Jennifer. 2010. Other People's Daughters: Critical Race Feminism and Black Girls' Education. *Educational Foundations*, no. 24 (1-2): 11-24.

Farris, Emily. M., & Holman, Mirya. R. 2014. Social capital and solving the puzzle of Black women's political participation. *Politics, Groups, and Identities*, no. 2(3): 331-349.

French, Bryana H. 2013. "More than Jezebels and freaks: Exploring how Black girls navigate sexual coercion and sexual scripts." *Journal of African American Studies* 17, no. 1: 35-50.

George, Janel. 2015. Stereotype and School Pushout: Race, Gender, and Discipline Disparities. *Arkansas Law Review*, no.68 (101): 102-129.

Gilens, Martin. 1996. "Race coding" and white opposition to welfare. *American Political Science Review*, no.90 (3): 593-604.

Gilliam, Walter S., Angela N. Maupin, Chin R. Reyes, Maria Accavitti, and Frederick Shic. 2016. Do early educators' implicit biases regarding sex and race relate to behavior expectations and recommendations of preschool expulsions and suspensions. *Yale University Child Study Center*, no. 9(28), 2016.

Hancock, Ange-Marie. 2004.*The politics of disgust: The public identity of the welfare queen*. New York, NY: NYU Press

Hancock, Ange-Marie. 2007a. When Multiplication Doesn't Equal Quick Addition: Examining Intersectionality as a Research Paradigm. *Perspectives on Politics*, no.*5*(1): 63-79.

Hancock, Ange-Marie. 2007b. "Intersectionality as a normative and empirical paradigm." *Politics & Gender* 3, no.2: 248.

Hines-Datiri, Dorothy, and Dorinda J. Carter Andrews. 2020. "The effects of zero tolerance policies on Black girls: Using critical race feminism and figured worlds to examine school discipline." *Urban Education* 55, no. 10: 1419-1440.

Holman, Mirya R., and Monica C. Schneider. 2018. "Gender, race, and political ambition: how intersectionality and frames influence interest in political office." *Politics, Groups, and Identities* 6, no. 2:264-280.

Huff, C., & Tingley, D. 2015. "Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Research and Politics*, July-September, 1– 12.

Kosciw, J. G., Greytak, E. A., Giga, N. M., Villenas, C., & Danischewski, D. J. 2016. The 2015 national school climate survey: the experiences of lesbian, gay, bisexual, transgender, and queer youth in our nation's schools. Gay, Lesbian and Straight Education Network (GLSEN). 121 West 27th Street Suite 804, New York, NY 10001.

Ly, Laura and Eric Levenson. February 2, 2021. "Rochester police officers handcuff and pepper-spray a 9-year-old girl after call of 'family trouble.'" CNN, retrieved from, https://www.cnn.com/2021/02/01/us/rochester-police-pepper-spray-child/index.html?utm_content=2021-02-01T10%3A00%3A12&utm_source=twCNN&utm_term=link&utm_medium=social

Lemi, Danielle Casarez, and Nadia E. Brown. 2019. "Melanin and curls: evaluation of Black women candidates." *Journal of Race, Ethnicity and Politics* 4, no. 2 (2019): 259-296.

Morris, Edward W. 2007. ""Ladies" or "loudies"? Perceptions and experiences of Black girls in classrooms." *Youth & Society* 38, no. 4: 490-515.

Nelson, Libby and Dora Lind. 2015. The School to Prison Pipeline, Explained. Justice Policy Institute. Retrieved from, http://www.justicepolicy.org/news/8775

Ocen, Priscilla. 2015. E(Racing) Childhood: Examining the Racialized Construction of Childhood and Innocence in the Treatment of Sexually Exploited Minors. *UCLA Law Review* no.62 (6): 1588-1640.

Okonofua, Jason A., and Jennifer L. Eberhardt. 2015. Two strikes: Race and the disciplining of young students. *Psychological Science*, no.26 (5): 617-624.

Peffley, Mark and Jon Hurwitz. 2010. *Justice in America: The Separate Realities of Blacks and Whites*. New York: Cambridge University Press.

Philpot, Tasha S., and Hanes Walton Jr. 2007. "One of our own: Black female candidates and the voters who support them." *American Journal of Political Science* 51, no. 1: 49-62.

Ritchie, Andrea J. 2007.*Invisible no more: Police violence against Black women and women of color*. Beacon Press.

Roberts, Dorothy. 2009. *Shattered bonds: The color of child welfare*. Civitas Books.

Schram, Sanford F., Joe Soss, Richard C. Fording, and Linda Houser. 2009. Deciding to discipline: Race, choice, and punishment at the frontlines of welfare reform. *American sociological review*, *74*(3), 398-422.

Sigelman, Lee, and Susan Welch.1984. "Race, gender, and opinion toward black and female presidential candidates." *Public Opinion Quarterly* 48, no. 2: 467-475.

Simien, Evelyn M. 2007. "Doing intersectionality research: From conceptual issues to practical examples." *Politics & Gender* 3, no. 2: 264.

Smooth, Wendy. 2006. "Intersectionality in electoral politics: A mess worth making." *Politics & Gender* 2, no. 3: 400.

Townsend, Tiffany G., Torsten B. Neilands, Anita Jones Thomas, and Tiffany R. Jackson. "I'm no Jezebel; I am young, gifted, and Black: Identity, sexuality, and Black girls." *Psychology of Women Quarterly* 34, no. 3 (2010): 273-285.

U.S. Department of Education Office for Civil Rights. (2014, March). *US Civil Rights Data Collection Data Snapshot: School Discipline.* Civil Rights Data Collection.

Weisse, C. S., Sorum, P. C., & Dominguez, R. E. 2003. The influence of gender and race on physicians' pain management decisions. *The journal of Pain, 4*(9), 505-510.

Wun, Connie. 2016a. "Unaccounted foundations: Black girls, anti-Black racism, and punishment in schools." *Critical Sociology* 42, no. 4-5 2016: 737-750.

Wun, Connie. 2016b. "Unaccounted foundations: Black girls, anti-Black racism, and punishment in schools." *Critical Sociology*42, no. 4-5 (2016): 737-750.

Appendix

Survey background/implementation

Time period of data collection:
The survey (HIT) was launched on Amazon's Mechanical Turk from 10:00AM CST to 12:00PM CST on 3/8/2020. Note that during this time, both the Democratic presidential primaries and the coronavirus were occurring at this time, although the coronavirus had not yet been designated a global health pandemic.

Survey recruitment materials:
The HIT was advertised to workers on Mechanical Turk with language mirroring that of the consent information, and included the following information:
HIT Title: Answer a short survey about current social and political issues in the United States.
HIT description: Take a 15-minute survey about current social and political issues in the United States. This survey is part of a national study being conducted by researchers in x. The survey should take about 15 minutes to complete. In appreciation of your participation, we will give you $2.50.
Keywords: survey, questionnaire, public opinion, social issues, politics

Participation criteria:
Workers had to meet two criteria to be eligible to participate: a US IP address and a HIT approval rate of >90. In addition to this, participants had to indicate being at least 18 years of age or older in the survey in order to complete the survey.

Payment:
Participants were paid $2.50 for their participation. Each participant received a randomly generated code at the end of the survey. In order to receive payment, participants had to enter this code into the Mechanical Turk submission page so that their identity and completion could be verified.

Consent
Participants begin surveys by reading introductory text and consent information. Participants read the consent language and will be told that clicking they agree to participate signifies their agreement and understanding of the consent document. [full consent document below]

Deception
The following language was provided upon completion:
Thank you for completing this study. During the course of the study, you read a story about a student or an adult. While you may have been told the story you read was published in print and Internet media, it was not real. The story, and the individual(s) and location included in the story were fabricated for research purposes. We were interested in learning how people react to different kinds of scenarios involving punishment, so these manipulations were essential.

If due to the deception involved in this study, you wish to delete your responses, you may check the box

below and click to the next page to receive your code for payment. Even if you delete your responses, you will be given credit for participating.

If you do not wish to delete your responses, simply click to the next page to receive your code for payment.

If you have any questions about the study, you may contact us at x.

Online Consent Form

I am conducting a research study to understand people's attitudes about current social and political issues. If you agree to participate, you will be asked to answer a 15-minute survey. The study contains basic questions about you (e.g., demographics), as well as a few questions about your thoughts on current social and political issues including those experienced by children at school and adults in the child-care and justice system.

Your participation in this study does not involve any risk to you beyond that of everyday life. The possible benefits to you from this study include the opportunity to participate in policy-relevant research. Taking part in this study may help researchers to both better understand what people think about some of the country's most pressing issues.

Participation in this study will involve no cost to you. If you complete this study, you will be paid $2.50. This amount will be automatically placed in your account 2 days after completion of the survey. Please read this consent document. If you decline to participate, or exit the study prematurely, you will still be given credit for participating.

The survey is anonymous. The survey is being hosted by Qualtrics and involves a secure connection. Terms of service, including privacy information, can be found here: http://www.qualtrics.com/terms-of-service/.

The results of the research study may be published, but your name will not be used. Your participation in this research study is completely voluntary. You can skip questions in the survey that you chose not to answer and you can withdraw at any time by just exiting the survey.

If you have any questions about this study, you may contact X Questions about your rights as a research subject may be directed to the Institutional Review Board (IRB) Office of X University at (312) 503-9338.

If you want a copy of this consent for your records, you can print it from the screen. If you would like documentation linking you to this research study, please email your request to the Principal Investigator at X

If you wish to participate, please select the Accept button below to begin the survey.

If you do not wish to participate in this study, please select the Decline button, and your session will end.

Sample characteristics

Of 1,644 participants who began the survey, 135 explicitly forbade use of their data and 43 did not reach this point in the survey to indicate whether any data they may have provided could be used. These 178 participants are excluded from all analyses, yielding a final sample of 1,466 participants. In total, 1,466 participants completed the survey. All participants were randomly assigned to one of 16 experimental conditions, with sample sizes ranging from 84 to 96 participants. Scenarios varied.

Table A.1 below provides data on the basic demographics of the final sample. The sample is primarily white, and relatively highly educated, Democratic, and liberal. Nevertheless, the sample is more diverse than many samples commonly used in experimental survey research and consistent with research on MTurk samples (see e.g., Berinsky, Huber, and Lenz).

Table A.1: Sample Demographics

Total n	1,466
Age (mean)	41.4 years (median=39)
Female	49.1%
Male	50.1%
Transgender/Gender non-conforming	0.8%
Education	
Less than high school diploma	0.4%
High school diploma	9.9%
Some college	28.4%
College graduate	43.3%
Graduate or professional school	17.9%
Annual Household Income	
Less than $10,000	2.9%
$10,000-$39,999	28.2%
$40,000-$69,999	33.7%
$70,000-$99,999	19.2%
$100,000 and above	15.9%
Race and ethnicity	
White	77.8%
Black	8.9%
Hispanic/Latinx	4.6%
Asian	6.3%
Other race or ethnicity	2.3%
% Living w/school-age children	34.7%
Partisanship	
Democrat/Lean Democrat	55.7%

Republican/Lean Republican	32.9%
Pure Independent	11.4%
Ideology (1-7; 7=extremely conservative)	
Mean	3.6
Median	3 (Slightly liberal)
Mode	2 (Liberal)

Experimental analyses: All participants in dress code conditions

The remaining analyses are restricted to only those participants assigned to experimental conditions about student dress code.

Dress Code Scenario

Consider an instance where a student named _____ is wearing shorts to school. When _____ arrives to class, the teacher tells _____ that _____ shorts violate the school's dress code policy. The teacher tells _____ to leave class and go to the office. This also isn't the first time that _____ has worn clothes that violate the school's dress code policy. For example, _____ has previously worn tank tops that are against school rules.

Dependent Variables

DV1 (adult2): How much do students like ${e://Field/name} act older than their age?
DV2 (adult3): How much are students like ${e://Field/name} danger to others
DV3 (adult5): How are students like ${e://Field/name} experienced with sex
DV4 (suspension): How harsh is suspension is for students like

Table A.2 displays the demographics and statistics checking randomization across conditions about student dress codes. Chi^2 tests indicate mostly effective randomization among participants across these conditions, with no statistically significant relationships between experimental conditions, and any of the demographic variables. As the table shows, however, there are some imbalances across conditions, particularly gender and education.

Table A.2: Demographics and Randomization Checks of Student Dress Code Conditions-All participants

	Condition 1: Emily	Condition 2: Jake	Condition 3: Jamal	Condition 4: Keisha
Total n	91	88	84	88
Mean Age (Chi^2=51.066; p=0.59)	40.0	39.4	41.1	38.3
% Female (Chi^2=5.187; p=0.16)	36.3%	46.6%	52.4%	40.9%
% White (Chi^2=3.585; p=0.31)	72.5%	73.9%	79.8%	67.1%
% Some college or less (Chi^2=7.6124; p=0.57)	43.9%	38.6%	27.4%	36.4%
% Income less than $40,000 ($Chi^2$=8.467; p=0.49)	26.4%	31.0%	35.7%	25.0%
% Democrat (Chi^2=4.636; p=0.59)	53.9%	59.1%	59.5%	50.0%

Mean Ideology (1-7; 7=extremely conservative; Chi^2=2.209; p=0.89)	3.8	3.6	3.4	3.8

Randomization and Manipulation Checks- All participants (dress code conditions only)

To confirm whether the racial manipulation was effective, Figure 1 below displays how likely it was participants thought the name of the student they read about was the name of a Black student. Responses were rescaled to range from 0 to 1. The figure also provides bands indicating 95% confidence intervals, which enable judgments from the figures about which responses are statistically distinguishable from one another. As the figure shows, both of the names Jamal and Keisha were significantly more likely (M=0.76 se=0.03; M=0.71 se= 0.03) than both of the names Jake and Emily (M=0.33 se=0.02; M=0.29 se=0.02) to be thought of as belonging to Black students; these differences are robust to regression analyses controlling for demographic variables. In addition, as the figure shows, there are no differences in the perceived race of the student by gender.

Figure A.1: Mean Likelihood That Name is of a Black Student by Experimental Condition

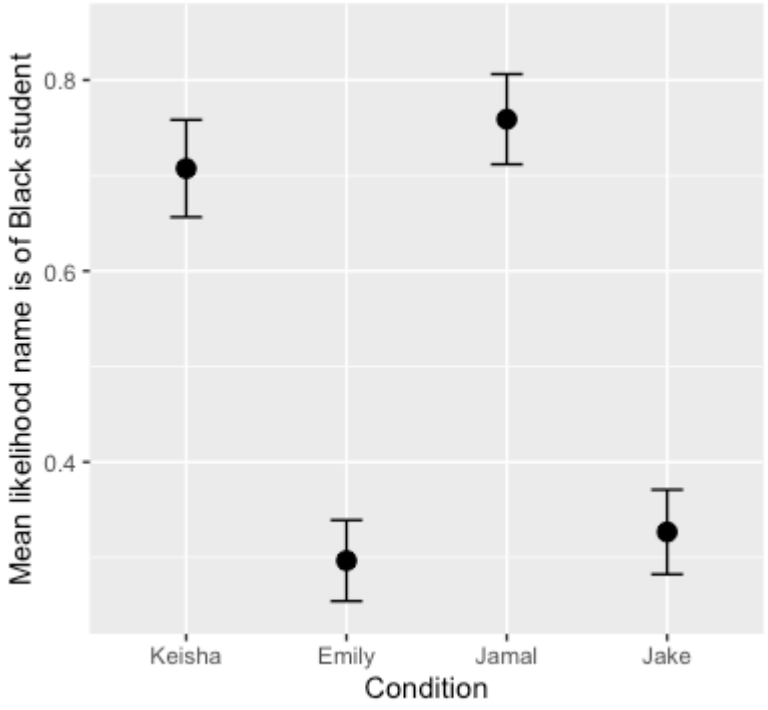
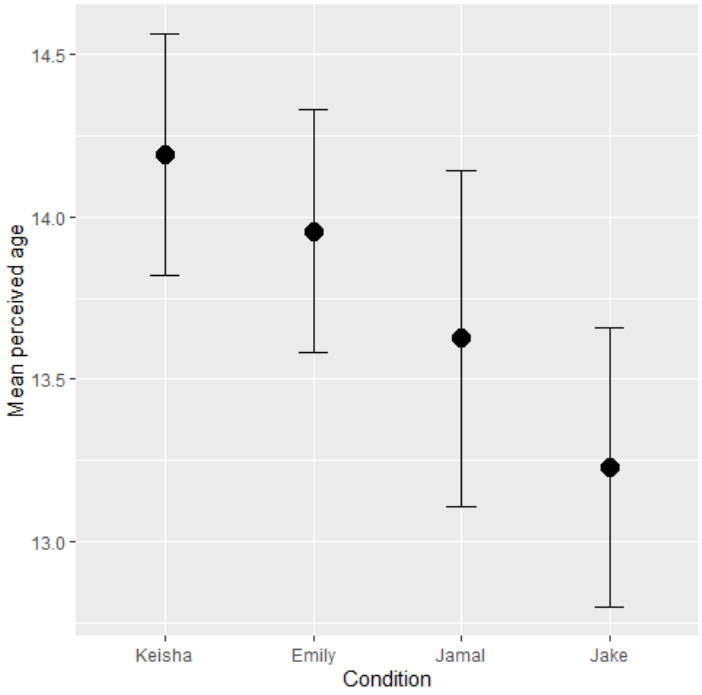
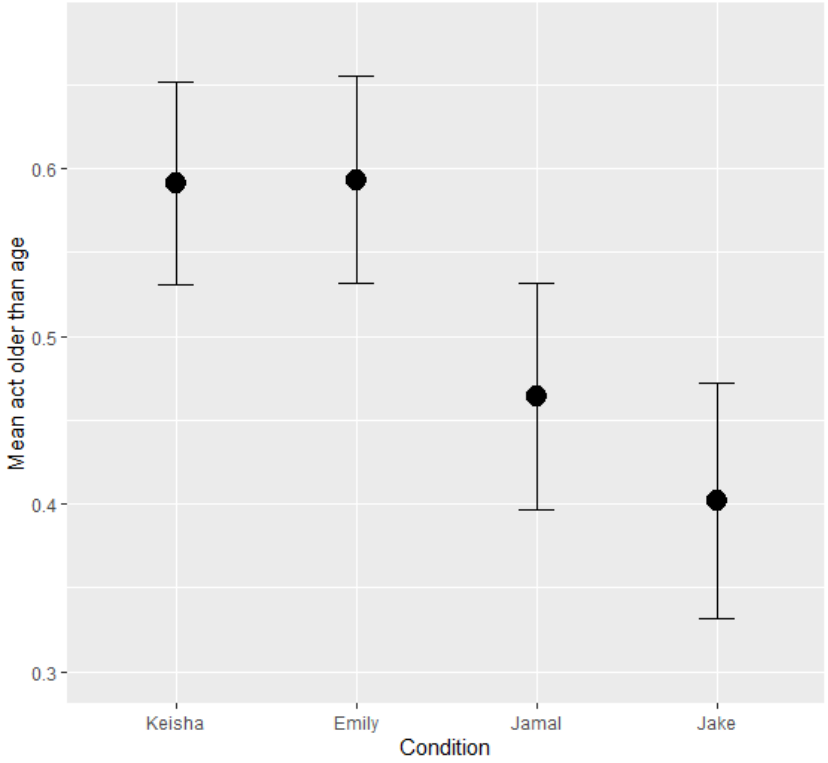


Figure A.2. Mean Perceived Student Age by Experimental Condition



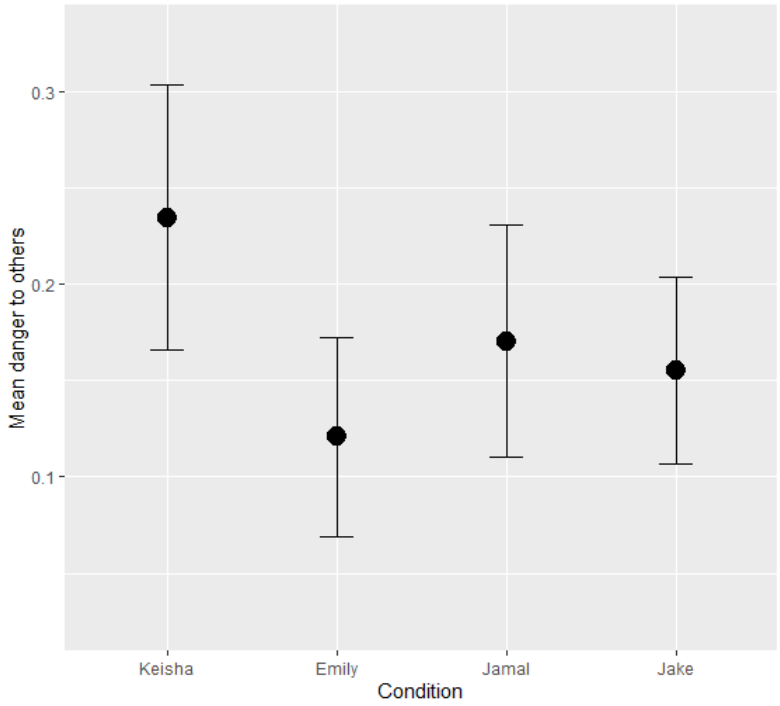
Condition	Mean	Std. Err.
Keisha	14.2	0.190
Emily	14.0	0.191
Jamal	13.6	0.264
Jake	13.2	0.220

Figure A.3. Mean Perception of Students' Acting Older than Age by Experimental Condition



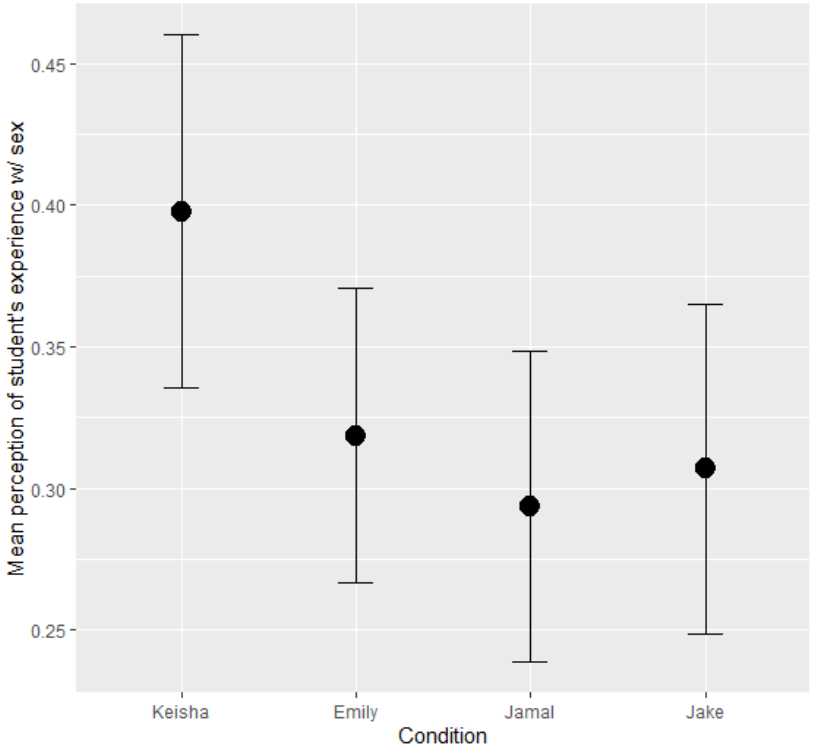
Condition	Mean	Std. Err.
Keisha	0.591	0.0308
Emily	0.593	0.0316
Jamal	0.464	0.0343
Jake	0.402	0.0358

Figure A.4. Mean Perception of Danger of Students to Others by Experimental Condition



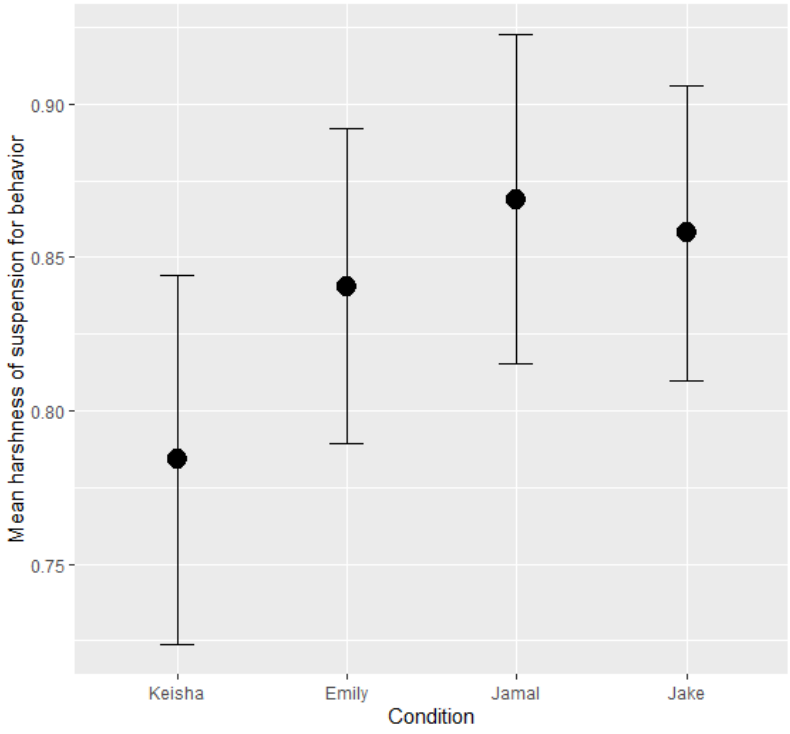
Condition	Mean	Std. Err.
Keisha	0.235	0.0350
Emily	0.121	0.0263
Jamal	0.171	0.0307
Jake	0.155	0.0247

Figure A.5. Mean Perception of Students' Experience with Sex by Experimental Condition



Condition	Mean	Std. Err.
Keisha	0.398	0.0318
Emily	0.319	0.0265
Jamal	0.294	0.0279
Jake	0.307	0.0296

Figure A.6. Mean Harshness of Suspension for Behavior by Experimental Condition



Condition	Mean	Std. Err.
Keisha	0.784	0.0321
Emily	0.841	0.0258
Jamal	0.869	0.0256
Jake	0.858	0.0242

Table A.3. Regression Analysis of Dress Code Scenario

	Dress Code Scenario							
	adult2	adult3	adult5	culp1	culp2	culp3	suspension	severity
Emily	0.008 (0.046)	-0.088** (0.040)	-0.063 (0.041)	-0.017 (0.040)	0.006 (0.031)	-0.034 (0.043)	0.058 (0.036)	-0.047 (0.037)
Jamal	-0.100** (0.047)	-0.016 (0.041)	-0.075* (0.042)	-0.004 (0.041)	-0.005 (0.032)	-0.065 (0.044)	0.058 (0.037)	-0.056 (0.038)
Jake	-0.195***(0.046)	-0.063 (0.040)	-0.080* (0.041)	-0.068* (0.040)	-0.015 (0.031)	-0.001 (0.043)	0.058 (0.036)	-0.043 (0.037)
Intercept	0.356*** (0.083)	0.125* (0.072)	0.285*** (0.074)	0.640*** (0.072)	0.639*** (0.057)	0.147* (0.078)	0.946*** (0.065)	0.179*** (0.067)

- Control variables: gender, race, education, age, income, ideology (coefficients and std. err. not reported here)
- Robust standard errors are in the parentheses (type = “HC3” following Hayes and Cai 2007).

Power Analysis

Given the (standardized) effect size, set significance level and sample size, how much power does our survey have (the probability of rejecting the null hypothesis when there is indeed a true effect – i.e., the probability of detecting a true effect?)

The conventional threshold is 0.8. I used pooled variance for power analysis since the ratio between standard deviations is close to 1 for almost all conditions ($S^2_{pooled} = \frac{(n-1)S_x^2+(m-1)S_y^2}{n+m-2}$). MTurk survey had sufficient power to detect medium-sized effects (Cohen's d, or the standardized effect size of approximately 0.5). At the 95% confidence level, the pilot survey has power of about 0.90 which is far above the conventional threshold of 0.8. The survey, however, did not have enough power to detect small-sized effects (Cohen's d of approximately 0.2); it has power of about 0.26 at the 95% confidence level, and the power of about 0.37 at the 90% confidence level – both of which are far below the threshold of 0.8. (Quick note: per https://www.statmethods.net/stats/power.html, I used variance as the denominator when computing Cohen's d. Using standard deviation as the denominator did not change the results much.)

Overall, the survey has enough power to detect effects for some variables, it was not the case for all dependent variables. The reason is that the standard deviations for these variables (i.e., culp1, culp2, culp3, severity) are larger than other variables – which led to smaller Cohen's d. Unlike other variables with smaller standard deviations, the questions where effects weren't detected had seven answer choices. While it allowed for more granular choices, precisely because respondents had more answer choices, the variance got larger. Underpowered analysis (shown in regression above) were not included in the paper.

Survey Background/Implementation

Sample Characteristics

Of 2,632 participants who began the survey,

- 5 were under the age of 17;
- 107 did not proceed to the survey, or they declined to proceed to the actual survey;
- 197 did not finish the survey; and
- 57 explicitly forbade use of their data.

These 366 responses are excluded from all analyses, yielding a final sample of 2,266 participants. Table 1 below provides data on the basic demographics of the final sample.

Table 1 below presents the basic demographics of the final sample. As with other online samples, our Prime Panel sample is less Republican, more highly educated (approximately 26% of the respondents said that they have post-baccalaureate degree), and earns higher income (about 30% of the respondents said that their household income in 2019 was $100,000 or above).

Table 1. Summary Statistics (unweighted)

Total n	2,266
Age (mean)	41.08 years (median: 38 years)
Female	60.5%*
Male	39.0%
Transgender/Gender non-conforming	0.35%
Education	
Less than high school diploma	2.3%
High school diploma	17.7%
Some college	27.3%
College graduate	27.2%
Graduate or professional school	25.5%
Annual Household Income	
Less than $10,000	7.5%
$10,000-$39,999	29.7%
$40,000-$69,999	19.1%
$70,000-$99,999	14.0%
$100,000 and above	29.5%
Race and ethnicity	
White	62.3%
Black	30.5%*
Hispanic/Latinx	3.2%

Asian	1.5%
Other race or ethnicity	2.3%
% Living with kids having online classes	51.4%
Partisanship	
Democrat/Lean Democrat	47.0%
Republican/Lean Republican	29.3%
Pure Independent	23.0%
Ideology (1-7; 7=extremely conservative)	
Mean	3.78
Median	4.00 (middle of the road)
Mode	4.00 (middle of the road)

** Note: we oversampled Black female respondents. This is not a nationally representative sample, and proper weight adjustment would be needed.

Experimental Analyses

Randomization and Manipulation Checks

We randomly assigned the 2,266 participants into one of the four conditions: Emily, Jake, Jamal, and Keisha. If the treatment assignment is truly random, we should see that covariates (e.g., respondent demographics) should be balanced across the four groups. Table 2 presents the summary of randomization check. The results from chi-square tests and Kruskal-Wallis test show randomization worked in general – overall, we do not see any relationship between the treatment group and the distribution of demographic variables.[1]

Table 2. Randomization Check (unweighted sample)

	Condition 1: Emily	Condition 2: Jake	Condition 3: Jamal	Condition 4: Keisha
Total n	570	570	565	561
Mean Age (Chi^2=70.415; p=0.43)	40.08 (15.69)	41.10 (14.99)	41.13 (16.14)	42.02 (15.64)
% Female (Chi^2=4.63; p=0.20)	58.88%	58.63%	60.71%	64.17%
% White (Chi^2=4.54; p=0.60)	59.75%	64.26%	63.54%	62.14%
% Some college or less (Chi^2=6.03; p=0.74)	46.32%	46.32%	47.61%	48.84%
% Income less than $40,000	36.8%	37.54%	36.41%	38.32%

[1] We do witness slight imbalances across conditions, particularly gender (% female) and race (% white).

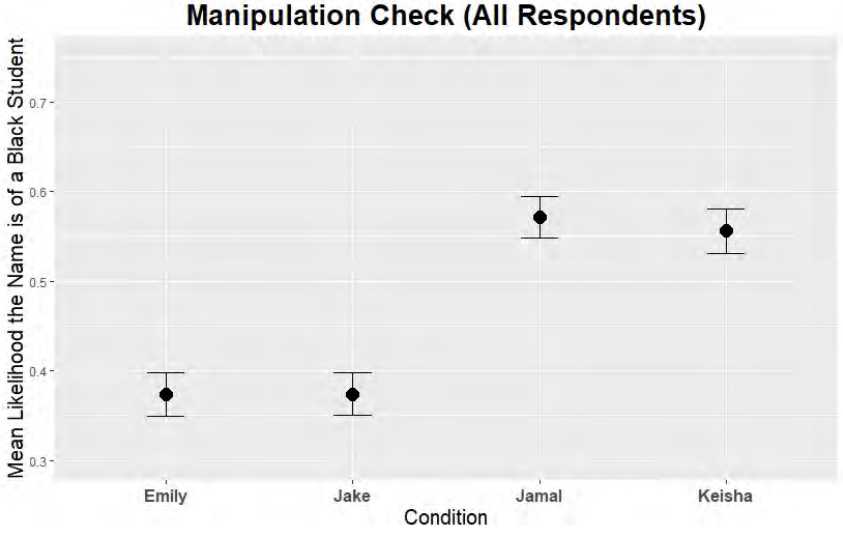
(Chi^2=23.459; p=0.8901)				
% Democrat (Chi^2=2.8849; p=0.4097)	50.09%	45.99%	47.60%	45.52%
Mean Ideology (1-7; 7=extremely conservative; Chi^2=5.8523; p=0.4399)	3.70 (1.84)	3.76 (1.74)	3.79 (1.85)	3.85 (1.87)

To confirm whether the racial manipulation was effective, we asked the respondents how likely it was that the name of the student they read about (Emily, Keisha, Jake, or Jamal) was the name of a Black student. Responses were rescaled to range from 0 to 1 for easier interpretation. Figure 1 also provides bands indicating 95% confidence intervals.

As the figure below shows, both the names Jamal and Keisha were significantly more likely ($\mu_{Jamal} = 0.57\ (0.01)$; $\mu_{Keisha} = 0.56\ (0.01)$ than both the names of Jake and Emily ($\mu_{Emily} = 0.38\ (0.01)$; $\mu_{Jake} = 0.37\ (0.01)$ to be thought of as names of Black students.[2]

However, the differences between Black student conditions and white student conditions are smaller compared to the MTurk survey done in March. This suggests that we may have more noncompliers, or fewer compliers – implying that it might be necessary to examine the responses from compliers and non-compliers separately, and see whether the overall treatment effect sizes are understated because of noncompliers. To see whether these "compliers" and "noncompliers" have different response patterns, we will examine these two groups separately.

Figure 1. Mean Likelihood that Name is of a Black Student by Experimental Condition



[2] Appendix: treatment worked better for Black participants than white participants. Black name cues (Jamal and Keisha) were clearer to Black respondents (i.e., Black respondents were more likely to "correctly" identify that Jamal and Keisha are names of Black students than white respondents).

We define compliers and noncompliers as follows:

- Compliers: those who thought Keisha and Jamal are names of Black students, and those who thought Jake and Emily are names of white students
- Noncompliers: those who thought Keisha and Jamal are names of white students, and those who thought Jake and Emily are names of Black students

Table 3 shows the summary statistics of compliers and noncompliers in our sample. It seems that noncompliers tend to be more educated, earn higher income, and Republican. We also see higher percentage of male and White among noncompliers.

Table 3. Compliers vs. Noncompliers: Summary Statistics (unweighted)

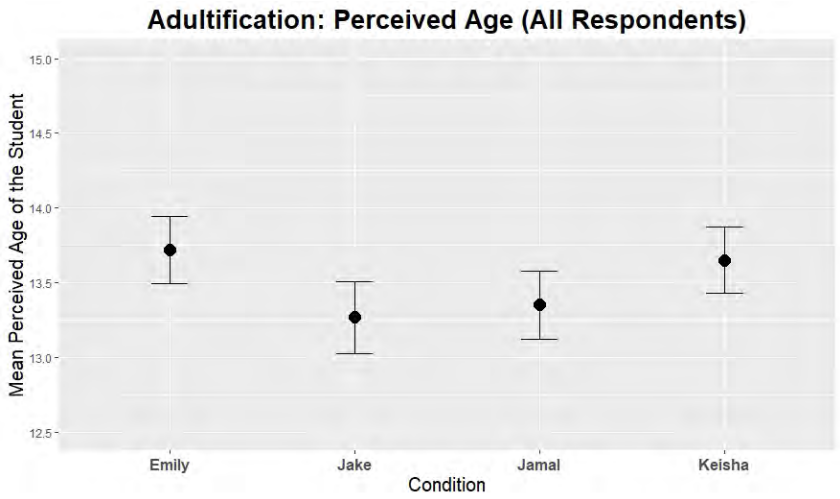
	Compliers	Noncompliers
Total n	1048	508
Age (mean)	41.46	40.94
Female	62.5%	53.9%
Male	37.5%	45.7%
Transgender/Gender non-conforming	0.2%	0.4%
Education		
Less than high school diploma	2.6%	1.2%
High school diploma	17.9%	14.8%
Some college	28.0%	22.4%
College graduate	28.1%	28.9%
Graduate or professional school	23.4%	32.7%
Annual Household Income		
Less than $10,000 (up til 1)	7.8%	6.3%
$10,000-$39,999 (2+3+4)	29.6%	24.8%
$40,000-$69,999 (5+6+7)	21.0%	17.5%
$70,000-$99,999 (8+9+10)	14.3%	13.6%
$100,000 and above (11+12_	27.0%	37.6%
Race and ethnicity		
White (7)	57.4%	69.5%
Black (3)	35.5%	23.4%
Hispanic/Latinx (4)	3.1%	4.3%
Asian (2)	1.3%	1.2%
Other race or ethnicity (1+5+10+6+8)	1.9%	1.4%
% Living with kids having online classes	50.2%	58.7%
Partisanship		

Democrat/Lean Democrat	50.5%	44.7%
Republican/Lean Republican	26.4%	36.6%
Pure Independent	22.1%	18.1%
Ideology (1-7; 7=extremely conservative)		
Mean	3.73	3.76
Median	4.0	4.0
Mode	4.0	4.0

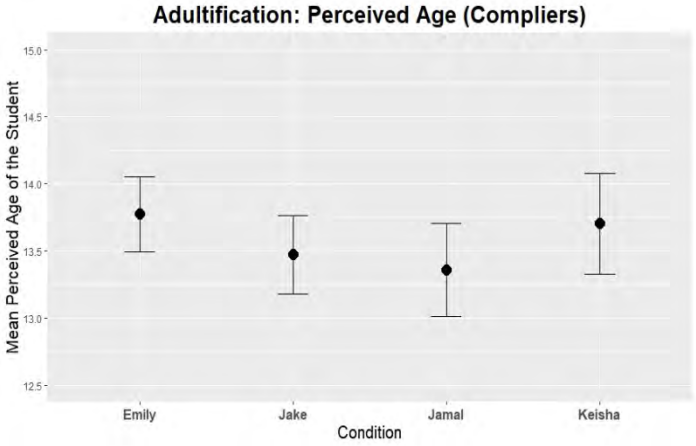
Effects of Experimental Treatments

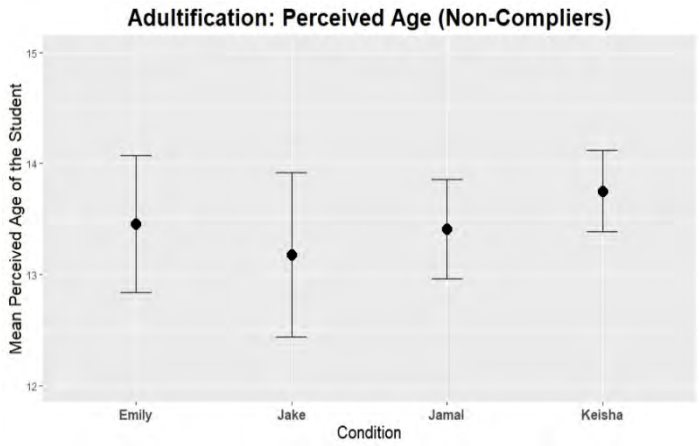
Adultification

Figure 2. Mean Perceived Student Age by Experimental Condition



While Emily is perceived to be slightly older than Jake and Jamal, the difference is not substantively large: 13.7 years old (Emily) vs. 13.3 years old (Jake). Compared with white respondents, Black participants thought Emily was slightly older (see Appendix). However, all students are thought to be roughly about the same age (between 13 and 14). We can see such response patterns among both compliers and noncompliers.



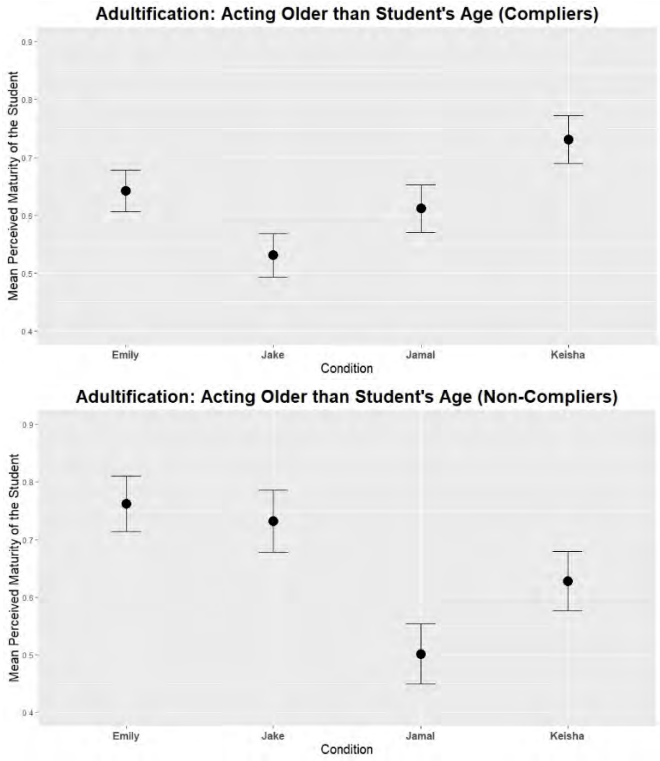


In sum, respondents did not perceive Keisha (or Emily) to be much older than her peers.

Figure 3. Mean Perception of Students' Acting Older than Age by Experimental Condition

As expected, we could see gendered adultification. For easier interpretation, we scaled the responses to range from 0 to 1. Higher values indicate that respondents thought the student in the given scenario acts older than their age.

Even though all four students were thought to be about the same age, respondents thought girls (Emily and Keisha) are more likely to act older than their age.

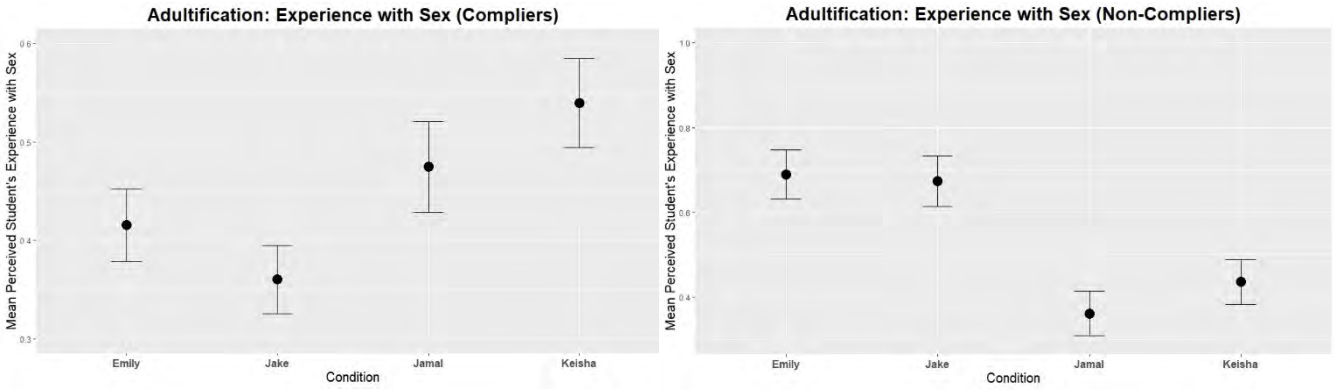


An interesting pattern emerges if we look at only compliers and noncompliers separately. Among compliers, Keisha is perceived to be the most mature, followed by Emily and Jamal. Jake is perceived to be least likely to act older than his age. Interestingly, the difference between Emily and Jamal is not that big (gendered *and racialized* perception of maturity?).

On the other hand, noncompliers thought Emily and Jake are more likely to act older than their age. Note that we defined noncompliers as the respondents who think that Emily and Jake are the names of Black students, and those who think that Jamal and Keisha are names of white students**. This means that noncompliers also think that Black students (in this case, Emily and Jake) act older than their age**.

Figure 4. Mean Perception of Experience with Sex

Again, as expected, people thought Keisha and Emily are more experienced with sex than Jake and Jamal in general.

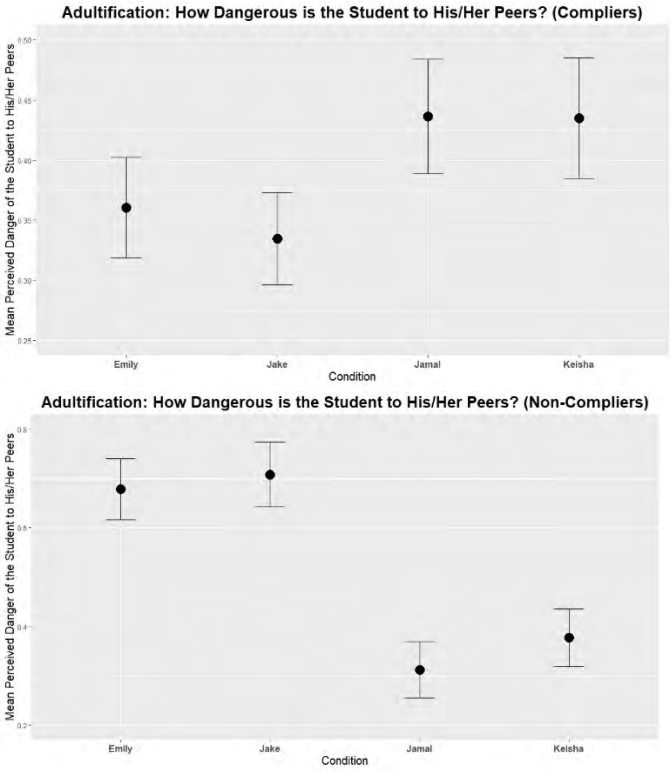


(note the different scales for noncompliers)

We can see a similar pattern as in Figure 3. Overall, we can see gendered adultification: girls (Emily and Keisha) were thought to be more experienced with sex than boys (Jake and Jamal) – even though they were thought to be about the same age.

When examining the responses of compliers and noncompliers separately, we can see more racialized responses. Both compliers and noncompliers though that Black kids are more experienced with sex than white kids (note that noncompliers thought Emily and Jake are names of Black students). We also see that Black girls (Keisha for compliers and Emily for noncompliers) are thought to be slightly more likely to have more experience with sex than their peers.

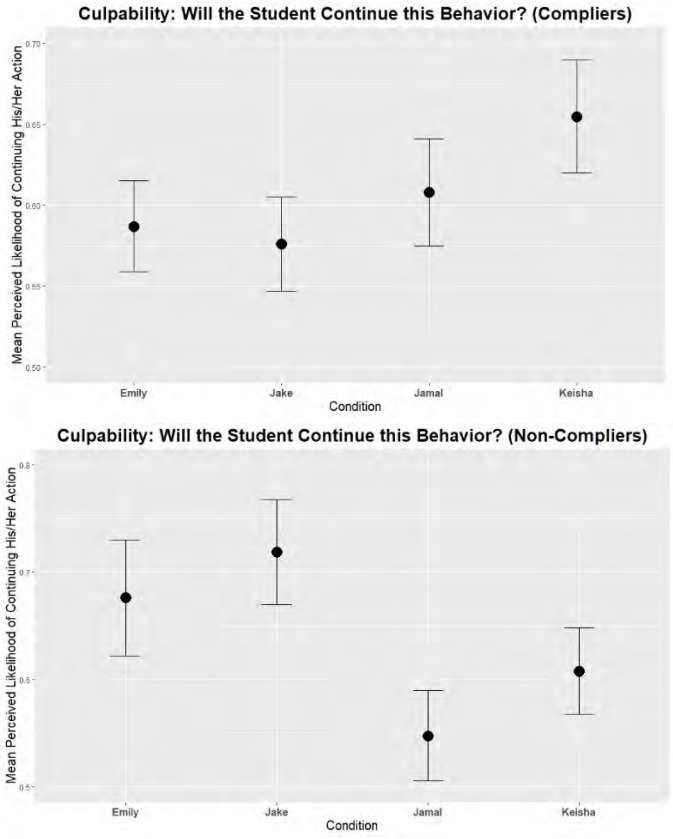
Figure 5. Mean Perception of Danger of Students to Others by Experimental Condition



We again see racialized responses. Compliers thought Black students are more dangerous to their peers than white students.

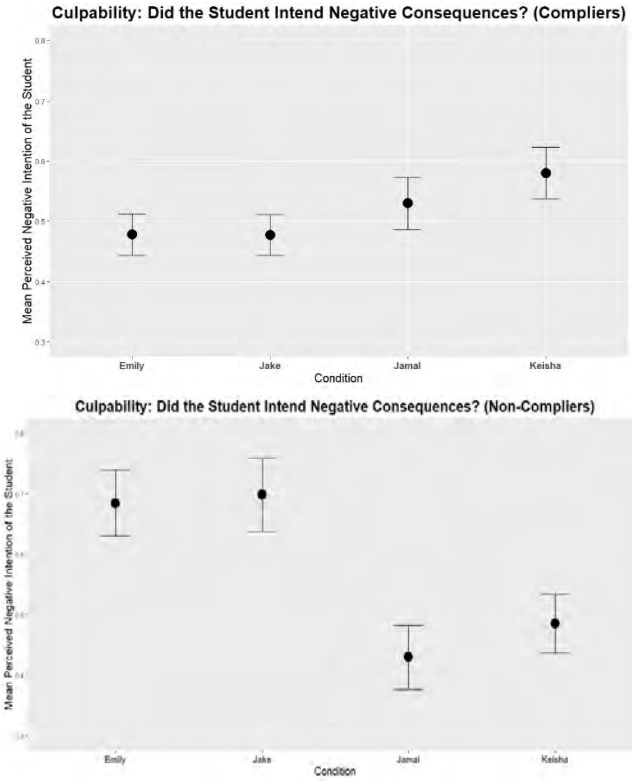
- Among compliers, Keisha was thought to be more mature and more experienced with sex than Jamal. Still, they think that Jamal is as dangerous to his peers as Keisha. (→ wonder whether respondents are thinking about adultification only, or whether they have other concerns in mind: e.g., criminalization of Black boys)

Figure 6-2. Culpability 2: Mean Perception of the Likelihood that Student will Continue His/Her Behavior



Among compliers, Keisha is thought to be the most likely to continue violating school dress code. In addition, we can see racialized responses among noncompliers as well: Emily and Jake (perceived to be Black students by noncompliers) are perceived to be more likely to continue their behaviors. However, Jake (i.e., Black male student) was thought to be more likely to continue violating the dress code than Emily (Black female student).

Figure 6-3. Culpability: Mean Perception of Student's Negative Intentions



- Gendered and racialized responses among compliers: Keisha is most likely to have intended negative consequences

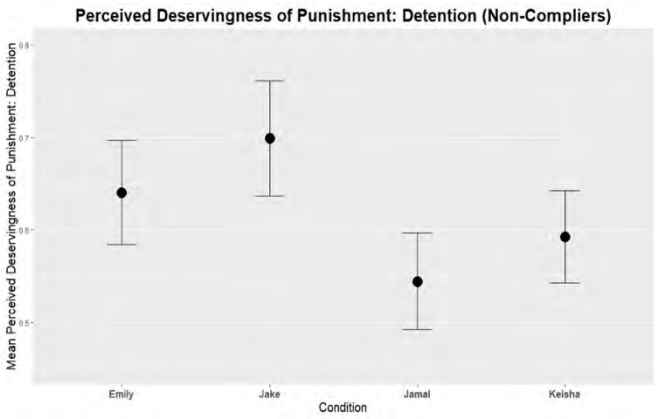
Compliers thought that (1) Keisha was more likely to continue her action (i.e., continue wearing "inappropriate clothes," and that (2) Keisha intended negative consequences. Though a little more racialized than gendered, we can see a similar pattern among noncompliers as well – Black students were thought to be more responsible for their actions than white students.

Figure 7. Mean Deservingness of Punishment (by Type) – among compliers

Figure 7-2. Detention

In general, we can see that respondents think Keisha "deserves" to get detention.

This response pattern is even clearer among compliers. Compliers agreed that Keisha should get detention as a punishment for violating school dress code. As with above, we can also see a racialized responses among noncompliers as well. In general, noncompliers were more supportive of punishing Black students with detention than white students. However, noncompliers were most likely to agree that a Black boy, Jake, should receive detention.



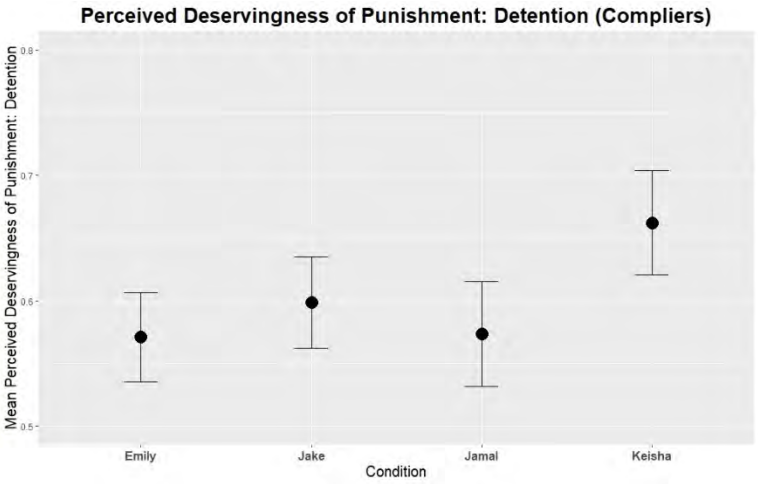
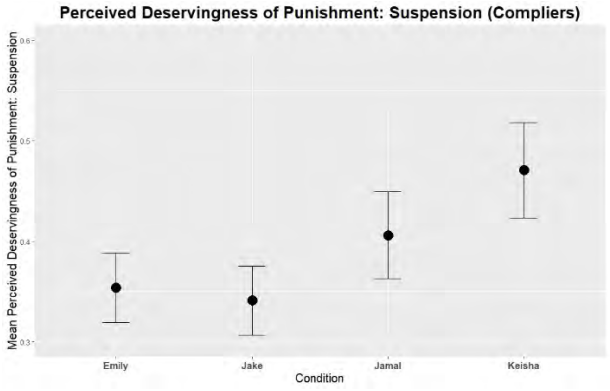


Figure 7-3. Suspension

We now look at how respondents think about the "deservingness" of suspension, a more severe form of punishment.

- In general, respondents agreed to punish Keisha with suspension more than other students. (Interpretation: "Figure 7-3 shows that suspension was thought to be the least harsh for Keisha than other students. In other words, suspension was seen as most appropriate when the student was Keisha.)
- The tendency is clearer among compliers and noncompliers.
    - Compliers: Keisha "deserves" to be suspended (followed by Jamal) – EVEN THOUGH their perceived responsibility of the student was about the same.
    - Noncompliers: again racialized – suspending a Black girl (Emily) was seen slightly less harsh than suspending a Black boy (Jake).
- Gendered response patterns: for both compliers and noncompliers -> suspending girls was seen (though slightly) less harsh than suspending boys of the same race.



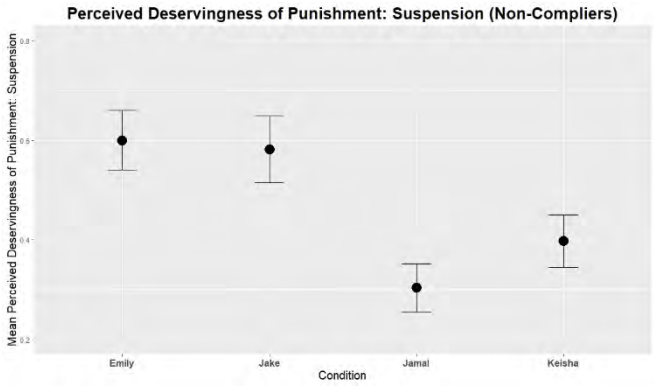
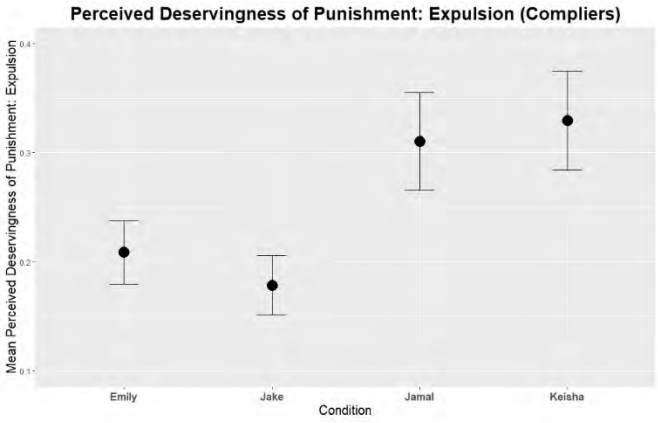


Figure 7-4. Expulsion

Lastly, we also asked whether respondents agree with expelling students for violating school dress code as their punishment. In general, respondents thought that students should not be expelled because of their actions – in other words, expulsion was too harsh for the students compared with other forms of punishment. However, people thought Emily's expulsion to be least harsh compared with other students.



Zooming in the responses of compliers and noncompliers tells us that respondents are more likely to agree with Black students being expelled as a punishment (slightly more so for Black girls).

In sum, we saw from the replication experiment that…

(1) While the respondents thought that the four students are about the same age, they thought girls acted older than boys of the same race. They also thought that Black students acted older than white students. In addition, respondents though that Black students are more experienced with sex than white students; compliers, in particular, thought that Keisha was the most experienced with sex as expected.
(2) Compliers also thought that Keisha was most responsible for her action, and the most likely to continue violating the school dress code.
(3) In a similar vein, more severe punishments (suspension and detention) were seen as being less harsh for Keisha than her peers – even though they all committed the same action.

All these findings support our hypotheses.

Stereotypes

Are Black girls thought to be more sexual? Are Black women seen as more sexually open than their white peers?

Table 1a. Stereotypes about Girls/Women (by Race)

	Black girls	Black women	White girls	White women
More sexual	4.113	4.145	4.594	4.434
More mature	4.455	----	4.067	----
More independent	4.924	4.922	4.202	4.521
Less respectful	3.827	3.624	3.986	4.038

Though slightly, Black girls are indeed seen as more independent and mature than white girls (adultification). This can lead to more punitive sentiments towards Black girls (Black girls are assumed to "know what they are doing," so they are more responsible than their "innocent" peers).

Table 1b. Stereotypes about Black and White Americans

	Black women	Black men	White women	White men
Sexually open	4.698	4.911	4.837	4.925
Violent	3.688	3.878	3.71	4.057

Contrary to the stereotype surrounding Black men (criminalization of Black men, violence of Black men), we can see that respondents rated white men to be more violent. Perhaps the racial uprisings in 2020 played a role – George Floyd and Breonna Taylor were killed by *white male* police officers. Daniel Holtzclaw, who raped multiple Black girls and women, was also half-white.

Table 2a and 2b show whether and how stereotypes surrounding Black girls (e.g., Black girls being more sexual or mature than their peers) affect people's punitive sentiments. We do see robust, significant, and positive coefficients for these stereotype variables – those who believe that Black girls are more sexual and mature than their peers were also more supportive of punishment, regardless of the experimental condition they were assigned. Note that *n* is around 550 – this is because we randomly assigned one of the four question sets (i.e., regarding Black women, Black girls, white women, and white girls) for survey length reasons.

Table 2a. Stereotypes Surrounding Black Girls and Punitive Sentiment

	Dependent variable:				
	severity	Change clothing	detention	suspension	expulsion
	(1)	(2)	(3)	(4)	(5)
Black girls are more sexual	0.059	0.061	0.090*	0.169***	0.141***
	(0.041)	(0.048)	(0.049)	(0.052)	(0.049)
age	-0.098*	-0.072	0.103	-0.184***	-0.193***
	(0.055)	(0.064)	(0.065)	(0.069)	(0.064)
income	0.023	-0.024	-0.007	0.031	-0.004
	(0.045)	(0.052)	(0.053)	(0.056)	(0.052)
education	0.148**	0.185**	0.218***	0.127	0.044
	(0.063)	(0.074)	(0.076)	(0.080)	(0.074)
Ideology (Lib-con)	0.070*	0.050	0.098**	0.098*	0.121**
	(0.041)	(0.047)	(0.049)	(0.051)	(0.048)
Party ID (Dem-Rep)	-0.029	0.049	-0.025	-0.022	-0.021
	(0.033)	(0.039)	(0.040)	(0.042)	(0.039)
Attend religion	0.087***	0.062*	0.009	0.113***	0.164***
	(0.032)	(0.038)	(0.039)	(0.041)	(0.038)
homeowner	0.014	0.028	-0.009	-0.048	0.016
	(0.026)	(0.030)	(0.031)	(0.033)	(0.030)
Employed (full)	0.068**	0.039	0.028	0.047	0.060*
	(0.027)	(0.031)	(0.032)	(0.034)	(0.031)
Disciplined at school	0.062***	0.026	0.056**	0.036	0.012
	(0.024)	(0.028)	(0.028)	(0.030)	(0.028)
Constant	0.242***	0.403***	0.314***	0.176***	0.078
	(0.043)	(0.051)	(0.052)	(0.055)	(0.051)
Observations	532	531	530	531	530
R^2	0.147	0.083	0.071	0.124	0.157
Adjusted R^2	0.131	0.065	0.054	0.107	0.141
Residual Std. Error	0.243 (df = 521)	0.283 (df = 520)	0.290 (df = 519)	0.307 (df = 520)	0.285 (df = 519)
F Statistic	8.992*** (df = 10; 521)	4.685*** (df = 10; 520)	3.993*** (df = 10; 519)	7.347*** (df = 10; 520)	9.672*** (df = 10; 519)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2b. Stereotypes Surrounding Black Girls and Punitive Sentiment

	Dependent variable:				
	severity	Change clothing	detention	suspension	expulsion
	(1)	(2)	(3)	(4)	(5)
Black girls are mature	0.112**	0.181***	0.190***	0.240***	0.193***
	(0.043)	(0.050)	(0.052)	(0.055)	(0.051)
age	-0.069	-0.042	0.142**	-0.126*	-0.140**
	(0.054)	(0.063)	(0.064)	(0.068)	(0.063)
income	0.028	-0.016	0.004	0.046	0.007
	(0.044)	(0.051)	(0.053)	(0.056)	(0.052)
education	0.132**	0.160**	0.190**	0.098	0.021
	(0.063)	(0.073)	(0.076)	(0.080)	(0.074)
Ideology (lib-con)	0.076*	0.057	0.108**	0.115**	0.138***
	(0.040)	(0.047)	(0.048)	(0.051)	(0.047)
Party ID (Dem-Rep)	-0.011	0.068*	-0.003	0.015	0.012
	(0.033)	(0.038)	(0.039)	(0.042)	(0.039)
Attend religion	0.077**	0.041	-0.011	0.094**	0.151***
	(0.032)	(0.037)	(0.038)	(0.041)	(0.038)
homeowner	0.014	0.032	-0.006	-0.042	0.018
	(0.026)	(0.030)	(0.031)	(0.033)	(0.030)
Employed (full))	0.071***	0.040	0.032	0.052	0.066**
	(0.027)	(0.031)	(0.032)	(0.034)	(0.031)
Disciplined at school	0.061**	0.020	0.051*	0.033	0.012
	(0.024)	(0.027)	(0.028)	(0.030)	(0.028)
Constant	0.199***	0.332***	0.245***	0.099*	0.012
	(0.047)	(0.054)	(0.056)	(0.059)	(0.055)
Observations	532	531	530	531	530
R^2	0.155	0.102	0.089	0.137	0.166
Adjusted R^2	0.139	0.085	0.072	0.120	0.150
Residual Std. Error	0.241 (df = 521)	0.280 (df = 520)	0.287 (df = 519)	0.304 (df = 520)	0.282 (df = 519)
F Statistic	9.579*** (df = 10; 521)	5.892*** (df = 10; 520)	5.077*** (df = 10; 519)	8.251*** (df = 10; 520)	10.349*** (df = 10; 519)
Note:					*p<0.1; **p<0.05; ***p<0.01