**Online Appendix**

**Table of Contents**

[**Sample Demographics for Study 1 and Study 2 1**](#_heading=h.50td52h1iqe8)

[**Survey 1 Details 2**](#_heading=h.iyxq597yjtmg)

[**Background and Implementation 2**](#_heading=h.t2sxowycf3rh)

[**Randomization Check 4**](#_heading=h.ouqbpbwrxh3c)

[**Manipulation Check 5**](#_heading=h.cujf6yd4xj4)

[**Survey 2 Details 8**](#_heading=h.7b4e77j9cn29)

[**Background and Implementation 8**](#_heading=h.wuv2xugrah34)

[**Randomization Check 8**](#_heading=h.43jbxpt0qaco)

[**Manipulation Check 9**](#_heading=h.kie3z594u1ye)

[**Compliers vs. Noncompliers 10**](#_heading=h.hc4wdn7e351o)

[**Numerical Values for Figure 2 (in paper) 15**](#_heading=h.4291acz6izkw)

[**Regression Coefficients for Figure 6 (in paper) 16**](#_heading=h.fug9hurf84jv)

[**Stereotypes and Real-World Policies and Scenarios 17**](#_heading=h.7f4ooyic2nfh)

[**Additional Measures: Factor-based Indicators 18**](#_heading=h.vddov87xz7ml)

# 

# Sample Demographics for Study 1 and Study 2

**Table A.1**

| **Demographics** | **Study 1** | **Study 2 (full)** | **Study 2 (compliers)** |
| --- | --- | --- | --- |
| Total n | 1,466 | 2,266 | 1,048 |
| Age (mean) | 41.4 years | 41.08 years | 41.46 years |
| Female | 49.1% | 60.5%\* | 62.5%\* |
| Male | 50.1% | 39.0% | 37.5% |
| Transgender/Gender non-conforming | 0.8% | 0.35% | 0.2% |
| **Education** | | | |
| Less than high school diploma | 0.4% | 2.3% | 2.6% |
| High school diploma | 9.9% | 17.7% | 17.9% |
| Some college | 28.4% | 27.3% | 28.0% |
| College graduate | 43.3% | 27.2% | 28.1% |
| Graduate or professional school | 17.9% | 25.5% | 23.4% |
| **Annual Household Income** | | | |
| Less than $10,000 | 2.9% | 7.5% | 7.8% |
| $10,000-$39,999 | 28.2% | 29.7% | 29.6% |
| $40,000-$69,999 | 33.7% | 19.1% | 21.0% |
| $70,000-$99,999 | 19.2% | 14.0% | 14.3% |
| $100,000 and above | 15.9% | 29.5% | 27.0% |
| **Race and ethnicity** | | | |
| White | 77.8% | 62.3% | 57.4% |
| Black | 8.9% | 30.5%\* | 35.5%\* |
| Hispanic/Latinx | 4.6% | 3.2% | 3.1% |
| Asian | 6.3% | 1.5% | 1.3% |
| Other race or ethnicity | 2.3% | 2.3% | 1.9% |
| **Partisanship** | | | |
| Democrat/Lean Democrat | 55.7% | 47.0% | 50.5% |
| Republican/Lean Republican | 32.9% | 29.3% | 26.4% |
| Pure Independent | 11.4% | 23.0% | 22.1% |
| **Ideology (1-7; 7=extremely conservative)** | | | |
| Mean | 3.6 | 3.78 | 3.73 |
| Median | 3.0 | 4.00 | 4.0 |
| Mode | 2.0 | 4.00 | 4.0 |

\* We oversampled Black female respondents.

# 

# 

# Survey 1 Details

## Background and 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, 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 who completed the survey. All participants were randomly assigned to one of 16 experimental conditions, with sample sizes ranging from 84 to 96 participants.

## Randomization Check

Table A.2 displays the demographics and statistics checking randomization across conditions about student dress codes. Chi2 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  (Chi2=51.066; p=0.59) | 40.0 | 39.4 | 41.1 | 38.3 |
| % Female  (Chi2=5.187; p=0.16) | 36.3% | 46.6% | 52.4% | 40.9% |
| % White  (Chi2=3.585; p=0.31) | 72.5% | 73.9% | 79.8% | 67.1% |
| % Some college or less  (Chi2=7.6124; p=0.57) | 43.9% | 38.6% | 27.4% | 36.4% |
| % Income less than $40,000  (Chi2=8.467; p=0.49) | 26.4% | 31.0% | 35.7% | 25.0% |
| % Democrat  (Chi2=4.636; p=0.59) | 53.9% | 59.1% | 59.5% | 50.0% |
| Mean Ideology  (1–7; 7=extremely conservative;  Chi2=2.209; p=0.89) | 3.8 | 3.6 | 3.4 | 3.8 |

# 

## Manipulation Check

After the completion of the survey, a manipulation check was conducted in order to determine the effectiveness of using racialized names. Figure A.1 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 95% confidence intervals to determine which responses are statistically distinguishable from one another more easily. Figure A.1 illustrates that respondents were significantly more likely to attribute the names Jamal and Keisha (M=0.76 se=0.03; M=0.71 se= 0.03), as opposed to Jake and Emily (M=0.33 se=0.02; M=0.29 se=0.02), to Black students. The difference between respondents’ selection of the Black and White names is about 40 percent, suggesting that the manipulation was effective. These differences are robust across regression analyses controlling for demographic variables.

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

**Chart

Description automatically generated**

**Table A.3.** Manipulation Check: Regression Analysis

| Dependent variable: Likelihood that the given name is of a Black student | |
| --- | --- |
|  |  |
| Condition: Emily | -0.411\*\*\* (0.033) |
| Condition: Jamal | 0.056 (0.034) |
| Condition: Jake | -0.377\*\*\* (0.033) |
| Woman | -0.009 (0.024) |
| Transgender or Gender Nonconforming | 0.119 (0.133) |
| Education | -0.005 (0.014) |
| Other race/ethnicity | -0.017 (0.102) |
| Asian | 0.069 (0.047) |
| Black/African-American | 0.054 (0.041) |
| Hispanic/Latino | -0.057 (0.044) |
| American Indian/Alaska Native | 0.142 (0.128) |
| Age | 0.001 (0.001) |
| Income | -0.007 (0.004) |
| Ideology | -0.002 (0.041) |
| Constant | 0.740\*\*\* (0.072) |
| Observations | 350 |
| R2 | 0.507 |
| Adjusted R2 | 0.486 |
| Residual Std. Error | 0.218 (df = 335) |
| F Statistic | 24.598\*\*\* (df = 14; 335) |

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

**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 ). 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 percent 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 percent confidence level, and the power of about 0.37 at the 90 percent confidence level, both of which are far below the threshold of 0.8. (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, but this 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 these allowed for more granularity, precisely because respondents had more answer choices, the variance got larger. Items of underpowered analysis (shown in regression above) were not included in the paper.

# Survey 2 Details

## Background and 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 A.5 below provides data on the basic demographics of the final sample. As with other online samples, our Prime Panel sample is less Republican, more highly educated (approximately 26 percent of the respondents said that they have post-baccalaureate degree), and higher income (about 30% of the respondents said that their household income in 2019 was $100,000 or above) than the average population.

**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 A.4 presents the summary of the 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 (we do witness slight imbalances across conditions, particular gender [% female] and race [% white]).

## Randomization Check

**Table A.4.** Randomization Check (unweighted sample)

|  | **Condition 1: Emily** | **Condition 2: Jake** | **Condition 3: Jamal** | **Condition 4: Keisha** |
| --- | --- | --- | --- | --- |
| Total n | 570 | 570 | 565 | 561 |
| Mean Age  (Chi2=70.415; p=0.43) | 40.08 (15.69) | 41.10 (14.99) | 41.13 (16.14) | 42.02 (15.64) |
| % Female  (Chi2=4.63; p=0.20) | 58.88% | 58.63% | 60.71% | 64.17% |
| % White  (Chi2=4.54; p=0.60) | 59.75% | 64.26% | 63.54% | 62.14% |
| % Some college or less  (Chi2=6.03; p=0.74) | 46.32% | 46.32% | 47.61% | 48.84% |
| % Income less than $40,000  (Chi2=23.459; p=0.8901) | 36.8% | 37.54% | 36.41% | 38.32% |
| % Democrat  (Chi2=2.8849; p=0.4097) | 50.09% | 45.99% | 47.60% | 45.52% |
| Mean Ideology  (1–7; 7=extremely conservative;  Chi2=5.8523; p=0.4399) | 3.70 (1.84) | 3.76 (1.74) | 3.79 (1.85) | 3.85 (1.87) |

## Manipulation Check

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 A.2 also provides bands indicating 95 percent confidence intervals.

As figure A.2 shows, both the names Jamal and Keisha were significantly more likely (; than the names of Jake and Emily (; to be thought of as names of Black students.[[1]](#footnote-0)

However, the differences between Black student conditions and white student conditions are smaller compared to Study 1 (the MTurk survey done in March 2020). This suggests that we may have more noncompliers, or fewer compliers. In the body of the paper, therefore, we have dropped the noncompliers and have focused our analyses on the results for the compilers – that is, on those respondents for which the manipulation was effective.

For the sake of due diligence, we conducted further analyses to try to better understand what may have been going on with the noncomplier group in Study 2. In what follows, we examine the responses from compliers and noncompliers separately to see whether the overall treatment effect sizes are understated because of noncompliers. Here, we examine these two groups separately to see whether these “compliers” and “noncompliers” have different response patterns.

**Figure A.2.** Mean Likelihood that Name is of a Black Student by Experimental Condition

Chart, box and whisker chart

Description automatically generated

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 A.6 shows the summary statistics of compliers and noncompliers in our sample. It seems that noncompliers tend to be more educated, earn higher income, and lean Republican. We also see higher percentage of male and white respondents among noncompliers.

**Table A.5.** Compliers vs. Noncompliers: Summary Statistics (unweighted)

|  | **Full sample** | **Compliers** | **Noncompliers** |
| --- | --- | --- | --- |
| Total n | 2,266 | 1048 | 508 |
| Age (mean) | 41.08 | 41.46 | 40.94 |
| Female | 60.5%\* | 62.5%\* | 53.9%\* |
| Male | 39.0% | 37.5% | 45.7% |
| Transgender/Gender non-conforming | 0.35% | 0.2% | 0.4% |
| **Education** | | | |
| Less than high school diploma | 2.3% | 2.6% | 1.2% |
| High school diploma | 17.7% | 17.9% | 14.8% |
| Some college | 27.3% | 28.0% | 22.4% |
| College graduate | 27.2% | 28.1% | 28.9% |
| Graduate or professional school | 25.5% | 23.4% | 32.7% |
| **Annual Household Income** | | | |
| Less than $10,000 (up til 1) | 7.5% | 7.8% | 6.3% |
| $10,000–$39,999 (2+3+4) | 29.7% | 29.6% | 24.8% |
| $40,000–$69,999 (5+6+7) | 19.1% | 21.0% | 17.5% |
| $70,000–$99,999 (8+9+10) | 14.0% | 14.3% | 13.6% |
| $100,000 and above (11+12) | 29.5% | 27.0% | 37.6% |
| **Race and ethnicity** | | | |
| White (7) | 62.3% | 57.4% | 69.5% |
| Black (3) | 30.5%\* | 35.5%\* | 23.4%\* |
| Hispanic/Latinx (4) | 3.2% | 3.1% | 4.3% |
| Asian (2) | 1.5% | 1.3% | 1.2% |
| Other race or ethnicity (1+5+10+6+8) | 2.3% | 1.9% | 1.4% |
| % Living with kids having online classes | 51.4% | 50.2% | 58.7% |
| **Partisanship** | | | |
| Democrat/Lean Democrat | 47.0% | 50.5% | 44.7% |
| Republican/Lean Republican | 29.3% | 26.4% | 36.6% |
| Pure Independent | 23.0% | 22.1% | 18.1% |
| **Ideology (1–7; 7=extremely conservative)** | | | |
| Mean | 3.78 | 3.73 | 3.76 |
| Median | 4.00 | 4.0 | 4.0 |
| Mode | 4.00 | 4.0 | 4.0 |

\* We oversampled Black female respondents.

## 

## 

## Compliers vs. Noncompliers

# 

**Effects of Experimental Treatment on Adultification and Punishment Variables:**

Perceptions of age are shown in figure A.2. 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.

**Figure A.3.** Mean Perception of Students’ Acting Older than Age by Experimental ConditionGraphical user interface, chart

Description automatically generatedGraphical user interface, chart, box and whisker chart

Description automatically generated

An interesting pattern emerges if we look at 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. 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**.

Regarding perceptions of the students’ experience with sex, Figure A.4 shows a similar pattern as in Figure A.3. Overall, we can see gendered adultification: girls (Emily and Keisha) were thought to be more experienced with sex than were boys (Jake and Jamal), even though they were thought to be about the same age.

**Figure A.3.** Mean Perception of Experience with Sex Chart, box and whisker chart

Description automatically generatedChart

Description automatically generated

When examining responses of compliers and noncompliers separately, we can see more racialized responses. Both compliers and noncompliers thought that Black kids are more experienced with sex than are 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.

For the question of perceived danger (Figure A.5), 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 is Keisha (e.g., criminalization of Black boys).

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

Chart, box and whisker chart

Description automatically generatedChart

Description automatically generated

Figures A.6-1, A.6-2, and A.6-3 show the mean deservingness of punishment (by type). In general (for compliers and noncompliers), 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 (figure A.6-1). As with above, we can also see 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 A.6-1.** Detention

Chart, box and whisker chart

Description automatically generatedChart, box and whisker chart

Description automatically generated

We now look at what respondents think about the “deservingness” of suspension, a more severe form of punishment (Figure A.6-2). In general, respondents agreed to punish Keisha with suspension more than other students. (Interpretation: “Figure A.6-2 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 report that Keisha “deserves” to be suspended (followed by Jamal), even though their perceived responsibility of the student was about the same. Noncompliers responses are again racialized; suspending a Black girl (Emily) was seen as slightly less harsh than suspending a Black boy (Jake).

Gendered response patterns: for both compliers and noncompliers, suspending girls was seen (though slightly) as less harsh than suspending boys of the same race.

**Figure A.6-2.** Suspension

Chart, box and whisker chart

Description automatically generatedChart, box and whisker chart

Description automatically generated

Lastly, we also asked whether respondents agree with expelling students for violating school dress code (Figure A.6.3). In general, respondents thought that students should not be expelled because of their actions; that is, expulsion was too harsh for the students compared with other forms of punishment. However, people thought Keisha’s expulsion to be less harsh compared with the other students.

**Figure A.6-3.** Expulsion

Chart, box and whisker chart

Description automatically generated

Zooming in on 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 thought 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 for her peers, even though they all committed the same action.

All these findings support the hypotheses.

# 

# Numerical Values for Figure 2 (in paper)

# 

Table A.6: Numerical values for Figure 2.

# 

**Study 1**

|  | ***Treatment*** | **Emily** | **Jake** | **Jamal** | **Keisha** |
| --- | --- | --- | --- | --- | --- |
| **Acts Older** | *Mean* | 0.59 | 0.40 | 0.46 | 0.59 |
|  | *SD* | 0.30 | 0.34 | 0.31 | 0.29 |
|  | *N* | 91 | 88 | 84 | 88 |
| **Danger** | *Mean* | 0.12 | 0.16 | 0.17 | 0.23 |
|  | *SD* | 0.25 | 0.23 | 0.28 | 0.33 |
|  | *N* | 91 | 88 | 84 | 88 |
| **Experienced w/Sex** | *Mean* | 0.32 | 0.31 | 0.29 | 0.40 |
|  | *SD* | 0.25 | 0.28 | 0.26 | 0.30 |
|  | *N* | 91 | 88 | 84 | 88 |

**Study 2**

|  | ***Treatment*** | **Emily** | **Jake** | **Jamal** | **Keisha** |
| --- | --- | --- | --- | --- | --- |
| **Acts Older** | *Mean* | 0.64 | 0.53 | 0.61 | 0.73 |
|  | *SD* | 0.31 | 0.32 | 0.32 | 0.32 |
|  | *N* | 295 | 284 | 242 | 223 |
| **Danger** | *Mean* | 0.36 | 0.33 | 0.44 | 0.43 |
|  | *SD* | 0.36 | 0.33 | 0.38 | 0.38 |
|  | *N* | 294 | 286 | 243 | 223 |
| **Experience w Sex** | *Mean* | 0.42 | 0.36 | 0.47 | 0.54 |
|  | *SD* | 0.32 | 0.30 | 0.36 | 0.34 |
|  | *N* | 296 | 286 | 243 | 223 |

# 

# Regression Coefficients for Figure 6 (in paper)

**Table A.7.** Regression Coefficients from Figure 6: Stereotype Predictors of Punitive Attitudes

OLS Regressions from Figure 6

DV=Punitive attitudes index

|  | (1) | (2) | (3) | (4) |
| --- | --- | --- | --- | --- |
| VARIABLES | Limited | Mediation | Limited | Mediation |
|  |  |  |  |  |
| Perceived Age |  | 0.0108 |  | 0.01 |
|  |  | (0.0103) |  | (0.010) |
| Acts older than age |  | 0.109\*\*\* |  | 0.11\*\*\* |
|  |  | (0.0299) |  | (0.030) |
| Danger to others |  | 0.367\*\*\* |  | 0.37\*\*\* |
|  |  | (0.0282) |  | (0.028) |
| Experienced with Sex |  | 0.181\*\*\* |  | 0.18\*\*\* |
|  |  | (0.0327) |  | (0.033) |
| t4 = 1, Emily | -0.0210 | -0.120\* | -0.02 | -0.12\* |
|  | (0.0829) | (0.0688) | (0.083) | (0.069) |
| t4 = 3, Jamal | 0.144\* | -0.0506 | 0.14\* | -0.05 |
|  | (0.0872) | (0.0722) | (0.087) | (0.072) |
| t4 = 4, Keisha | 0.240\*\*\* | -0.0364 | 0.24\*\*\* | -0.04 |
|  | (0.0892) | (0.0752) | (0.089) | (0.075) |
| Constant | -0.0790 | -1.621\*\*\* | -0.08 | -1.62\*\*\* |
|  | (0.0592) | (0.149) | (0.059) | (0.149) |
|  |  |  |  |  |
| Observations | 1,042 | 1,033 | 1,042 | 1,033 |
| R-squared | 0.011 | 0.345 | 0.011 | 0.345 |

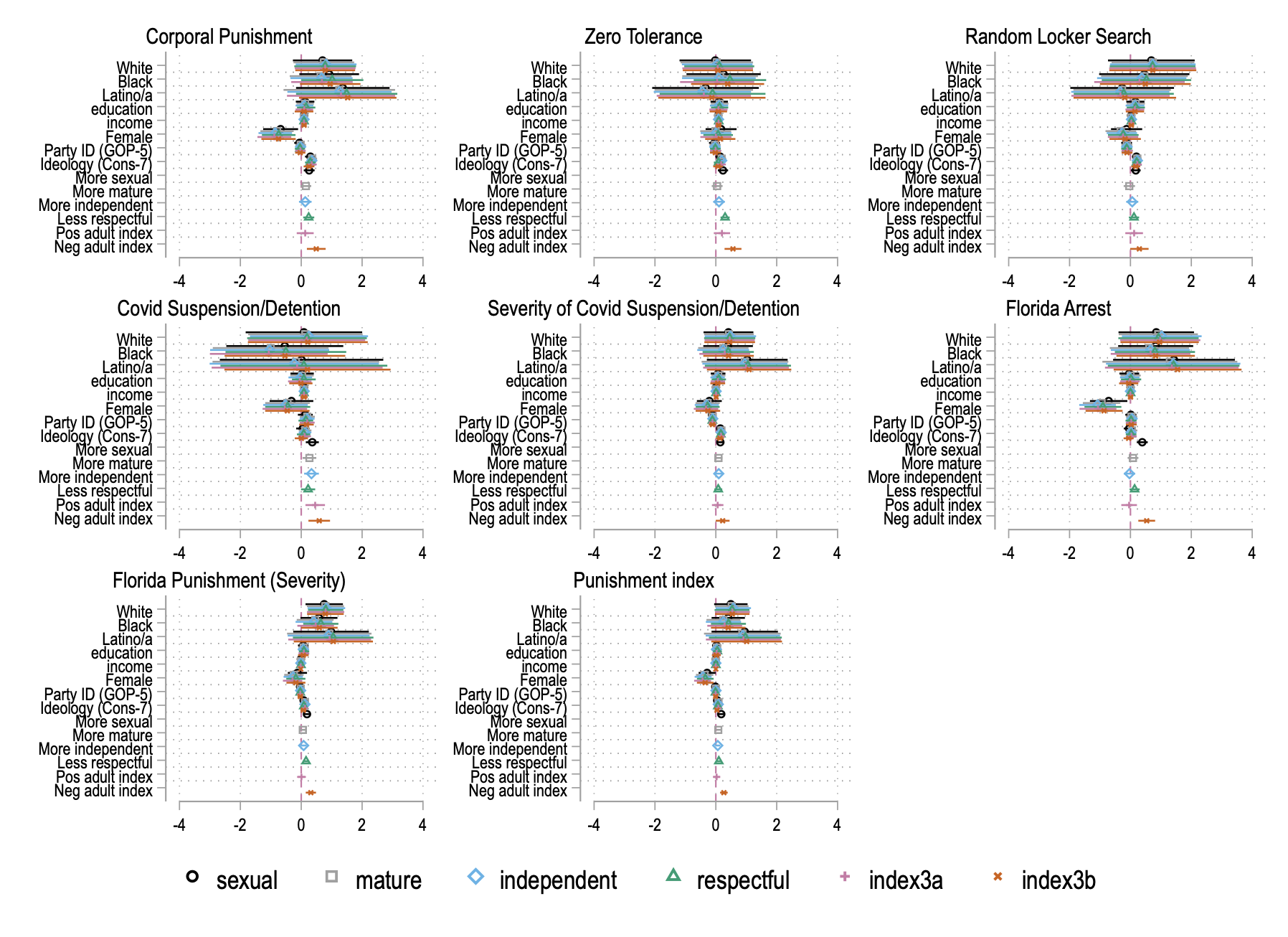
Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

# Stereotypes and Real-World Policies and Scenarios

Figure A.7 graphically presents the coefficients and 95% confidence intervals for multivariate analyses building on the bivariate analyses in Figure 4 in the body of the text. Each subfigure presents estimates from an OLS regression of the dependent variable listed at the top on one of the stereotypical views along with the demographic covariates listed at left (only one stereotypical view was included in each model, so each subfigure includes estimates for six models, as indicated by the key at the bottom. In nearly every instance, the relationship between the negative stereotypical view and the view on punishment remains positive and significant. *All* relationships remain positive and significant if only demographic covariates are included (i.e, if ideology and partisanship are excluded; the well-known association between racial attitudes, ideology, and partisanship suggest that these measures may be post-treatment to stereotypes, but we include them here to present the most conservative presentation of our key relationships. More detailed figures and numerical presentations are available upon request to the authors or via the replication data that accompanies this paper.

**Figure A.7 Coefficient estimates, OLS regression of support for punishment (top of each subfigure) on explanatory covariates listed at left.**



# Additional Measures: Factor-based Indicators

For three concepts included in the paper’s analysis – punitive attitudes in the vignette (Fig 3), stereotypical attitudes about Black girls (Fig 4), and general support for school punishment (Figure 4) – we constructed index estimates based on factor analysis. The following section provides the Stata input and output that generated those measures, including Eigen values and factor loadings.

**.\*Index of support for punishment of student in vignette**

. factor severity changeclothing detention suspension expulsion, pcf

. rotate, varimax

Factor analysis/correlation Number of obs = 1,042

Method: principal-component factors Retained factors = 1

Rotation: orthogonal varimax (Kaiser off) Number of params = 5

--------------------------------------------------------------------------

Factor | Variance Difference Proportion Cumulative

-------------+------------------------------------------------------------

Factor1 | 1.98070 . 0.3961 0.3961

--------------------------------------------------------------------------

LR test: independent vs. saturated: chi2(10) = 570.26 Prob>chi2 = 0.0000

Rotated factor loadings (pattern matrix) and unique variances

---------------------------------------

Variable | Factor1 | Uniqueness

-------------+----------+--------------

severity | 0.7740 | 0.4009

changeclot~g | 0.3291 | 0.8917

detention | 0.5382 | 0.7103

suspension | 0.6796 | 0.5381

expulsion | 0.7223 | 0.4783

---------------------------------------

Factor rotation matrix

-----------------------

| Factor1

-------------+---------

Factor1 | 1.0000

-----------------------

**.\*Index of adultification stereotypes about Black girls**

. factor blackgirl\_sexual blackgirl\_mature blackgirl\_independent blackgirl\_respectful, pcf

rotate, varimax

Factor analysis/correlation Number of obs = 239

Method: principal-component factors Retained factors = 2

Rotation: orthogonal varimax (Kaiser off) Number of params = 6

--------------------------------------------------------------------------

Factor | Variance Difference Proportion Cumulative

-------------+------------------------------------------------------------

Factor1 | 1.64218 0.06947 0.4105 0.4105

Factor2 | 1.57271 . 0.3932 0.8037

--------------------------------------------------------------------------

LR test: independent vs. saturated: chi2(6) = 221.28 Prob>chi2 = 0.0000

Rotated factor loadings (pattern matrix) and unique variances

-------------------------------------------------

Variable | Factor1 Factor2 | Uniqueness

-------------+--------------------+--------------

bg\_sexual | 0.0819 0.8954 | 0.1915

bg\_mature | 0.8975 0.0102 | 0.1945

bg\_indepen~t | 0.8927 -0.0964 | 0.1939

bg\_respect~l | -0.1822 0.8726 | 0.2053

-------------------------------------------------

Factor rotation matrix

--------------------------------

| Factor1 Factor2

-------------+------------------

Factor1 | 0.7775 -0.6289

Factor2 | 0.6289 0.7775

--------------------------------

**.\*Index of general support for school punishment**

. factor corpun zerotol locksearch covid\_pun covid\_punS arrestkid arrestkidS , pcf

. rotate, varimax

Factor analysis/correlation Number of obs = 1,041

Method: principal-component factors Retained factors = 2

Rotation: orthogonal varimax (Kaiser off) Number of params = 13

--------------------------------------------------------------------------

Factor | Variance Difference Proportion Cumulative

-------------+------------------------------------------------------------

Factor1 | 2.42359 0.76987 0.3462 0.3462

Factor2 | 1.65373 . 0.2362 0.5825

--------------------------------------------------------------------------

LR test: independent vs. saturated: chi2(21) = 1654.59 Prob>chi2 = 0.0000

Rotated factor loadings (pattern matrix) and unique variances

-------------------------------------------------

Variable | Factor1 Factor2 | Uniqueness

-------------+--------------------+--------------

corpun | 0.5706 0.3961 | 0.5176

zerotol | 0.1431 0.8078 | 0.3270

locksearch | 0.0820 0.8216 | 0.3182

covid\_pun | 0.5338 0.3111 | 0.6183

covid\_punS | 0.6880 0.2343 | 0.4717

arrestkid | 0.7975 0.0057 | 0.3639

arrestkidS | 0.8225 0.1325 | 0.3059

-------------------------------------------------

Factor rotation matrix

--------------------------------

| Factor1 Factor2

-------------+------------------

Factor1 | 0.8404 0.5419

Factor2 | -0.5419 0.8404

--------------------------------

1. The treatment worked better for Black participants than for 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 were white respondents). [↑](#footnote-ref-0)