What's in a Mane?: Appearance Norms and Racialized Hair Bias*

Tarikua Erda† Jeffrey Shrader‡

March 26, 2026

[Latest version]

Abstract

Mainstream professionalism norms favor straight hair, imposing private costs of conformity on Black individuals whose kinky hair texture deviates furthest from this standard. We conduct incentive-compatible experiments in which subjects rate headshots of hypothetical candidates for STEM jobs on professionalism, competence, and agreeableness. Relative to straightened hair, wearing one's curly/kinky hair texture carries larger penalties for Black women compared to white women, primarily for perceived professionalism and competence. For Black men, such racialized penalties are concentrated in agreeableness. Smiling attenuates racialized penalties for Black models, consistent with costly signaling of warmth to offset unfavorable priors. Leveraging the clean identification that hairstyle affords, our study speaks to a broader class of malleable identity markers that individuals may alter to conform to dominant norms but only at ongoing private cost. Because these costs recur in daily life rather than at discrete episodes like hiring, they fall largely outside the existing literature's focus, implying that current estimates understate the welfare toll of discrimination. Our findings also inform ongoing legal discourse around recognizing hair texture as a protected racial characteristic and outlawing hair-based discrimination in the U.S.

JEL Codes: J15, Z10
Keywords: Norms, Implicit bias, Discrimination, Assimilation, Statistical Discrimination, Identity, STEM, Innovation and R&D

*We acknowledge generous funding from the Columbia Experimental Lab for the Social Sciences (CELSS) and the Institute for Social and Economic Research and Policy (ISERP). For feedback and comments, we thank Douglas Almond, Elizabeth Ananat, Sandra Black, Mark Dean, Alex Eble, Judd Kessler, Corinne Low, Suresh Naidu, Ebonya Washington, Jack Willis, and participants in the CELSS Experimental Lunch and the Workshop on Entrepreneurial Finance and Innovation Fellows (WEFI) group. This study was carried out under Columbia IRB (Protocol: AAAU0417). AEA-Preregistry information: AEARCTR-0010183. All errors and omissions are ours.

†Erda: New York University, tarikua.erda@nyu.edu.
‡Shrader: Columbia University, jgs2103@columbia.edu.

"As someone with naturally curly hair, I spent decades damaging it with flat irons to conform to a mainstream, 'professional' appearance, feeling that I had little choice if I wanted to advance up the corporate ladder." — Survey Respondent

1 Introduction

The existing literature on discrimination predominantly identifies bias through markers that evaluators treat as fixed, such as names on resumes or facial features (Bertrand & Mullainathan, 2004; Kahn & Davies, 2011). But many markers that shape daily experiences—how one speaks, dresses, or wears their hair—are mutable to varying degrees. Where conforming to dominant group norms carries economic returns (Akerlof & Kranton, 2000), these traits generate recurring private costs that existing studies focused on discrete episodes like hiring and lending[1] do not capture. We provide direct evidence on these costs, focusing on hair, a trait where conformity costs are directly observable, large, and extend beyond pecuniary margins.

Hair *texture*—whether strands are straight, wavy, curly, or kinky—is an immutable feature, yet hair *style* can be continuously altered using chemicals, heat, and various accessories, often in ways central to self-expression, cultural identity, and religious practice. Bias against the naturally kinky hair texture of populations of African origin dates to the Transatlantic Slave Trade, and is one of many cases across societies where hair has been used to enforce cultural conformity or dominance.[2] Mainstream norms that view the natural curly/kinky hair of Black individuals as inappropriate for the workplace impose costs along several margins: Black Americans spend about a 30% larger share of their income, and dedicate roughly 64 hours—8 work days—per year on personal grooming relative to their white counterparts (see Figure 1). About 80% of Black women report changing their hair from its natural kinky state to fit in at the office (JOY Collective, 2019), and they are three times as likely as white women to skip exercise to preserve straightened hairstyles (Johnson et al., 2017). Hair straightening also carries health risks, including cancer-related ones.[3] These costs primarily fall on women, but Black boys and young men also face penalties, includ-

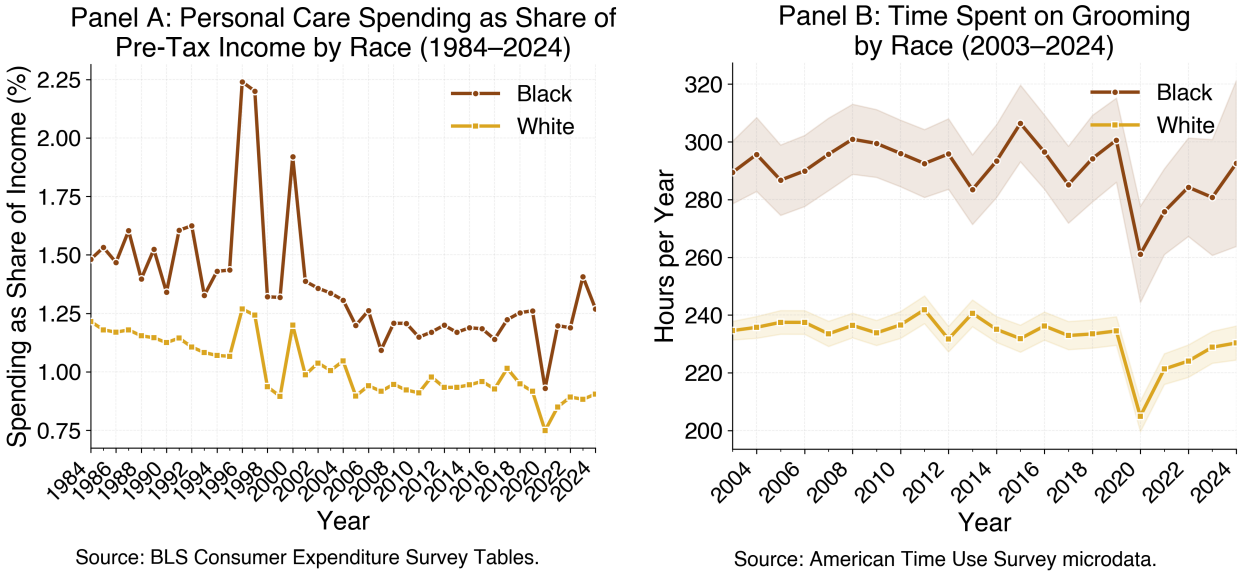
[1]The existing empirical literature on racial discrimination centers on discrete decision points such as in hiring (Bertrand & Mullainathan, 2004; Agan & Starr, 2018), criminal justice (Arnold et al., 2018; Rehavi & Starr, 2014), lending (Bartlett et al., 2022), medical treatment decisions (Hoffman et al., 2016).

[2]European slave traders shaved the heads of enslaved Africans upon arrival in the Americas (Byrd & Tharps, 2002). US boarding schools forcibly cut Native American children's hair as part of assimilation campaigns. San Francisco's Queue Ordinance of 1876 targeted Chinese immigrants by mandating hair cutting in jails. Nazi concentration camps shaved Jewish prisoners' heads upon arrival (Levi, 1958).

[3]See Eberle et al. (2020); De Sá Dias et al. (2007); Woolford et al. (2016); Callender et al. (2004).

ing suspensions from school or sports for wearing dreadlocks.[4] The economic significance of these costs has gained growing recognition in policy: the *CROWN Act*, which protects Afro-textured hair and hairstyles associated with race like braids and locs in employment and schools, has been adopted in more than 20 US states since 2019.

Figure 1: Large, persistent Black-white gaps in money and time spent on personal care[5]



We conduct incentive-compatible, pre-registered experiments to assess appearance norms in STEM academia—an economically important setting where the role of appearance is ex-ante ambiguous. In principle, given that academia prides itself on meritocracy, appearance may not matter at all. Yet, attractiveness predicts success even in ostensibly merit-based contexts like economics faculty careers (Hale et al., 2023) and entrepreneurial pitch competitions (Brooks et al., 2014). Hair-based bias documented in industry[6] may thus extend to academia, but it is still unclear whether 'unconventional' hair may be penalized or rewarded as a signal of competence (invoking the mad-scientist trope (Chambers, 1983)). The setting also carries high stakes for both efficiency and equity: STEM fields train the innovation and R&D workforce that drives productivity growth and also command large wage premiums, so discrimination that deters talented Black individuals from entering or advancing can distort aggregate growth and worsen racial income gaps.[7]

[4] See APNews (2023) and Reuters (2024).

[5] In Panel (A), "White" denotes the BLS category "White and Other" prior to 2003 and "White" thereafter. Panel (B) plots annual hours spent on washing, dressing, and grooming oneself (ATUS activity code 010201) by race, using ATUS microdata for $N \approx 203{,}000$ non-Hispanic White and Black respondents. Shaded bands in Panel (B) represent 95% confidence intervals around survey-weighted means.

[6] See Opie and Phillips (2015); Koval and Rosette (2021).

[7] Hsieh et al. (2019) estimate that declining barriers faced by women and Black Americans account for

We survey university students and a representative sample of US adults, asking each respondent to rate 32 headshots of hypothetical STEM professor candidates on professionalism, competence, and agreeableness. Students are embedded in the academic norms they are evaluating; lay adults provide a benchmark for whether our findings are specific to academia or reflect broader societal attitudes. The headshots feature four models—a Black female, a white female, a Black male, and a white male—generated through an original photoshoot with editing monitored by the authors, and depict three hair looks: natural curly/kinky hair, dreadlocks, and straightened hair (clean-cut for men), with additional variation in attire, facial expression, and photo resolution to obscure our central focus on hair.[8]

Our design incorporates several features to ensure both clean identification and honest responses. First, before they start rating, we prime respondents with casual weekend photos of the models to establish that all models have naturally curly/kinky hair, so that subsequent hairstyle variation is understood as styling choices relative to a shared baseline texture. Second, we employ higher-order belief elicitation: respondents were offered one of three bonus $100 cash prizes for accurately predicting how university staff we had previously surveyed rated the same photos, making it safer to express potentially stigmatized views. Third, we include non-experimental photos of other-race models to reduce fixation on Black-white comparisons. Our main estimates come from OLS regressions that exploit within-model variation in hairstyle, comparing how the same individual is rated in natural versus straightened hair across Black and white models of the same gender.

We find pronounced penalties in perceived professionalism and competence for the Black female model when she wears her naturally kinky hair instead of straightened hair. In the student sample, these racialized penalties represent about 5% of the baseline rating and far exceed the penalty for a blurry, low-resolution photograph—a feature objectively within individuals' control. Students rate the white female model's comparable curly hairstyle as *more* competent and agreeable than her straight-hair photo, while penalizing the Black female. Lay adults show a similar pattern, though the racialized gap is attenuated. That students show penalties at least as large suggests that proximity to the ivory tower does not blunt appearance-based bias; if anything, it sharpens it.

For the Black male model, the excess penalty is concentrated in agreeableness: students award the white male a significant *premium* for curly hair while penalizing the Black male, consistent with natural Black hair on men activating priors associated with perceived threat

a significant share of U.S. productivity growth since 1960, while Cook (2014) and Cook and Yang (2018) document innovation losses from the historical exclusion of Black Americans.

[8]The natural hair texture looks of the models were rated to represent comparable levels of effort in prior focus group surveys of lay adults with matching (self-reported) race and hair texture.

(Hester & Gray, 2018). On professionalism and competence, the pattern reverses, with the white male penalized more for deviating from clean-cut hair. Evaluators appear to map the same deviation from straight-hair norms onto different deficits depending on race and gender—less polished and competent for Black women, less approachable for Black men.

Our results also point to a statistical discrimination channel behind the agreeableness penalties. The racialized agreeableness penalty is large when Black models wear a neutral expression but vanishes when they smile, consistent with evaluators updating unfavorable priors in response to a warmth signal (Phelps, 1972). One source of these priors may be historical: during the Civil Rights era of the 1960s, the "Black is Beautiful" movement encouraged Black Americans to embrace their natural hair as an act of self-acceptance, but wearing an Afro was widely perceived by mainstream society as a political statement and a sign of militancy (Byrd & Tharps, 2002), an association that experimental work in psychology finds persists today in the form of perceived dominance attributed to Afrocentric hairstyles (Opie & Phillips, 2015). Smiling does not, however, reduce the professionalism or competence penalties, which remain anchored in conformity to appearance norms. That Black candidates can offset interpersonal penalties only by supplying warmth that white or straight-haired candidates need not provide represents a form of emotional labor that existing estimates of discrimination costs do not capture.

When it comes to dreadlocks, the racial gap in ratings is reversed as both student and lay adult respondents severely penalize white models, especially on professionalism and competence, with many citing 'cultural appropriation.' Despite these stated beliefs about dreadlocks as a hairstyle of particular cultural significance to Black people, respondents still penalize Black models in absolute terms.

Finally, we test two information interventions: a vignette on workplace authenticity and productivity, and an explainer on the history of hair-based racial discrimination. Apart from conditions that induced experimenter demand per (de Quidt et al., 2018), neither treatment reduced racialized hair penalties. This suggests that simple informational interventions to mitigate hair bias may be ineffective on their own.

In sum, our results reveal that racialized hair penalties are large, intersectional, and operate through distinct channels for Black women and Black men. We also identify emotional labor as a previously undocumented component of these costs, where Black individuals can offset hair penalties but only by bearing asymmetric burdens of signaling warmth. More broadly, while hair affords unusually clean identification, it is one of many traits that individuals did not choose but can alter to access economic opportunity (e.g., skin bleaching in South Asia and Africa, accent modification among immigrants). Each presents the same

double bind: deviate and face penalties, or assimilate and bear recurring private costs. The welfare toll of discrimination is thus likely understated by a literature focused on markers that evaluators treat as fixed. Finally, that hair penalties resist light-touch informational correction is relevant to ongoing policy debates, including the CROWN Act, over whether hair texture warrants formal legal protection.

Related Literature: This paper makes four key contributions. First, we extend the empirical literature on racial discrimination in labor markets by shifting attention from immutable racial markers to one that is malleable but nonetheless penalized. Prior work documents returns to appearance-related attributes such as beauty (Hamermesh & Biddle, 1994), height (Persico et al., 2004), weight (Cawley, 2004), and dental health (Glied & Neidell, 2008; Gallego et al., 2024), as well as the costs of adjusting identity signals through surname changes (Arai & Skogman Thoursie, 2009) or resume whitening (Kang et al., 2016). We show that the malleability of hair does not attenuate the evaluation penalties associated with natural Black hairstyles; several penalties we estimate for the Black female model's natural kinky hair texture in fact exceed those for attributes squarely within individuals' control, such as smiling and photo resolution. Hair straightening imposes recurring costs, suggesting that the welfare losses from discrimination are larger than studies focused on discrete episodes like hiring would indicate.

Second, we broaden the scope of identity-specific costs documented in prior work. Existing estimates of such costs—most notably, gender-based price discrimination (i.e., the pink tax) (de Blasio & Menin, 2015)—have centered on pecuniary margins. That smiling, a costly signal of agreeableness, attenuates the racialized hair penalty for Black models is suggestive of non-pecuniary costs, a sort of warmth-tax. Additionally, consistent with recent empirical work showing that labor market penalties associated with multiple marginalized identities interact in non-additive ways (Paul et al., 2022), our use of an explicitly intersectional approach (Crenshaw, 1989) reveals that hair penalties fall disproportionately on Black women, and cannot be recovered by examining race or gender in isolation.

Third, this study advances the methodology of studying unconscious bias and contributes new evidence on the costs and mechanisms of hair bias, and if/how it can be mitigated. Prior work in adjacent social science disciplines has used experimental approaches to document hair bias in settings such as advertising and consulting (Koval & Rosette, 2021) and corporate employment (Opie & Phillips, 2015). We build on this foundation by introducing several innovations: our experiment is the first to use pre-registered, incentive-compatible designs; to examine hair bias as it applies to both Black women and Black men; and to explicitly account for social desirability bias. Our experimental photos are designed

to achieve high realism and precisely controlled variation, and our study is unique in incorporating information interventions to evaluate whether hair bias can be moderated. Rather than using methods that reveal the bias under study to subjects—such as the Implicit Association Test (IAT)—our approach obfuscates the survey's central focus on hair through a holistic headshot rating exercise, offering a more reliable measure of unconscious bias.

Fourth, by situating our experiment in STEM academia, we contribute to the literature on barriers to minority representation in innovation-intensive fields. This literature examines whether representation gaps reflect comparative advantage or discriminatory environments, including stereotype threat and discriminatory evaluation (Steele et al., 2002; Francis & Darity, 2020), discrimination in professional network formation (Ajzenman et al., 2025), the absence of same-race and same-gender senior faculty who could serve as role models for minority students (Bettinger & Long, 2005; Redding, 2019). Our findings identify appearance-based bias as a previously undocumented channel through which the academic environment may deter Black scholars.

2 Background

2.1 Historical Overview on Hair Bias

For centuries, hair was and continues to be used as a marker of various cultural indications such as age, marital status, social status and so on in Africa (Byrd & Tharps, 2002). Since the onset of Slave Trade, however, institutionalized racism has affected, and in part dictated, how Black people in the US navigate hair. Europeans shaved the heads of enslaved Africans upon arrival to the Americas to assert domination, reallocate time toward labor (Chapman, 2007; Dash, 2006; Morrow, 1973). In 1786, the Tignon Law required *all* women of color–including emancipated ones–in New Orleans to wear a scarf over their hair as a visible sign of belonging to the 'slave class.' From late 1800s on, various hair products, mainly chemical and heat-based tools to straighten Black hair were introduced to the market. It was encouraged, even within commonly held wisdom in Black communities, that one had to 'tame' their kinky hair and straighten it to get and maintain jobs.

Later, during the 1960s' Civil Rights Movement, more Black people embraced the act of simply wearing their natural hair texture as an Afro as an act of self-acceptance (Byrd & Tharps, 2002). However, it was not too long before hair straightening products gained traction again, and other styles like locs, braids etc on the rise, with greater 'natural hair movement' since early 2000s. Today, the CROWN Act, introduced in 2019, outlaws hair

Figure 2: An advertisement for a hair relaxer from the 1970s asks whether the pictured Black woman got her job because of her resume or her straightened hair



bias & discrimination and is now adopted in more than 20 states).

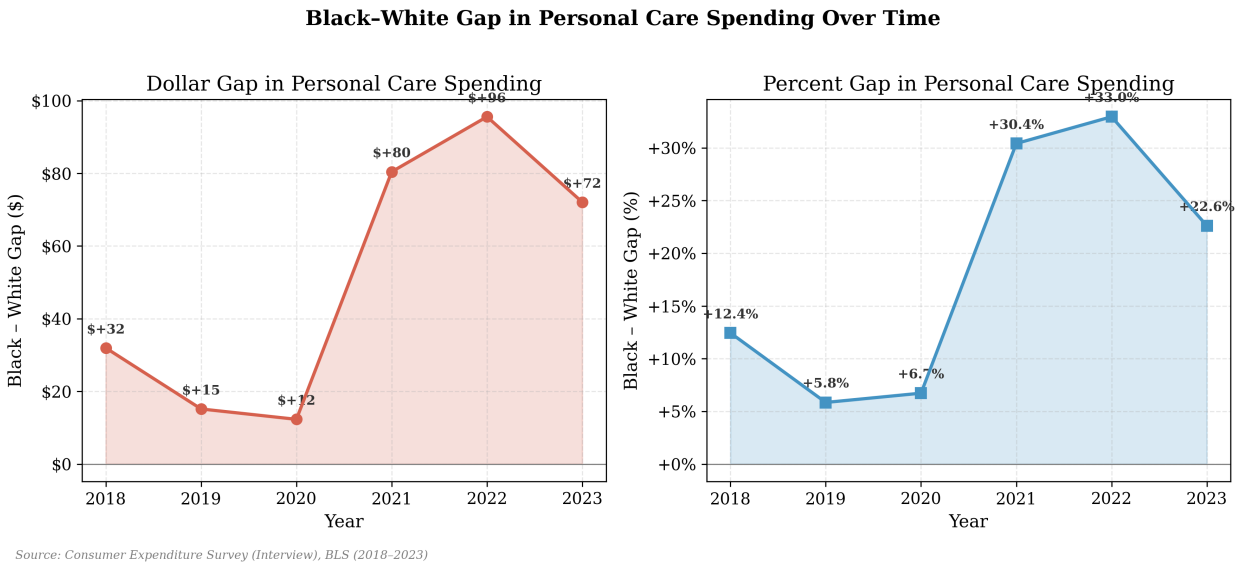
2.2 Economic Implications of Hair Bias

Hair bias carries substantial and multifaceted costs. While no nationally representative survey isolates hair care spending by race, we show in Figure 1 (left panel) using the Consumer Expenditure Survey—which tracks total personal care products and services—that, over the past four decades (1984–2024), Black households have devoted roughly 1.34 times as much of their income to personal care as white households. This ratio that has remained remarkably stable across business cycles, recessions, and shifting consumer trends. In the most recent year (2023), the gap stands at 0.52 percentage points—Black households spend 1.41 percent of income on personal care versus 0.88 percent for white households—one of the widest gaps in the post-2000 period. This gap is doubly regressive because Black households both spend more in absolute dollars and earn less than whites.

Further, regression-adjusted analysis of CES Interview Survey microdata ($N = 17{,}130$ consumer units) confirms that the gap is not an artifact of demographic composition: controlling for income, age, household size, education, region, and urban versus rural status,

Black households spend $146–$173 more per year on personal care than observably similar white households, with the difference driven almost entirely by salon and barbershop services (see Figure 3).

Figure 3: Gap in spending on personal care



Beyond the time and monetary costs, chemical hair straighteners (called "relaxers") have been linked to several toxic, and even carcinogenic, outcomes[9] while thermal straightening can cause heat damage that permanently alters one's curl pattern and results in hair loss.[10] Therefore, norms of professionalism impose a dual burden on Black Americans, with ongoing costs that compound over a lifetime.

3 Experimental Design

3.1 Experimental Visual Stimuli

Since our study is centered on appearance, the visual stimuli in our survey are a key element in our survey. We conducted an original photoshoot in May 2022 with four main models: a Black female, a white female, a Black male, and a white male. Models were photographs in different attires and facial expressions and the authors also closely supervised the photoshoot to minimize variation in elements that would not be directly controlled for in our experimental analyses. We subsequently hired a photo editor through the online platform Fiverr

[9] See, Eberle et al. (2020); França-Stefoni et al. (2015); De Sá Dias et al. (2007); James-Todd et al. (2011); Woolford et al. (2016).

[10] Callender et al. (2004).

to generate pictures of them in different hairstyles. The blurring effect on the originally high resolution photos was added after all photo editing was completed.

A key potential confounder when studying the perception of hair texture, particularly that of women, in large samples of respondents from various gender and racial backgrounds is the level of effort signaled by hairstyle, which we accounted for by running smaller online focus group surveys related to the perception of effort. This is because individuals are more likely to be familiar with their own hair texture and the effort involved with certain hair styles in their gender and racial demographic group, compared to the hair texture and effort of other demographic groups. This is especially relevant to women's hair they face a larger grooming burden in many societies, and hair straightening is common in the US among women of various races. Thus, to account for the role of effort, we surveyed white and Black lay adults who (self-reportedly) had curly and kinky hair rated the level of effort suggested by the models' curly and kinky hairstyles, respectively. We asked subjects to think of effort put into the pictured hairstyle as a combination of (1) time spent on the pictured hair, (2) the amount of hair products applied, and (3) the physical energy the pictured individual spent working with different tools and manipulating their hair to create the pictured hairstyle. The white and Black female models' hair were rated to signal comparable levels of effort by white and Black female focus group respondents.

Since men face a lower grooming burden and are far less likely to manipulate their hair texture (e.g., straighten curly/kinky hair) relative to women, what matters for our study of the perception of men's hair is that there are palpable differences in effort along the lines of the recency of a hair cut.[11] As such, the non-dreadlock hairstyle variations for the male models in our study are based on whether the model has longer/bigger curly/kinky-Afro hair or shorter hair that looks clean-cut/polished, suggesting a more recent visit to a barber.

The final set of experimental visual stimuli were prepared by combining, for each of the four models, 3 hairstyle variations (kinky/curly, dreadlocks, and straight/clean-cut hair), 2 clothing variations (suit and t-shirt), 2 facial expressions (smiling and non-smiling) and, for all of the above 24 combinations but three of them, 2 photo-resolution levels (high-resolution and blurry). Because of this last step of not including blurry versions for 3 photos per model, we ended up with 84 instead of 96 unique experimental photos, which meaningfully increases each photo's probability of being shown to and rated by respondents while retaining decreasing our ability to recover the role of photo resolution on respondent rating through the combinations of estimates on the rest of the model's photos.

[11] Unsurprisingly, the aforementioned definition of 'effort' we gave to focus group participants—which indeed accomplished its purpose of distinguishing effort from hair texture for the female models' photos—resulted in seemingly highly subjective and varied estimates, particularly for the white male.

To add racial diversity in the set of photos that subjects rated, we also included 12 stock photos of individuals who presented as neither white nor Black but whose photo characteristics (e.g., photo background, clothing style and angle) generally resembled those of our experimental photos. This brought the total number of visual stimuli from which respondents would rate to 98. Thus, on average, respondents in our experiment encountered the main experimental photos of our four white and Black models about 86% of the time.

3.2 Survey Procedure

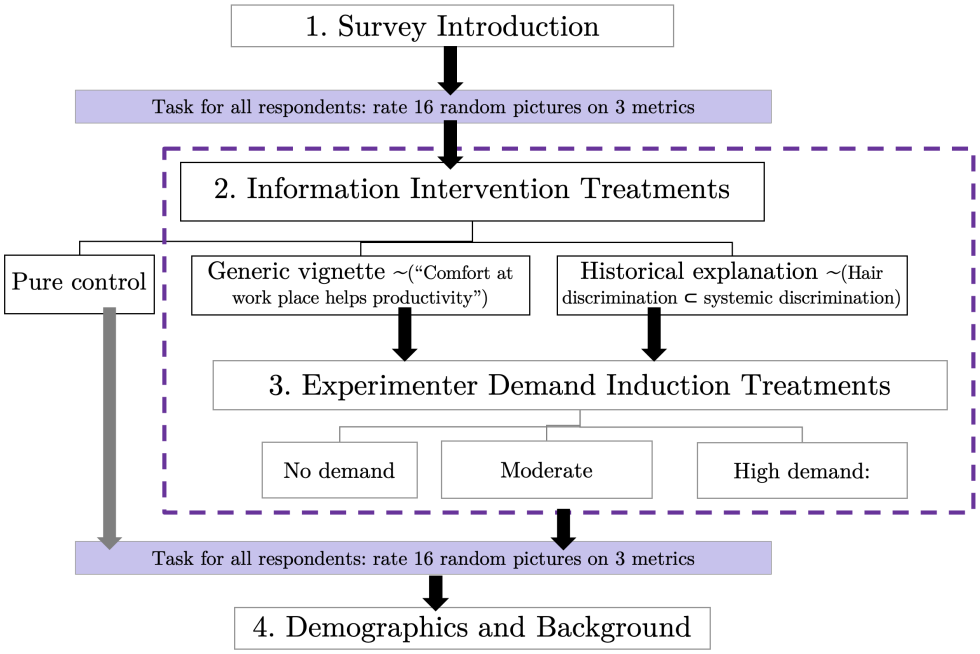
University students were recruited from the Online Recruitment System for Economic Experiments (ORSEE) subject pool while lay adults were recruited through CloudResearch. The main survey activity was rating 32 headshots in four blocks of 8 photos each. Since student participants rated headshots imagining that the pictured models were candidates for a university professor position in the field of their major, they were asked about their major at the start of the survey. Lay adults were all asked to imagine that pictured models were hypothetical candidates for STEM professor jobs. After completing the first two blocks of ratings (after 16 photo ratings), respondents were randomly assigned to one of three categories at an even 1/3 chance: a pure control group, an information treatment involving a vignette about how authenticity at the workplace enhances diversity, and another information intervention that explains the history of hair bias in the US and the CROWN Act. Before proceeding to the last 16 photos, treated subjects were subsequently assigned to one of three groups that received no experimenter demand induction (with 50% chance), a moderate or high experimenter demand (each with 25% chance). The survey flow is presented in Figure 4.

The rating activity involved three metrics: professionalism, competence, and agreeableness. We defined professionalism for them as the state of being in line with the norms of the university as a workplace, competence as being qualified, credible, and capable to rise to the demands of an academic job (teaching and doing research), and agreeableness as the quality of being collegeal and collaborative with other scholars as well as approachable to students seeking answers to their questions or mentorship from faculty. While respondents were asked to rate a Likert scale from 1-10, we also provided a non-response option ("N/A") in case they wanted to skip a rating.[12]

To ensure incentive compatibility, we gave respondents an explanation of the bonus cash prize we were offering for accurate responses right after we introduced the headshot rating activity and rating metrics. We explained that we had conducted a similar headshot

[12]For instance, the 'competence' rating, many subjects used this 'N/A' option or at least expressed discomfort with evaluating competence based on a appearance).

Figure 4: Survey Flow



rating exercise on university administrative staff who were in a position to train and advise university students in an approach that highly encouraged truthful responses. This was indeed true – we had conducted a headshot-rating exercise with university staff from a New York area higher education institution in May 2023, prior to the launch of the main surveys on students and lay adults in October 2023. As a reward for the participation of the university staff in our survey, we contributed money to a professional headshot session held for students of that institution and also shared aggregated insight from the survey responses as input for the headshot session.[13]

Prior to the start of the rating activity, we also presented subjects with a primer to 'acquaint' them with the 8 models (four main and four joker models) whose photos they would subsequently rate using small photos that showed the models on 'a casual weekend morning.' This was intended to implicitly give respondents a benchmark as to the natural state of each model on a casual day lest they make other, potentially heterogeneous, assumptions about the experimental models that interfere with our inference of the photo ratings. For example, without a benchmark, subjects may assume the natural hair texture of the white female model is in fact straight but that she manipulated it to look curly, which would impair our investigation of the perception of similar non-straight natural hair textures in the Black and

[13]This was intended to encourage truthful responses, in a fashion similar to the Incentivized Resume Rating approach pioneered by (Kessler et al., 2019) in which employers rate resumes they know to be hypothetical in return for matches with *real* job candidates.

white female models relative to straightened hair.

4 Data and Empirical Strategy

4.1 Sample Descriptions

As Table A.1 shows, there are pronounced differences in the composition of our student and lay adult samples. As it is drawn from a representative sample of the US population, the lay adult sample mirrors the racial break-down of the US population and also has a roughly even split across male and female respondents. In contrast, more than 70% of students identify as female, which partly follows from the fact that a decent share of our respondents attend Barnard College, a women's college. The student sample, which was recruited through the ORSEE which does not allow targeted stratification of the sample, is primarily composed of students with white (46%) and Asian (39%) background, which are both much higher than the Columbia student population distribution. Black students, other multi-racial or minority groups made up a smaller proportion of the student sample, including compared to the relative shares in the Columbia and Barnard student populations. Additionally roughly half of lay adult respondents reported being democrats, and 22% republicans, while 71% of students reported being democrats and less than 3% republican.

About 67% of students had a STEM major. These majors largely include computer science, engineering, chemistry, physics, life and environmental sciences as well as cognitive science and quantitative-driven majors such as economics and data science. All students rated photos imagining that the pictured model was a candidate for a professor position in their major discipline. Our main results presented in Section 5 juxtapose the views of these students and contrast them against those of lay adults, who were all asked to imagine STEM professors in their ratings.

4.2 Estimation Strategy

Using the data from the first half of the ratings, we estimate the below specification.

$$\begin{aligned} V_{ij} = \beta_1 B_j + \beta_2 G_j + \beta_3 H_{2,j} + \beta_4 H_{3,j} + \beta_5 B_j G_j + \beta_6 B_j H_{1,j} + \beta_7 B_j H_{2,j} \\ + \beta_8 B_j H_{3,j} + \beta_9 G_j H_{1,j} + \beta_{10} G_j H_{2,j} + \beta_{11} G_j H_{3,j} + \beta_{12} B_j G_j H_{1,j} \\ + \beta_{13} B_j G_j H_{2,j} + \beta_{14} B_j G_j H_{3,j} + X_{ij}'\theta + \alpha_i + \varepsilon_{ij} \end{aligned}$$

V_{ij} denote respondent i's rating of a photo j on the 1–10 Likert scale. B denotes Black model, G denotes female model, H_1, H_2 and H_3 are hair style indicators for curly/kinky hair, dreadlocks, and straightened (and for the male models, clean-cut hair), respectively. The other variables in the regression are controls. These include a subject-level fixed effects α_i and, in additional specifications, controls for other photo attributes (represented by X): an indicator for whether the model is wearing a suit jacket (C), and indicator for whether the photo is in high resolution (R), and an indicator for whether the model is smiling (E).

The main hypotheses we will test are if β_{12}=0 and β_{14}=0 to understand whether the Black female model rated the same for curly or loc hairstyles as for straight. Additionally, we will compare the relative penalties faced by Black and white models of of the same gender for curly/kinky hair relative to straight/clean-cut hair.

5 Results

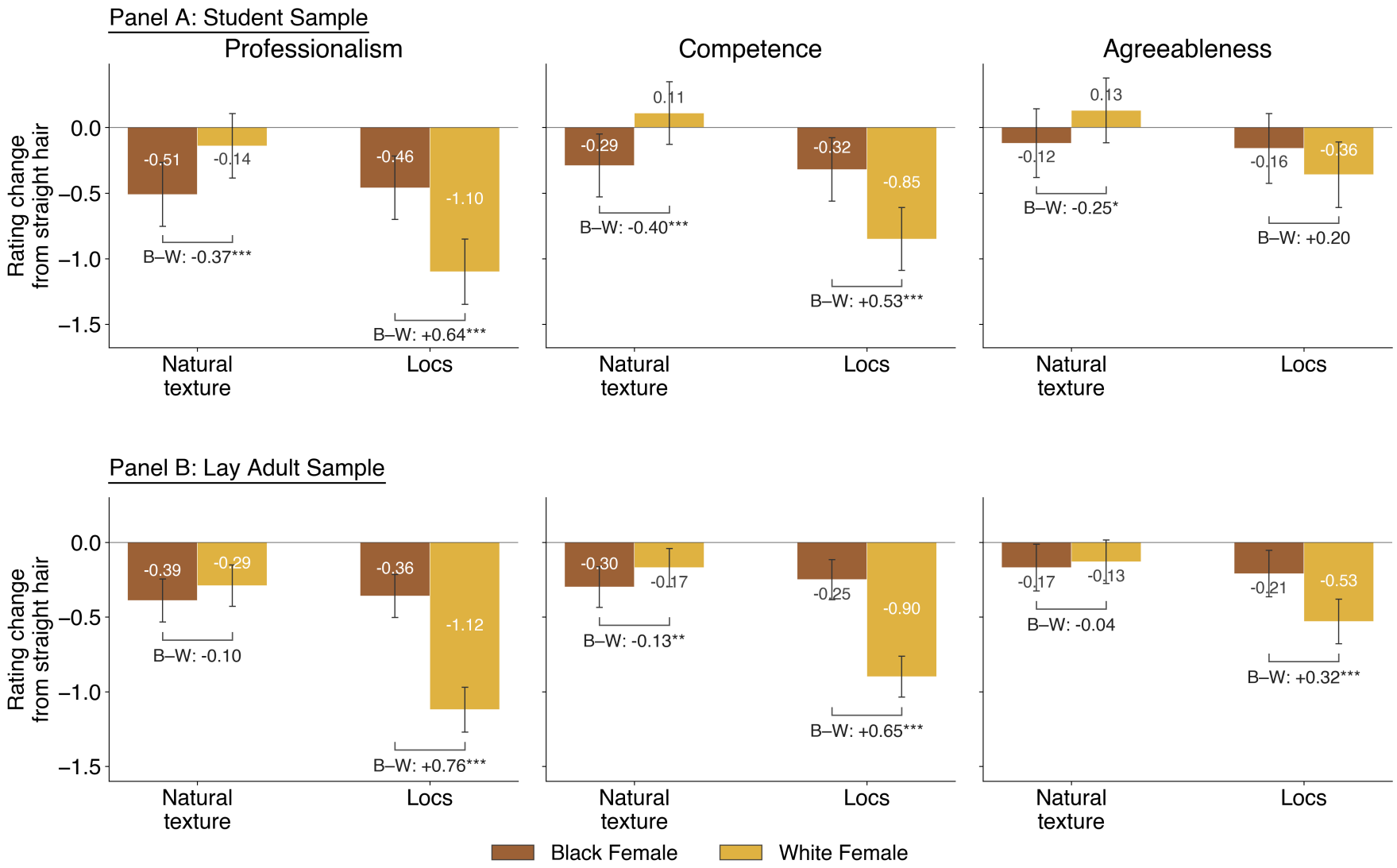
5.1 Baseline ratings for Female Models

Figure 5 presents regression-adjusted predicted ratings for the Black and White female models across the three hairstyle conditions, separately for the student (Panel A) and lay adult (Panel B) samples. These predicted ratings are derived from estimates of Equation (4.2), which regresses each rating outcome on the full set of race-by-hairstyle interactions along with controls for clothing, facial expression, and photo resolution, and subject fixed effects. The figures plot predicted values at the mean of the control variables. Each figure depicts the racialized hair penalty, i.e., the difference-in-differences between the Black model's hairstyle differential and the White model's equivalent differential, at the bottom left. Full regression tables with detailed coefficients for all photo attributes are presented in Appendix Tables A.2–A.4.

Panel A of Figure 5 shows a clear asymmetry in how natural hair texture affects ratings of the Black versus White female models in the university student sample. Relative to straightened hair, wearing naturally curly or kinky hair reduces the Black female model's predicted professionalism rating by 0.51 points ($p < 0.01$) and her competence rating by 0.29 points ($p < 0.01$). The White female model's corresponding curly hair penalties are far smaller: 0.14 points for professionalism, statistically insignificant. On competence, the White female model actually receives a *premium* of 0.11 points for wearing curly hair relative to straight. Students rate the Black female model as less competent for wearing her natural hair texture while simultaneously rating the White female model as *more* competent for

wearing hers. The racialized curly hair penalty—the difference-in-differences between the Black model's curly-versus-straight gap and the White model's—is -0.37 points ($p < 0.01$) for professionalism and -0.39 points ($p < 0.01$) for competence.

Figure 5: Average respondent ratings for Female Models by Hairstyle and Race



Bars show deviations from each female model's own straight-hair predicted rating (set to zero), based on Equation (4.2) evaluated at control-variable means. Error bars are 95% CIs. Brackets report the "B–W gap" (difference-in-differences: Black model's hair penalty minus White model's), with significance from the race×hairstyle interaction coefficient. All regressions include subject fixed effects and controls for clothing, facial expression, and photo resolution. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The pattern for dreadlocks is more nuanced. The White female model bears the heavier penalty: her predicted professionalism falls by 1.10 points when pictured in locs rather than straight hair, compared to 0.47 points for the Black female model. On competence, the gap is similarly lopsided: the White model loses 0.85 points versus the Black model's 0.32. The racialized locs gap is large and positive: +0.63 for professionalism and +0.54 for competence. In our open-ended response questions, several students expressed strong opposition to White individuals wearing dreadlocks, characterizing it as cultural appropriation. Despite this asymmetry, the Black female model is still penalized in absolute terms for wearing locs relative to straightened hair on every metric.

Panel B shows that lay adults penalize natural hair for both models, but the racialized gap in curly hair penalties is attenuated relative to the student sample. The Black female model loses 0.39 points on professionalism and 0.30 on competence when pictured in curly versus straightened hair (both highly statistically significant), while the White female model loses 0.29 and 0.17 points, respectively. Both models are penalized, but the gap between them is smaller: the racialized curly hair penalty is −0.10 for professionalism and −0.14 for competence, neither of which survives correction for multiple comparisons. This attenuation may partly reflect the greater demographic diversity of the lay adult sample and possibly different baseline priors about natural hair in professional settings.

On dreadlocks, lay adults follow the same pattern as students. The White female model's professionalism drops by 1.12 points when pictured in locs, compared to 0.36 points for the Black female model; on competence, the penalties are 0.90 versus 0.25 points. The racialized locs gap is large and significant across all three outcomes: +0.76 for professionalism, +0.65 for competence, and +0.33 for agreeableness .

Across both panels, the pattern is consistent: naturally curly hair is penalized more when worn by a Black woman, while it may actually be rewarded on the White model as a signal of competence or approachability. Respondents disproportionately penalize White models for wearing dreadlocks, but the Black female model is still penalized for the hairstyle in absolute terms.

5.2 Baseline Ratings for Male Models

Panel A of Figure 6 shows that for male models in the student sample, the racialized hair penalty operates along a different margin than for females: agreeableness rather than professionalism or competence. Relative to clean-cut hair, wearing longer curly hair reduces the Black male model's predicted agreeableness by just 0.04 points, while the White male model

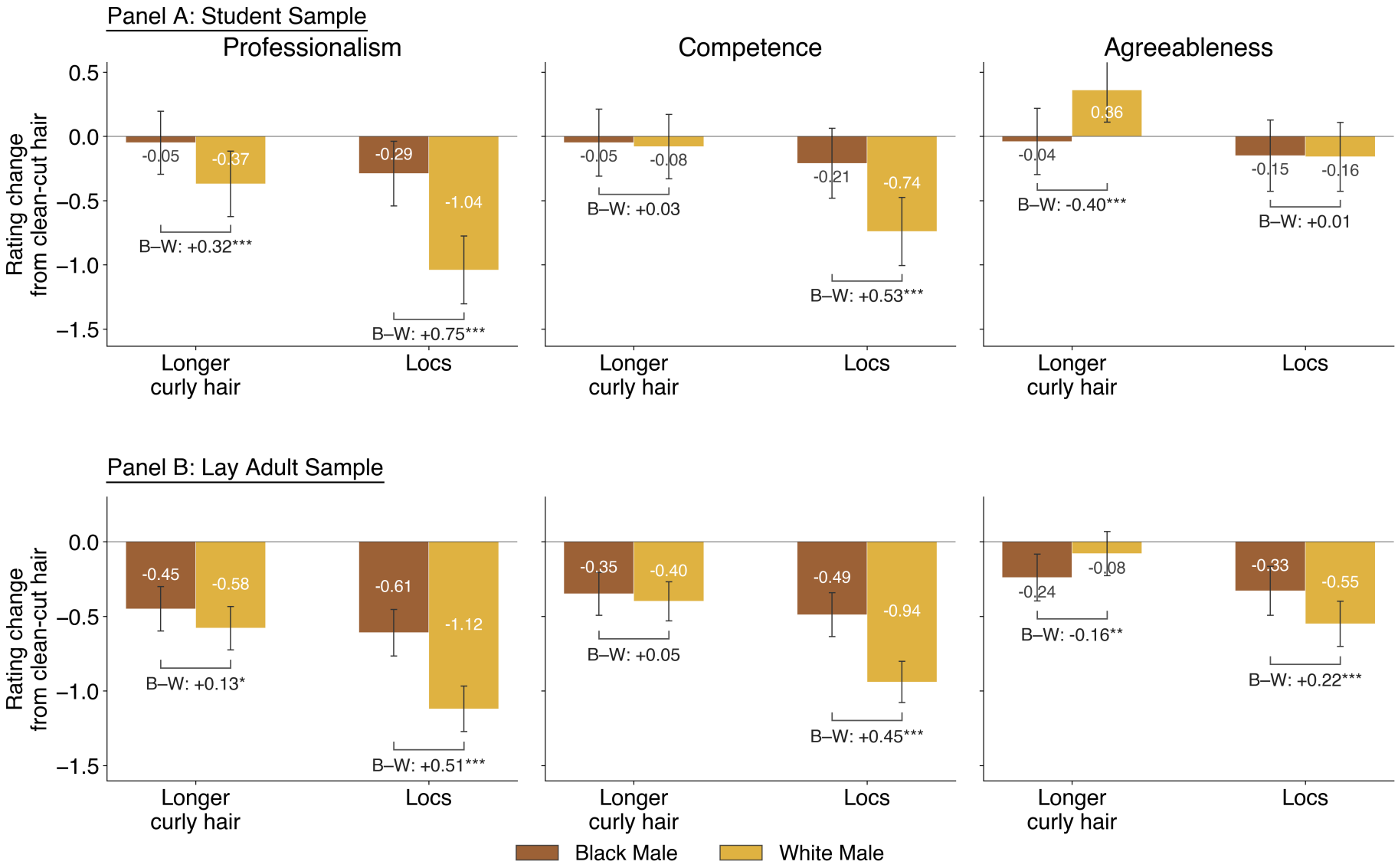
receives a *premium* of 0.36 points for the equivalent curly hairstyle. The racialized curly hair penalty on agreeableness is thus −0.40 points. On professionalism and competence, by contrast, the Black male model's Afro penalties are small and no larger than the White model's: the racialized gaps are +0.32 and a statistically insignificant +0.03, respectively.

Dreadlocks follow the same pattern as the female models: the White male model bears the heavier penalty. His professionalism drops by 1.04 points when pictured in locs versus clean-cut hair, compared to 0.29 points for the Black male model. On competence, the penalties are 0.74 versus 0.21 points. The racialized locs gaps are large and positive: +0.75 for professionalism and +0.54 for competence. On agreeableness, however, the racialized locs gap is near zero (+0.01), a contrast with the large negative Afro gap on the same metric. The Black male model is still penalized in absolute terms for wearing locs relative to clean-cut hair across all three outcomes.

Panel B shows that lay adults follow a similar pattern, with agreeableness again the primary site of racialized penalties for the Afro hairstyle. The Black male model loses 0.24 points on agreeableness when pictured in a bigger Afro versus clean-cut hair, while the White male model loses only 0.08 points. The racialized Afro gap on agreeableness is −0.16. On professionalism, the racialized Afro gap is +0.12, and on competence it is an imprecisely estimated +0.06, confirming no disproportionate penalty for the Black male on these dimensions.

For dreadlocks, lay adults penalize the White male model more heavily than the Black male model across the board. The White model's professionalism drops by 1.12 points in locs versus clean-cut, compared to 0.61 for the Black model; on competence, 0.94 versus 0.49. The racialized locs gaps are +0.50 for professionalism, +0.45 for competence, and +0.22 for agreeableness. Unlike in the student sample, the lay adult racialized locs gap on agreeableness is positive and significant, suggesting that lay adults penalize the White male for locs on agreeableness as well.

Figure 6: Average respondent ratings for Male Models by Hairstyle and Race



Bars show deviations from each male model's own clean-cut-hair predicted rating (set to zero), based on Equation (4.2) evaluated at control-variable means. Error bars are 95% CIs. Brackets report the "B–W gap" (difference-in-differences: Black model's hair penalty minus White model's), with significance from the race×hairstyle interaction coefficient. All regressions include subject fixed effects and controls for clothing, facial expression, and photo resolution. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

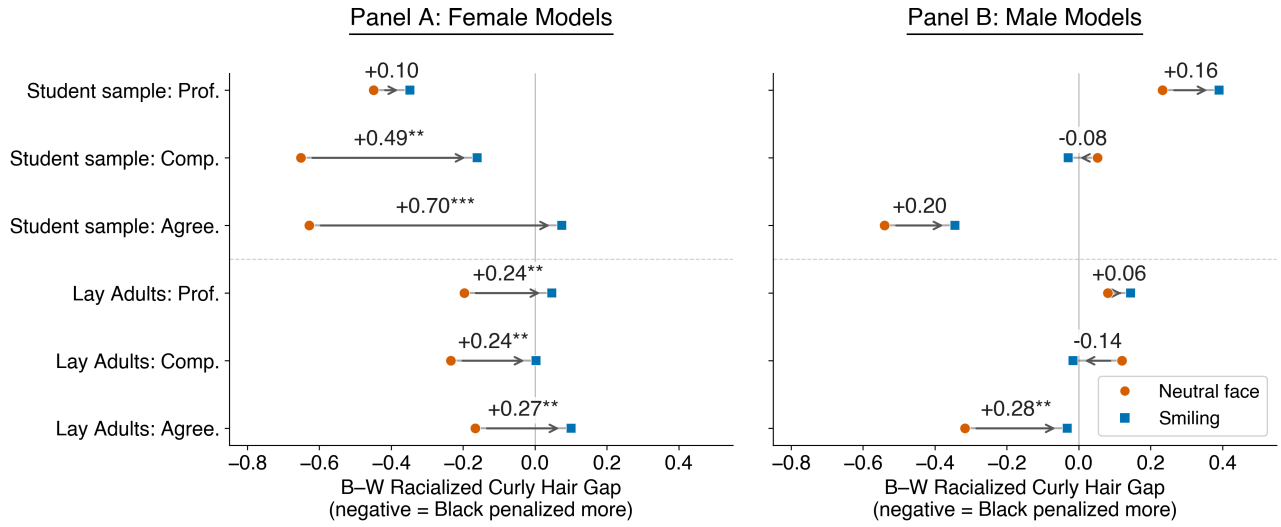
The male models' results present a pattern that is, at first glance, puzzling. While the Black female model faces racialized penalties concentrated in professionalism and competence, the Black male face them primarily on agreeableness, and he in fact faces lower penalties for professionalism and competence than white male model. Since male appearance norms have received considerably less empirical attention than female ones, we offer two tentative interpretations. First, the apparent advantage may reflect expectation congruence: the Afro reads as phenotypically natural on the Black male model, whereas the equivalent curly style on the White male model may register as unkempt or lacking effort in a way that respondents understand as less professional. Second, model-specific factors may play a role: we only use one model per race x gender category, so it may well be that case the Black male model simply photographed as more conventionally professional across hairstyles than his White counterpart. The within-model comparisons remain clean, but cross-model level differences could amplify or dampen the racialized gaps. But overall, the dreadlocks pattern is consistent across genders: white models are penalized far more than Black models, yet Black models are still penalized in absolute terms.

5.2.1 Facial Expression as a Moderator of the Racialized Hair Penalty

Smiling substantially attenuates the racialized curly hair penalty, particularly on interpersonal dimensions related to agreeableness. Among university students, the Black–White gap in the curly hair penalty on agreeableness falls from -0.63 points when the model is pictured with a serious expression to $+0.07$ when smiling—a moderation effect of 0.7 points. On competence, the gap narrows from -0.65 to -0.16. Lay adults exhibit a similar pattern: smiling eliminates the racialized curly gap on all three metrics, with moderation effects of 0.24 (professionalism), 0.24 (competence), and 0.27 (agreeableness). For Black men, the attenuation is concentrated on agreeableness among lay adults, where the racialized gap moves from -0.32 to -0.03.

That smiling narrows the racialized gap does not, however, mean it eliminates the Black model's own penalty. Among students, the Black female model's professionalism penalty for curly hair actually *increases* when she is smiling (-0.46 to -0.60), even as the cross-race gap narrows because smiling simultaneously depresses the White model's ratings. The within-model attenuation is instead concentrated on agreeableness: the Black female model's curly hair penalty falls from -0.27 when serious to an insignificant $+0.03$ when smiling, while the Black male model's near-zero penalty on agreeableness flips to a small positive ($+0.10$). Smiling thus selectively helps Black models on the warmth dimension, leaving the professionalism and competence penalties largely intact.

Figure 7: Smiling attenuates racialized penalties



Each row shows the Black–White racialized curly hair penalty (curly minus straight, Black minus White) under serious (circle) and smiling (square) conditions. Values above arrows report the smile moderation effect (change in B–W gap). $*p < 0.10, **p < 0.05, ***p < 0.01$.

We interpret this pattern as consistent with costly signaling of warmth to offset unfavorable priors (Phelps, 1972). If evaluators hold prior beliefs that natural Black hair signals lower agreeableness, a smile provides a countervailing positive signal that is strong enough to neutralize the penalty—but only on the dimension where the signal is directly relevant. That the same smile does not help on professionalism or competence, and may even hurt, suggests these judgments are anchored to visual markers of conformity to appearance norms rather than to interpersonal cues. Overall, the smile moderation effect we identify imply that the cost of overcoming hair-based discrimination is borne partly through emotional labor: Black individuals must actively signal warmth to achieve the same baseline evaluation that straight-haired or White individuals receive by default.

5.3 Information Interventions

To estimate the treatment effect of our interventions, we interact each hairstyle and race combination in Equation (4.2) with a treatment status and post-treatment indicator (i.e., TREAT ×POST). We calculate these as $\left(\bar{y}^{POST}_{\text{TREAT}} - \bar{y}^{PRE}_{\text{TREAT}}\right) - \left(\bar{y}^{\text{POST}}_{\text{CONTROL}} - \bar{y}^{PRE}_{\text{CONTROL}}\right)$ with Y being the rating differential between curly (or locs) vs. straight/clean-cut hair for the Black model relative to the White model of the same gender. Overall, our treatments do not appear to have substantially or statistically significantly shifted respondents' racialized evaluation of hairstyle. The few individually significant estimates that do emerge are con-

centrated among female locs in the lay adult sample, where the historical video treatment is associated with a modest reduction in the locs penalty for professionalism and competence. However, these are sporadic, do not survive multiple hypothesis testing, and are not echoed across genders or in the student sample. The curly hair penalty, meanwhile, is largely unresponsive to either intervention in both samples. This pattern is consistent with the view that racialized hair perceptions are deeply ingrained and that resistant to low-touch informational nudges. Detailed results are presented in Tables A.5–A.7 in the Appendix.

6 Robustness Checks

Respondent effort. Our pre-analysis plan specified checks for respondent effort based on survey completion time. First, we flag respondents below the 3rd or above the 97th percentile of completion time and re-estimate our main specifications with these observations dropped. Second, we exclude respondents whose completion time fell below one-fifth of the sample median, a threshold designed to identify those who may have clicked through the survey without engaging with the headshot ratings. Our main results are robust to both restrictions. Point estimates remain stable in magnitude and direction, and statistical significance is preserved across the primary outcomes.

Multiple Hypothesis Tests Our design tests for racialized hair penalties across 2 races, 2 genders, 2 hairstyle contrasts (curly vs. straight, locs vs. straight), and 3 outcome measures — yielding a large family of simultaneous tests. To guard against false positives from multiple comparisons, we apply the Romano-Wolf step-down procedure (500 bootstrap replications) within each sample, defining families of 3 outcomes (professionalism, competence, agreeableness) for each gender × hair-type cell. This approach controls the familywise error rate while preserving power by exploiting the joint dependence structure across outcomes. The results are shown in Appendix Figure A.5.

The racialized curly hair penalty for Black women—our central finding—survives correction on all three metrics in the student sample (RW p = 0.002, 0.004, and 0.048 for professionalism, competence, and agreeableness). Among lay adults, this penalty does not survive, with point estimates near zero. The locs result is more pervasive: the racialized gap favoring Black models survives correction across both samples, both genders, and nearly every metric. That the curly penalty is concentrated among evaluators with direct exposure to the professional norms they are reporting on, while the locs finding holds regardless of evaluator background, suggests these hairstyles activate different evaluative channels—curly

hair penalties appear tied to institutional appearance norms that students have absorbed firsthand, whereas the locs gap reflects broader cultural associations legible to any evaluator.

We apply the same correction to the smile moderation results, where the coefficient of interest is the triple interaction of expression, hair type, and photo race. For curly hair, the female agreeableness moderation survives among students (RW p = 0.03), consistent with a warmth signal strong enough to offset the penalty on that dimension but not on professionalism or competence. The moderation is more broadly robust for locs: among students, the female locs moderation survives on competence and agreeableness (RW p = 0.022 and 0.002), and among lay adults it survives on all three metrics (RW p = 0.018, 0.016, and 0.002). The overall pattern reinforces the paper's central finding: the racialized curly hair penalty, and the partial ability of smiling to attenuate it, are most clearly detected by evaluators who are closest to academic institutions, precisely the context where appearance-based discrimination carries the highest stakes.

7 Conclusion

This paper provides direct experimental evidence that professionalism norms penalize natural Black hair in STEM academia, imposing assimilation costs that differ by race, gender, and evaluative dimension. Our findings add to a growing body of work showing that systemic barriers have hindered the US economy via missing patents, foregone output, and greater inequality (Cook & Yang, 2018; Hunt et al., 2012), while diverse perspectives enhance patent novelty (Yang et al., 2022) and broaden research attention to underserved populations (Dossi, 2024). A diverse body of educators in STEM higher education can also attract and retain a diverse student body through role-model effects (Redding, 2019; Baker, 2013). Our results suggest that even as representation grows, more remains to be done on the inclusivity of workplace norms that govern who is perceived as professional, competent, and approachable.[14]

We note two caveats. First, in realistic settings there is often more information available about a person than a photograph: voice, mannerisms, accent, credentials. Our estimates may therefore overstate or understate bias depending on how these additional signals interact

[14]For example, in a survey about the experiences of underrepresented minorities in economics, a Black associate professor reported that she is usually the first Black economist [her students] have ever seen. And that in and of itself has its own issues, like are you qualified? Are you competent?" Moreover, she noted that student evaluations about her and other minority faculty veer off [into] personality… They might like my class [but] they'd be like, 'But I don't like.' And it's usually something really personal or how I dress" (Bayer et al., 2020).

with appearance. Second, each race-by-gender cell is represented by a single model. We take this concern seriously but note that several results are hard to square with model-specific confounds. The smile moderation effect is one example: the racialized agreeableness penalty vanishes when the model smiles, but the professionalism and competence penalties do not. This selective attenuation tracks what statistical discrimination theory predicts (Phelps, 1972) and is not easily generated by any idiosyncrasy of one model's appearance. The dreadlocks results show a consistent cross-gender pattern in which white models are penalized more, consistent with cultural-appropriation priors operating at the group level. Our focus group calibration confirmed that the hairstyle stimuli were rated as reflecting comparable effort across race-matched respondents. We see these results as establishing the existence and structure of racialized hair penalties with enough precision to warrant further work with multiple models per cell.

Despite affecting the daily lives of millions of Black Americans, imposing well-documented monetary, time, and health costs, and bearing directly on legislation now adopted in more than twenty US states, racialized hair bias has remained outside the scope of empirical economics. This paper introduces the first pre-registered, incentive-compatible experimental design for measuring these penalties; documents that they are large, intersectional, and operate through distinct channels for Black women and Black men; uncovers a statistical discrimination mechanism through which smiling attenuates warmth-related penalties but not conformity-based ones; and shows that light-touch information interventions do not reduce them. Together, these findings broaden the welfare analysis of discrimination by showing that its costs extend well beyond the discrete decision points the existing literature has studied, into the ongoing expenditure of money, time, health, and, as we now show, emotional labor required to conform.

References

Agan, A., & Starr, S. (2018). Ban the box, criminal records, and racial discrimination: A field experiment. *Quarterly Journal of Economics*, *133*(1), 191–235. doi: 10.1093/qje/qjx028

Ajzenman, N., Ferman, B., & Sant'Anna, P. C. (2025). Discrimination in the formation of academic networks: A field experiment on #EconTwitter. *American Economic Review: Insights*, *7*(3), 357–375.

Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *Quarterly Journal of Economics*, *115*(3), 715–753. doi: 10.1162/003355300554881

APNews. (2023, September). *A Black student was suspended for his hairstyle. The school says it wasn't discrimination.* Retrieved 2026-03-10, from https://apnews.com/article/hairstyles-dreadlocks-racial-discrimination-crown-act-034a59b9f2652881470dc606b39e5243 (Section: Education)

Arai, M., & Skogman Thoursie, P. (2009). Renouncing personal names: An empirical examination of surname change and earnings. *Journal of Labor Economics*, *27*(1), 127–147.

Arnold, D., Dobbie, W., & Yang, C. S. (2018). Racial bias in bail decisions. *Quarterly Journal of Economics*, *133*(4), 1885–1932. doi: 10.1093/qje/qjy012

Baker, C. N. (2013, December). Social Support and Success in Higher Education: The Influence of On-Campus Support on African American and Latino College Students. *The Urban Review*, *45*(5), 632–650. doi: 10.1007/s11256-013-0234-9

Bartlett, R., Morse, A., Stanton, R., & Wallace, N. (2022). Consumer-lending discrimination in the FinTech era. *Journal of Financial Economics*, *143*(1), 30–56. doi: 10.1016/j.jfineco.2021.05.047

Bayer, A., Hoover, G. A., & Washington, E. (2020, August). How You Can Work to Increase the Presence and Improve the Experience of Black, Latinx, and Native American People in the Economics Profession. *Journal of Economic Perspectives*, *34*(3), 193–219. doi: 10.1257/jep.34.3.193

Bertrand, M., & Mullainathan, S. (2004, September). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, *94*(4), 991–1013. doi: 10.1257/0002828042002561

Bettinger, E. P., & Long, B. T. (2005). Do faculty serve as role models? The impact of instructor gender on female students. *American Economic Review*, *95*(2), 152–157.

Brooks, A. W., Huang, L., Kearney, S. W., & Murray, F. E. (2014, March). Investors prefer entrepreneurial ventures pitched by attractive men. *Proceedings of the National*

Academy of Sciences, *111*(12), 4427–4431. doi: 10.1073/pnas.1321202111

Byrd, A. D., & Tharps, L. L. (2002). *Hair story : untangling the roots of Black hair in America*. New York : St. Martin's Griffin. Retrieved 2025-04-17, from http://archive.org/details/isbn_9780312283223

Callender, V. D., McMichael, A. J., & Cohen, G. F. (2004). Medical and surgical therapies for alopecias in black women. *Dermatologic Therapy*, *17*(2), 164–176. doi: 10.1111/j.1396-0296.2004.04017.x

Cawley, J. (2004). The impact of obesity on wages. *Journal of Human Resources*, *39*(2), 451–474.

Chambers, D. W. (1983). Stereotypic images of the scientist: The draw-a-scientist test. *Science Education*, *67*(2), 255–265. doi: 10.1002/sce.3730670213

Chapman, Y. (2007). *"I am not my hair! Or am I?": Black women's transformative experience in their self perceptions of abroad and at home.* (Thesis).

Cook, L. (2014). Violence and Economic Growth: Evidence from African American Patents, 1870- 1940. *Journal of Economic Growth*, *19*(2), 221–257.

Cook, L., & Yang, Y. (2018). Missing Women and African Americans, Innovation, and Economic Growth. *Working paper.*.

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. In *University of chicago legal forum* (Vol. 1989, pp. 139–167).

Dash, P. (2006, January). Black hair culture, politics and change. *International Journal of Inclusive Education*. Retrieved 2025-04-17, from https://www.tandfonline.com/doi/full/10.1080/13603110500173183 doi: 10.1080/13603110500173183

de Blasio, B., & Menin, J. (2015, December). From Cradle to Cane: The Cost of Being a Female Consumer.

de Quidt, J., Haushofer, J., & Roth, C. (2018, November). Measuring and Bounding Experimenter Demand. *American Economic Review*, *108*(11), 3266–3302. doi: 10.1257/aer.20171330

De Sá Dias, T. C., Baby, A. R., Kaneko, T. M., & Robles Velasco, M. V. (2007). Relaxing/straightening of Afro-ethnic hair: Historical overview. *Journal of Cosmetic Dermatology*, *6*(1), 2–5. doi: 10.1111/j.1473-2165.2007.00294.x

Dossi, G. (2024). *Race and Science.*

Eberle, C. E., Sandler, D. P., Taylor, K. W., & White, A. J. (2020, July). Hair dye and chemical straightener use and breast cancer risk in a large US population of black and white women. *International Journal of Cancer*, *147*(2), 383–391. doi: 10.1002/ijc.32738

França-Stefoni, S. A., Dario, M. F., Sá-Dias, T. C., Bedin, V., de Almeida, A. J., Baby, A. R., & Velasco, M. V. R. (2015). Protein loss in human hair from combination straightening and coloring treatments. *Journal of Cosmetic Dermatology*, *14*(3), 204–208. doi: 10.1111/jocd.12151

Francis, D. V., & Darity, W. A., Jr. (2020). Isolation: An alternative to the "acting white" hypothesis in explaining black under-enrollment in advanced courses. *Journal of Economics, Race, and Policy*, *3*(2), 117–122.

Gallego, F. A., Larroulet Philippi, C., & Repetto, A. (2024). What's behind her smile? Health, looks, and self-esteem. *American Economic Journal: Applied Economics*, *16*(2), 359–388.

Glied, S., & Neidell, M. (2008). *The economic value of teeth* (Working Paper No. 13879). National Bureau of Economic Research.

Hale, G., Regev, T., & Rubinstein, Y. (2023, March). *Do Looks Matter for an Academic Career in Economics?* (SSRN Scholarly Paper No. ID 3805308). Rochester, NY: Social Science Research Network.

Hamermesh, D. S., & Biddle, J. E. (1994). Beauty and the labor market. *American Economic Review*, *84*(5), 1174–1194.

Hester, N., & Gray, K. (2018). For black men, being tall increases threat stereotyping and police stops. *Proceedings of the National Academy of Sciences*, *115*(11), 2711–2715. doi: 10.1073/pnas.1714454115

Hoffman, K. M., Trawalter, S., Axt, J. R., & Oliver, M. N. (2016). Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between Blacks and Whites. *Proceedings of the National Academy of Sciences*, *113*(16), 4296–4301. doi: 10.1073/pnas.1516047113

Hsieh, C.-T., Hurst, E., Jones, C. I., & Klenow, P. J. (2019). The allocation of talent and U.S. economic growth. *Econometrica*, *87*(5), 1439–1474.

Hunt, J., Garant, J.-P., Herman, H., & Munroe, D. J. (2012, March). *Why Don't Women Patent?* (Tech. Rep. No. w17888). National Bureau of Economic Research. doi: 10.3386/w17888

James-Todd, T., Terry, M. B., Rich-Edwards, J., Deierlein, A., & Senie, R. (2011, June). Childhood Hair Product Use and Earlier Age at Menarche in a Racially Diverse Study Population: A Pilot Study. *Annals of epidemiology*, *21*(6), 461–465. doi: 10.1016/j.annepidem.2011.01.009

Johnson, A. M., Godsil, R. D., MacFarlane, J., Aronson, J., Balcetis, E., Barreto, M., ... Okonofua, J. (2017). PERCEPTION INSTITUTE. , 18.

JOY Collective, . (2019). *The CROWN Research Study.*

Kahn, K. B., & Davies, P. G. (2011, July). Differentially dangerous? Phenotypic racial stereotypicality increases implicit bias among ingroup and outgroup members. *Group Processes & Intergroup Relations*, *14*(4), 569–580. doi: 10.1177/1368430210374609

Kang, S. K., DeCelles, K. A., Tilcsik, A., & Jun, S. (2016). Whitened résumés: Race and self-presentation in the labor market. *Administrative Science Quarterly*, *61*(3), 469–502.

Kessler, J. B., Low, C., & Sullivan, C. D. (2019). Incentivized resume rating: Eliciting employer preferences without deception. , *109*(11), 3713–3744. Retrieved 2021-04-14, from http://www.aeaweb.org/articles?id=10.1257/aer.20181714 doi: 10.1257/aer.20181714

Koval, C. Z., & Rosette, A. S. (2021, July). The Natural Hair Bias in Job Recruitment. *Social Psychological and Personality Science*, *12*(5), 741–750. doi: 10.1177/1948550620937937

Levi, P. (1958). *Survival in Auschwitz*. New York: Orion Press. (Originally published as *Se questo è un uomo*, De Silva, 1947)

Morrow, W. L. (1973). *400 Years Without a Comb*. Black Publishers of San Diego.

Opie, T. R., & Phillips, K. W. (2015). Hair penalties: The negative influence of Afrocentric hair on ratings of Black women's dominance and professionalism. *Frontiers in Psychology*, *6*. doi: 10.3389/fpsyg.2015.01311

Paul, M., Zaw, K., Hamilton, D., & Darity, W. A., Jr. (2022). Returns in the labor market: A nuanced view of penalties at the intersection of race and gender in the US. *Feminist Economics*, *28*(2), 1–31.

Persico, N., Postlewaite, A., & Silverman, D. (2004). The effect of adolescent experience on labor market outcomes: The case of height. *Journal of Political Economy*, *112*(5), 1019–1053.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *American Economic Review*, *62*(4), 659–661.

Redding, C. (2019, August). A Teacher Like Me: A Review of the Effect of Student–Teacher Racial/Ethnic Matching on Teacher Perceptions of Students and Student Academic and Behavioral Outcomes. *Review of Educational Research*, *89*(4), 499–535. doi: 10.3102/0034654319853545

Rehavi, M. M., & Starr, S. B. (2014). Racial disparity in federal criminal sentences. *Journal of Political Economy*, *122*(6), 1320–1354. doi: 10.1086/677255

Reuters. (2024, February). Darryl George: Texas judge rules school district can discipline Black teen over hairstyle. *Reuters*. Retrieved 2026-03-10, from https://www.reuters.com/legal/texas-judge-rules-school-district-can

-discipline-black-teen-over-hairstyle-2024-02-22/

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379–440). Academic Press.

Woolford, S. J., Woolford-Hunt, C. J., Sami, A., Blake, N., & Williams, D. R. (2016, July). No sweat: African American adolescent girls' opinions of hairstyle choices and physical activity. *BMC Obesity*, *3*(1), 31. doi: 10.1186/s40608-016-0111-7

Yang, Y., Tian, T. Y., Woodruff, T. K., Jones, B. F., & Uzzi, B. (2022, September). Gender-diverse teams produce more novel and higher-impact scientific ideas. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(36), e2200841119. doi: 10.1073/pnas.2200841119

A Appendix

A.1 Additional Tables and Figures

Table A.1: Summary statistics

Variable	Mean	Std. Dev.	N
Student sample			
STEM major	0.671	0.47	642
Asian	0.391	0.488	481
Black	0.071	0.257	481
Native_American	0.006	0.079	481
White	0.464	0.499	481
Other	0.069	0.253	481
Female	0.714	0.452	500
Male	0.244	0.43	500
Nonbinary	0.042	0.201	500
Democrat	0.709	0.455	419
Republican	0.024	0.153	419
Independent	0.267	0.443	419
Lay adult sample			
Asian	0.049	0.216	2,412
Black	0.127	0.333	2,412
Native American	0.013	0.114	2,412
White	0.830	0.376	2,412
Other	0.012	0.111	2,412
Female	0.501	0.500	2,418
Male	0.497	0.500	2,418
Nonbinary	0.002	0.050	2,418
Democrat	0.490	0.500	2,390
Republican	0.235	0.424	2,390
Independent	0.275	0.447	2,390

Observation counts change due to item non-response.

Table A.2: All students and Lay adults: Professionalism Ratings

	All students		Lay adults	
	Female	Male	Female	Male
Suit	2.358***	2.427***	2.744***	2.798***
	(0.084)	(0.085)	(0.052)	(0.052)
High resolution	0.138**	0.170***	0.195***	0.161***
	(0.059)	(0.059)	(0.030)	(0.030)
Smile	0.198***	0.011	0.379***	0.250***
	(0.048)	(0.052)	(0.030)	(0.029)
Black X Curly	-0.109	0.501***	-0.352***	-0.067
	(0.093)	(0.087)	(0.053)	(0.055)
Black X Locs	-0.063	0.254***	-0.322***	-0.224***
	(0.091)	(0.091)	(0.053)	(0.060)
Black X Straight	0.404***	0.551***	0.038	0.391***
	(0.083)	(0.090)	(0.051)	(0.052)
White X Curly	-0.143	-0.370***	-0.291***	-0.580***
	(0.094)	(0.091)	(0.050)	(0.053)
White X Locs	-1.100***	-1.044***	-1.120***	-1.119***
	(0.096)	(0.098)	(0.058)	(0.058)
White X Straight	.	.	.	.
	.	.	.	.
Constant	5.036***	4.904***	4.583***	4.514***
	(0.083)	(0.092)	(0.050)	(0.052)
Rating differentials				
Black: Curly–straight/clean-cut hair	-0.512***	-0.050	-0.390***	-0.458***
	(0.084)	(0.084)	(0.049)	(0.048)
Black: Locs–straight/clean-cut hair	-0.467***	-0.297***	-0.360***	-0.615***
	(0.077)	(0.082)	(0.049)	(0.054)
Black vs white: Curly–straight/clean-cut	-0.369***	0.320**	-0.098	0.123*
	(0.122)	(0.124)	(0.068)	(0.067)
Black vs white: Locs–straight/clean-cut	0.633***	0.747***	0.760***	0.504***
	(0.118)	(0.126)	(0.074)	(0.072)
Observations	3841	3845	12753	12942

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All models contain subject FE

Table A.3: All students and Lay adults: Competence Ratings

	All students		Lay adults	
	Female	Male	Female	Male
Suit	1.205***	1.114***	1.517***	1.510***
	(0.069)	(0.065)	(0.040)	(0.040)
High resolution	0.123**	0.203***	0.221***	0.234***
	(0.051)	(0.054)	(0.027)	(0.027)
Smile	0.354***	0.103**	0.550***	0.302***
	(0.050)	(0.048)	(0.029)	(0.028)
Black X Curly	0.167*	0.668***	-0.336***	0.068
	(0.089)	(0.094)	(0.050)	(0.052)
Black X Locs	0.134	0.516***	-0.283***	-0.071
	(0.090)	(0.102)	(0.049)	(0.055)
Black X Straight	0.449***	0.721***	-0.031	0.418***
	(0.084)	(0.094)	(0.048)	(0.051)
White X Curly	0.109	-0.082	-0.169***	-0.405***
	(0.089)	(0.091)	(0.048)	(0.048)
White X Locs	-0.853***	-0.742***	-0.907***	-0.938***
	(0.090)	(0.101)	(0.054)	(0.053)
White X Straight	.	.	.	.
	.	.	.	.
Constant	5.734***	5.697***	5.536***	5.543***
	(0.083)	(0.090)	(0.044)	(0.046)
Rating differentials				
Black: Curly–straight/clean-cut hair	-0.281***	-0.053	-0.305***	-0.350***
	(0.075)	(0.078)	(0.044)	(0.044)
Black: Locs–straight/clean-cut hair	-0.315***	-0.206***	-0.252***	-0.489***
	(0.067)	(0.079)	(0.044)	(0.049)
Black vs white: Curly–straight/clean-cut	-0.390***	0.029	-0.136**	0.056
	(0.116)	(0.122)	(0.062)	(0.064)
Black vs white: Locs–straight/clean-cut	0.538***	0.536***	0.654***	0.449***
	(0.110)	(0.123)	(0.067)	(0.068)
Observations	3798	3791	12744	12929

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All models contain subject FE

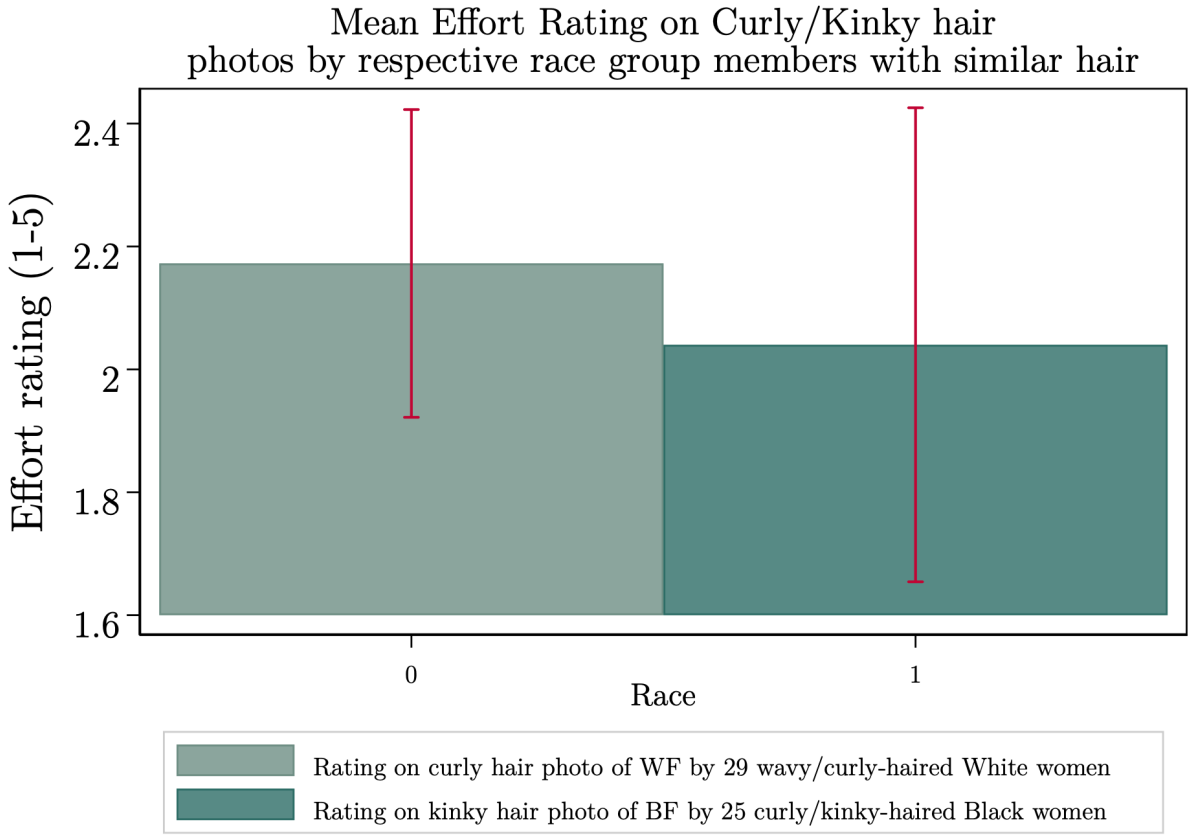
Table A.4: All students and Lay adults: Agreeableness Ratings

	All students		Lay adults	
	Female	Male	Female	Male
Suit	0.430***	0.276***	0.604***	0.606***
	(0.061)	(0.060)	(0.039)	(0.037)
High resolution	-0.118**	0.152***	-0.022	0.229***
	(0.051)	(0.051)	(0.031)	(0.028)
Smile	1.946***	1.689***	2.151***	1.818***
	(0.073)	(0.071)	(0.046)	(0.043)
Black X Curly	0.482***	1.556***	-0.087	0.555***
	(0.094)	(0.091)	(0.058)	(0.056)
Black X Locs	0.445***	1.452***	-0.130**	0.464***
	(0.097)	(0.105)	(0.057)	(0.061)
Black X Straight	0.603***	1.599***	0.077	0.803***
	(0.095)	(0.095)	(0.055)	(0.057)
White X Curly	0.126	0.360***	-0.127**	-0.085
	(0.094)	(0.089)	(0.055)	(0.055)
White X Locs	-0.363***	-0.159	-0.532***	-0.555***
	(0.096)	(0.101)	(0.057)	(0.057)
White X Straight	.	.	.	.
	.	.	.	.
Constant	5.054***	5.024***	5.058***	5.174***
	(0.084)	(0.092)	(0.050)	(0.052)
Rating differentials				
Black: Curly–straight/clean-cut hair	-0.121	-0.042	-0.164***	-0.248***
	(0.083)	(0.077)	(0.051)	(0.046)
Black: Locs–straight/clean-cut hair	-0.158*	-0.147*	-0.207***	-0.339***
	(0.083)	(0.080)	(0.049)	(0.051)
Black vs white: Curly–straight/clean-cut	-0.246*	-0.402***	-0.037	-0.163**
	(0.127)	(0.120)	(0.075)	(0.071)
Black vs white: Locs–straight/clean-cut	0.205	0.012	0.325***	0.216***
	(0.126)	(0.120)	(0.073)	(0.073)
Observations	3831	3831	12751	12941

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All models contain subject FE

A.2 Focus group survey findings

Figure A.1: The curly hair photo of the white female model and the kinky hair photo of the Black female model used in our experiment were rated to represent comparable levels of effort by white and Black women with similar hair textures



A.3 Subjects' self-reported approach to ratings

Figure A.2: Student Respondents' explanations about how they made photo ratings
Note that this wordcloud is made after removing filler words, rating-related metrics, and other common phrases respondents used, in particular ["professionalism professional professionalness competency competence completeness competent agreeableness agreeable agreeability based chose also "]

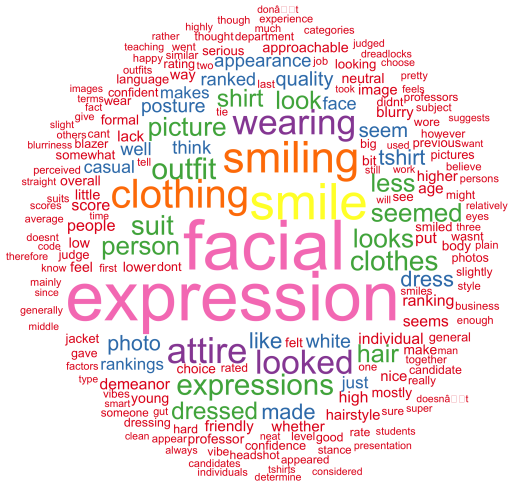


Figure A.3: Lay adult participants' explanations about how they made photo ratings



A.4 Treatment Effects

Table A.5: Treatment effects: Professionalism Ratings

	STEM students						Lay adults					
	No demand		Moderate demand		High demand		No demand		Moderate demand		High demand	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Post-Generic Treatment: Black CURLY penalty	-0.013	-0.001	0.384	0.391	0.286	0.314	0.053	0.055	0.065	0.062	0.353**	0.350*
	(0.203)	(0.202)	(0.257)	(0.253)	(0.305)	(0.300)	(0.104)	(0.104)	(0.159)	(0.159)	(0.179)	(0.179)
Generic TE: Black vs white CURLY penalty	-0.055	0.133	0.331	0.529	0.238	0.451	-0.005	0.077	0.003	0.081	0.295	0.377*
	(0.277)	(0.308)	(0.319)	(0.345)	(0.358)	(0.379)	(0.166)	(0.153)	(0.205)	(0.195)	(0.221)	(0.212)
Post-Historical Treatment: Black CURLY penalty	0.140	0.135	0.306	0.285	0.365	0.331	0.093	0.098	0.206	0.220	0.233	0.239
	(0.186)	(0.186)	(0.323)	(0.322)	(0.276)	(0.275)	(0.108)	(0.108)	(0.171)	(0.171)	(0.171)	(0.171)
Historical TE: Black vs white CURLY penalty	0.098	0.268	0.253	0.424	0.316	0.468	0.036	0.120	0.144	0.239	0.175	0.266
	(0.266)	(0.297)	(0.374)	(0.397)	(0.334)	(0.361)	(0.168)	(0.156)	(0.214)	(0.204)	(0.214)	(0.205)
Post-Generic treatment: Black LOC penalty	-0.013	-0.001	0.384	0.391	0.286	0.314	0.053	0.055	0.065	0.062	0.353**	0.350*
	(0.203)	(0.202)	(0.257)	(0.253)	(0.305)	(0.300)	(0.104)	(0.104)	(0.159)	(0.159)	(0.179)	(0.179)
Generic TE: Black vs white LOC penalty	0.085	0.007	0.126	0.110	0.068	0.057	-0.099	-0.140	0.233	0.190	0.183	0.140
	(0.303)	(0.311)	(0.331)	(0.339)	(0.394)	(0.397)	(0.165)	(0.161)	(0.219)	(0.216)	(0.225)	(0.223)
Post-Historical treatment: Black LOC penalty	-0.015	-0.005	0.509*	0.514*	0.858***	0.849***	0.429***	0.436***	0.272	0.283*	0.060	0.064
	(0.202)	(0.199)	(0.295)	(0.288)	(0.226)	(0.216)	(0.110)	(0.110)	(0.171)	(0.171)	(0.191)	(0.191)
Historical TE: Black vs white LOC penalty	-0.065	-0.176	0.449	0.355	0.801***	0.685**	0.352**	0.317**	0.191	0.162	-0.018	-0.057
	(0.290)	(0.295)	(0.359)	(0.362)	(0.306)	(0.307)	(0.163)	(0.159)	(0.208)	(0.205)	(0.226)	(0.223)
Observations	7273	7299	4832	4858	4933	4959	25191	25189	16864	16862	16854	16852

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All models contain subject FE

Table A.6: Treatment effects: Competence Ratings

	STEM students						Lay adults					
	No demand		Moderate demand		High demand		No demand		Moderate demand		High demand	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Post-Generic Treatment: Black CURLY penalty	0.112	0.156	0.240	0.303	0.215	0.299	0.010	0.013	-0.022	-0.025	0.269	0.270
	(0.170)	(0.174)	(0.234)	(0.240)	(0.205)	(0.215)	(0.108)	(0.108)	(0.135)	(0.135)	(0.169)	(0.169)
Generic TE: Black vs white CURLY penalty	0.141	0.410	0.263	0.557*	0.247	0.553*	0.034	-0.027	-0.002	-0.065	0.292	0.235
	(0.236)	(0.274)	(0.286)	(0.322)	(0.263)	(0.300)	(0.154)	(0.147)	(0.174)	(0.167)	(0.201)	(0.196)
Post-Historical Treatment: Black CURLY penalty	0.056	0.073	-0.004	0.024	0.365	0.370	0.021	0.019	0.237*	0.237*	0.080	0.079
	(0.190)	(0.191)	(0.256)	(0.255)	(0.244)	(0.242)	(0.107)	(0.107)	(0.137)	(0.136)	(0.151)	(0.151)
Historical TE: Black vs white CURLY penalty	0.086	0.327	0.019	0.278	0.398	0.624*	0.045	-0.021	0.258	0.197	0.104	0.044
	(0.252)	(0.285)	(0.304)	(0.330)	(0.292)	(0.322)	(0.153)	(0.146)	(0.176)	(0.168)	(0.186)	(0.181)
Post-Generic treatment: Black LOC penalty	0.112	0.156	0.240	0.303	0.215	0.299	0.010	0.013	-0.022	-0.025	0.269	0.270
	(0.170)	(0.174)	(0.234)	(0.240)	(0.205)	(0.215)	(0.108)	(0.108)	(0.135)	(0.135)	(0.169)	(0.169)
Generic TE: Black vs white LOC penalty	0.358	0.450	0.074	0.181	-0.104	0.035	-0.084	-0.197	0.316	0.202	0.191	0.080
	(0.254)	(0.291)	(0.309)	(0.342)	(0.318)	(0.345)	(0.148)	(0.147)	(0.206)	(0.205)	(0.195)	(0.194)
Post-Historical treatment: Black LOC penalty	-0.095	-0.086	0.046	0.069	0.681***	0.654***	0.258**	0.255**	0.247*	0.247*	-0.022	-0.028
	(0.167)	(0.167)	(0.273)	(0.271)	(0.219)	(0.216)	(0.101)	(0.101)	(0.147)	(0.146)	(0.162)	(0.162)
Historical TE: Black vs white LOC penalty	-0.164	-0.102	-0.033	0.060	0.609**	0.636**	0.287**	0.171	0.273	0.162	0.004	-0.113
	(0.245)	(0.281)	(0.326)	(0.353)	(0.281)	(0.313)	(0.145)	(0.144)	(0.180)	(0.179)	(0.192)	(0.192)
Observations	7152	7185	4717	4750	4844	4877	25170	25169	16846	16845	16832	16831

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All models contain subject FE

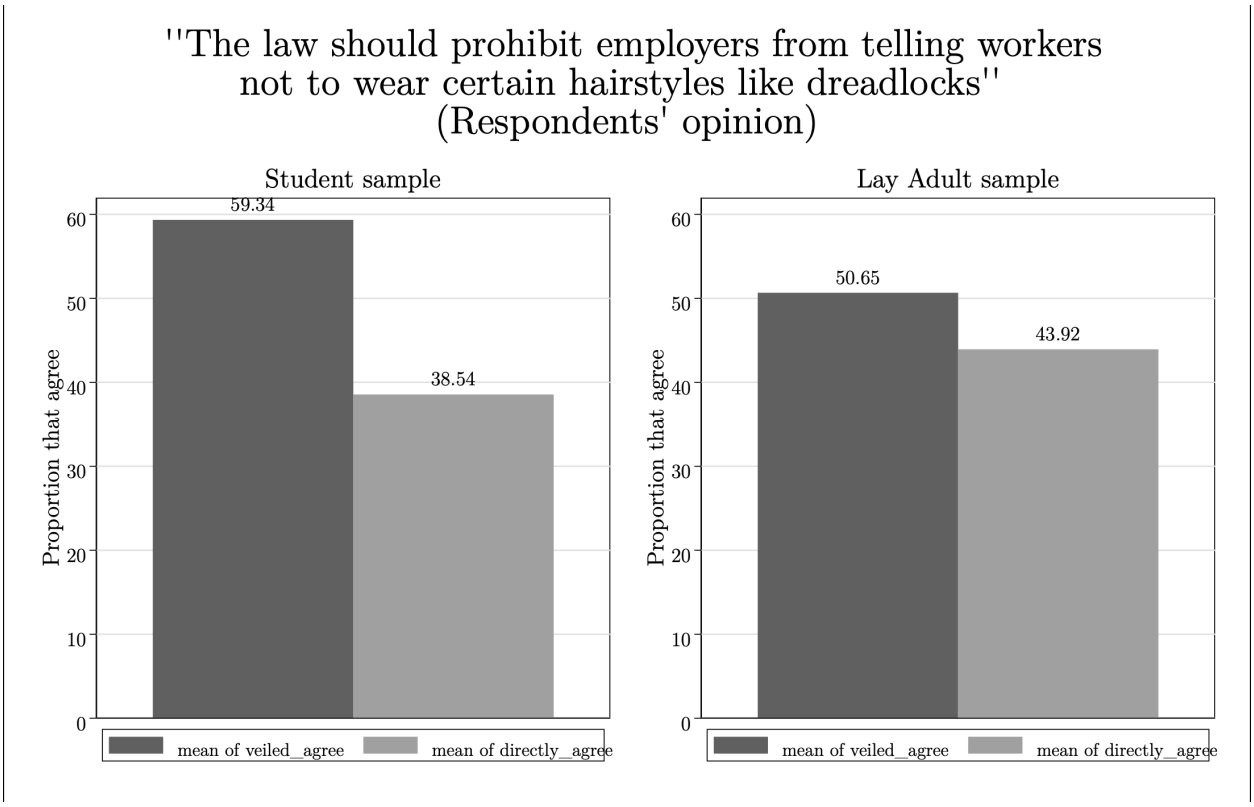
Table A.7: Treatment effects: Agreeableness Ratings

	STEM students						Lay adults					
	No demand		Moderate demand		High demand		No demand		Moderate demand		High demand	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Post-Generic Treatment: Black CURLY penalty	0.002	0.059	0.134	0.273	0.073	0.216	0.049	0.053	-0.017	-0.019	-0.120	-0.112
	(0.183)	(0.182)	(0.290)	(0.288)	(0.266)	(0.263)	(0.127)	(0.126)	(0.186)	(0.185)	(0.191)	(0.189)
Generic TE: Black vs white CURLY penalty	0.019	0.133	0.142	0.343	0.100	0.297	0.077	0.121	0.004	0.046	-0.091	-0.036
	(0.279)	(0.282)	(0.360)	(0.360)	(0.340)	(0.341)	(0.177)	(0.167)	(0.223)	(0.216)	(0.227)	(0.218)
Post-Historical Treatment: Black CURLY penalty	-0.068	-0.055	-0.349	-0.370	0.121	0.093	0.021	0.018	0.226	0.224	0.065	0.071
	(0.188)	(0.191)	(0.288)	(0.294)	(0.273)	(0.281)	(0.123)	(0.122)	(0.176)	(0.175)	(0.164)	(0.163)
Historical TE: Black vs white CURLY penalty	-0.051	0.019	-0.341	-0.300	0.148	0.174	0.048	0.087	0.247	0.288	0.094	0.147
	(0.283)	(0.288)	(0.357)	(0.364)	(0.345)	(0.354)	(0.173)	(0.165)	(0.215)	(0.207)	(0.204)	(0.196)
Post-Generic treatment: Black LOC penalty	0.002	0.059	0.134	0.273	0.073	0.216	0.049	0.053	-0.017	-0.019	-0.120	-0.112
	(0.183)	(0.182)	(0.290)	(0.288)	(0.266)	(0.263)	(0.127)	(0.126)	(0.186)	(0.185)	(0.191)	(0.189)
Generic TE: Black vs white LOC penalty	0.073	0.202	-0.502	-0.301	-0.247	-0.034	0.016	-0.042	0.345	0.288	0.124	0.072
	(0.277)	(0.306)	(0.376)	(0.395)	(0.333)	(0.358)	(0.169)	(0.166)	(0.232)	(0.229)	(0.219)	(0.216)
Post-Historical treatment: Black LOC penalty	-0.303	-0.284	-0.247	-0.218	0.856***	0.828***	0.146	0.138	0.016	0.017	0.017	0.014
	(0.205)	(0.201)	(0.316)	(0.310)	(0.239)	(0.236)	(0.120)	(0.119)	(0.185)	(0.183)	(0.180)	(0.179)
Historical TE: Black vs white LOC penalty	-0.461*	-0.364	-0.408	-0.290	0.688**	0.741**	0.237	0.173	0.104	0.048	0.106	0.050
	(0.277)	(0.306)	(0.366)	(0.385)	(0.305)	(0.333)	(0.168)	(0.164)	(0.220)	(0.216)	(0.215)	(0.212)
Observations	7238	7264	4829	4855	4923	4949	25184	25183	16868	16867	16847	16846

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All models contain subject FE

A.5 List Experiment Results

Figure A.4: Our list experiment suggests that subjects underreport support for legislation like the CROWN Act, which would ban employers from instituting grooming policies that prohibit certain hairstyles



A.6 Multiple Hypothesis Testing Robustness Checks

Figure A.5: Racialized Hair Penalty estimates with Romano-Wolf Correction

