

## Leveraging Genomic Data to Document Within-Race Attractiveness Penalties Among Black Americans

Beza Taddess<sup>1\*</sup>

Luyin Zhang<sup>2\*</sup>

Sam Trejo<sup>1,2†</sup>

1. Department of Sociology, Princeton University
2. Office of Population Research, Princeton University

\*Denotes equal work

†Send correspondence to [samtrejo@princeton.edu](mailto:samtrejo@princeton.edu)

### Abstract

In recent years, scholars of racial inequality have increasingly sought to move beyond simply quantifying *discrete* racial disparities and instead measure social stratification as a function of *continuous* racialized characteristics which vary both within and between racial groups. In this paper, we draw on a sample of genotyped respondents from the Add Health study and construct genetic similarity proportions, individual-level measures that correlate with racialized physical features that vary across the expansive family tree of humanity (skin tone, facial structure, hair texture, etc.). We then investigate the relationship between these proportions and interviewer-rated physical attractiveness among self-identified Black Americans ( $N=2,087$ ). Our findings highlight the existence of substantial attractiveness penalties related to having higher levels of Sub-Saharan African (as opposed to European) genetic similarity.

## Introduction

A large body of empirical research documents average disparities between White and Black Americans in a host of valued outcomes, ranging from childhood educational opportunity (1) to success in the dating market (2) to life expectancy (3). In recent years, however, scholars of racial inequality have increasingly sought to move beyond simply quantifying *discrete* disparities – for instance, the average difference in a variable of interest between two racial groups (e.g., the Black-White achievement gap) – to instead measure social stratification as a function of *continuous* racialized characteristics (4, 5),\* which vary both within and between racial groups (e.g., skin tone (6)). An overemphasis on discrete racial categories can serve to obscure meaningful within-race inequalities, whereas shifting the focus to various racialized characteristics highlights that there exists variation in how individuals with the same racial identity experience race and its consequences. Though this line of inquiry offers much promise, studies of racialized characteristics have, thus far, faced two key methodological challenges: [i] it is difficult to know *a priori* which specific individual-level physical and social characteristics have become imbued with racial meaning and stigma, and [ii] a relatively limited number of racialized characteristics (often measured with substantial error) are currently available in existing nationally representative data sources.† To address these challenges, and in light of the increased availability of molecular genetic data, we propose a new tool for the empirical exploration of racial stratification: genetic similarity proportions.

Genetic similarity proportions (GSPs (9)), also known as genetic ancestry proportions, have a wide range of applications in human genomics, including in the study of gene-ancestry interaction effects (10, 11) and in the analysis of historical human admixture events (12). (Genetic admixture refers to genomic mixing of previously isolated populations.) However, recent social scientific research has demonstrated that GSPs may also be used to index variation in racialized physical features in the United States today (13). GSPs, which are readily constructed from genotype and/or sequence data, provide estimates of the fraction of a person’s DNA that is categorized into various (often geographically defined) genomic reference populations (14, 15). Notably, GSPs vary continuously among admixed populations (such as Black and Hispanic Americans), are fixed at birth, and are estimated with little measurement error.‡ Moreover, the use of GSPs does not require researchers to prespecify which specific racialized characteristics might be relevant to a given social process; because racial ideologies are socially constructed using

---

\* We use the term racial group to describe the intersection of traditional racial categories (e.g., Black, White, Native American, Asian/Pacific Islander, etc.) and Hispanicity. Note, these racial/ethnic groups – henceforth simply ‘racial groups’ – are constructed to be mutually exclusive (see *Materials and Methods*). For the remainder of the paper, when we use the terms ‘Black Americans’ and ‘White Americans’, we are describing the group of individuals who self-identify as Black or White, respectively, and who do *not* also self-identify as Hispanic.

† Despite their wide use in the study of colorism, existing measures of skin tone contain a substantial amount of measurement error (7, 8), which serves to attenuate estimates of within-race stratification towards zero.

‡ In principal component analysis and related methods (including global GSP estimation), measurement error arises when the estimated SNP-level weights for a given component meaningfully vary across models fit on different finite samples of the same underlying population. Past work has shown that the amount of measurement error in the four genetic similarity proportions used in this study, which correspond to the first four axes of genetic variation in the Add Health data, is relatively small (13). However, as the number of GSPs estimated (i.e.,  $K$ ) increases, so does that amount of measurement error (16). So, while our approach allows us to accurately distinguish between Sub-Saharan African and European genetic similarity, there is no guarantee that the same methods and data could accurately distinguish between Northern and Southern European genetic similarity (or even more granular ancestral differences).

physical features that vary across the expansive family tree of humanity (17, 18), variation in GSPs will tend to capture phenotypic variation in these characteristics.<sup>§</sup>

In this paper, we use GSPs to explore social stratification in externally rated physical attractiveness. Historically, sex-specific selection on attractiveness may have played an important role in the evolution of humans (20), and – in the modern era – there are significant social advantages to being viewed as physically attractive (21, 22). Individuals perceived to be attractive tend to marry earlier (23), earn more (24–28), live longer (29), and report higher subjective well-being than their less attractive counterparts (30). Notably, exactly which physical features are viewed as attractive is subjective and culturally dependent, with past work documenting the existence of variation in preferences across the globe (31). A recent analysis of dating app data found that Black Americans were the only racial group where both men and women were systematically viewed as less desirable than their White counterparts (2). We draw on a sample of genotyped Americans from the National Longitudinal Study of Adolescent to Adult Health (Add Health) and construct GSPs linked to four present-day reference populations (32, 33): Sub-Saharan Africa (Yoruba in Nigeria;  $P^{AFR}$ ), Europe (Europeans in Utah;  $P^{EUR}$ ), East Asia (Han Chinese in Beijing;  $P^{EAS}$ ), and Indigenous America (Pima & Maya in Mexico;  $P^{IAM}$ ). Then, we empirically test whether – among members of a single self-identified racial group – individuals with certain GSPs are systematically viewed as more attractive than others. While the bulk of our analyses focus on within-race variation in GSPs among Black Americans, the empirical results we present nonetheless inform understandings of average between-race disparities.

Our study makes numerous important contributions. First, we provide robust empirical evidence of racialized stratification in attractiveness among Black Americans, with individuals with higher amounts of Sub-Saharan African (as opposed to European) genetic similarity receiving the lowest ratings; this result implies that racial attractiveness disparities arise – not merely due to stigmatization of individuals based on their perceived race – but also due to a broader societal stigmatization of the physical features associated with Blackness. Second, our results point to methodological issues with survey-based physical attractiveness ratings and suggest such measures likely substantially understate the true magnitude of Black-White attractiveness disparities. Third, we demonstrate that GSPs represent a new tool for social scientists interested in studying racial inequality. Finally, our results have important implications for the interpretation of genome-wide association study results for complex traits, particularly among admixed populations.

## Results

We begin by examining racial disparities in interviewer-rated attractiveness using the Add Health data. In Waves I, II, III, and IV of the Add Health study, in-home interviewers were asked to rate the physical attractiveness of each respondent using a Likert scale measure with five categories: ‘very unattractive’, ‘unattractive’, ‘about average’, ‘attractive’, ‘very attractive’. Figure 1 contains four bar charts, each of which displays the average attractiveness ratings of four self-identified racial groups: White Americans, Black Americans, Asian/Pacific Islander Americans, and Hispanic Americans. Panels A and B display each race’s average attractiveness score, whereas Panels C and D display the fraction of respondents of a given race rated in the highest attractiveness

---

<sup>§</sup> Social scientists use the term racial formation to describe the process through which racial meanings extend to previously unclassified relationships, practices, or groups (19). Through racial formation, physical and social characteristics are transformed into markers of racial group membership that influence social perceptions and experiences.

category. While Panels A and C display unadjusted values, Panels C and D display values which are residualized on interviewer fixed effects. Although White Americans tend to have the highest attractiveness ratings, we observe only small average differences across racial groups – in line with previous work using interviewer-reported survey measures (28). For instance, the unadjusted difference in the attractiveness score of White and Black Americans is just 0.095 standard deviations (SD). These modest results stand in stark contrast to the substantial attractiveness disparities observed in the large-scale revealed preference analyses of dating app data (2);\*\* fortunately, the results from our subsequent GSP analysis offer some insight into this apparent empirical puzzle.

Next, we turn to our within-race analysis of Black Americans. Figure S1 displays ternary plots of respondent GSPs by racial group; consistent with prior work (13), the genomes of most Black Americans in our sample are a mix of Sub-Saharan African ( $\bar{P}^{AFR}=0.80$ ) and European genetic similarity ( $\bar{P}^{EUR}=0.18$ ).†† Table 1 presents results from linear and logistic regressions of a respondent’s rated attractiveness on their GSPs. Among Black Americans, a 10 percentage point (pp) increase in Sub-Saharan African genetic similarity is associated with a 0.11 SD ( $p<0.001$ ) decrease in attractiveness score and a 3.6 pp (25%;  $p<0.001$ ) decrease in the probability of being rated as ‘very attractive’. The magnitude of these associations slightly attenuates after controlling for various physical features (skin tone, hair color, and eye color), racial classification, family and neighborhood socioeconomic status, but remains highly statistically significant.‡‡ According to the estimates from Model 1, the difference in average attractiveness score between an individual at the 5<sup>th</sup> and 95<sup>th</sup> percentiles, respectively, of the Black distribution of Sub-Saharan African genetic similarity ( $P^{AFR}=0.46$ ;  $P^{AFR}=0.93$ ) is equal to -0.52 SD; notably, this number is over *five times as large* as the average Black-White attractiveness disparity presented in Figure 1.

Figure 2 presents a pair of binned scatterplots that graphically display the relationship between Sub-Saharan African genetic similarity and attractiveness among Black Americans; the Y-axis in Panel A displays attractiveness scores, whereas the Y-axis in Panel B displays the fraction of respondents who are rated as *at least* a given attractiveness category (i.e., ‘about average’, ‘attractive’, and ‘very attractive’). Note, these plots contain both the observed average attractiveness ratings of White Americans (solid blue markers) and the predicted average attractiveness scores of White Americans (hollow blue markers); the predicted values are useful for gaining a sense of how attractive we would expect White Americans to be on average given the positive relationship between European (rather than African) genetic similar observed among

---

\*\* For instance, Monk, Esposito, and Lee 2021(28) use interviewer-rated attractiveness measures and find a Black-White disparity of just 0.08 SD. On the other hand, Bruch and Newman 2018 (2) apply a link analysis algorithm to dating app data and find White Americans are, on average, ranked at the 53<sup>rd</sup> percentile of the attractiveness distribution, whereas Black Americans are ranked at the 40<sup>th</sup> percentile. While it is difficult to directly compare standard deviation units to percentiles, a 13 percentile change from the mean of a standard normal distribution is equal to 0.33 SD. Results derived from dating app behavior, however, likely partly reflect sorting on dimensions of desirability other than physical attractiveness (like, for instance, educational attainment and socioeconomic status) and may also suffer from bias from differential selection into app usage across race.

†† The genomes of Hispanic Americans in our sample are, on the other hand, a combination of European ( $\bar{P}^{EUR}=0.58$ ), Indigenous American ( $\bar{P}^{EUR}=0.29$ ), and Sub-Saharan African genetic similarity ( $\bar{P}^{AFR}=0.11$ ). See Figure S2 for histograms of Sub-Saharan African, European, and Indigenous American genetic similarity among Black and Hispanic Americans.

‡‡ Note, while the GSPs are correlated with skin tone, hair color, and eye color, we include these three measures as covariates in order to test the extent to which the GSPs explain variation in attractiveness *net* of currently available physical feature measures. Figure S8 graphically illustrates the fact that  $P^{AFR}$  explains variation in attractiveness among Black Americans of the same interviewer-rated skin tone.

Black Americans. In Panel A, it can be plainly seen that there exist large average differences between the observed attractiveness scores of White Americans (solid blue marker) and the predicted attractiveness scores of White Americans (hollow blue marker); thus, while the observed Black-White disparity is only -0.095 SD, extrapolation of the relationship between GSP and attractiveness among Black Americans yields a predicted Black-White disparity that is far greater<sup>§§</sup>. This, as we discuss below, casts some doubt on the extent to which interviewer-reported attractiveness ratings accurately measure underlying perceptions of attractiveness. See Table S4 and Figure S4 for an analogous set of results regarding the relationship between respondent GSPs and attractiveness among Hispanic Americans; in general, a far more limited set of associations exists in our Hispanic sample, all of which fall to statistical insignificance after the inclusion of physical and social covariates.

Importantly, there exists meaningful heterogeneity regarding exactly how intensely different Black respondents are penalized. In Panel A of Figure 3, we decompose our results regarding the relationship between Sub-Saharan African genetic similarity and attractiveness across six dimensions: interviewer race, gender, and age and respondent gender, age, and census region. After implementing a 10% Benjamini-Hochberg false discovery rate correction (34), only two of these dimensions – interviewer race and respondent census region – capture statistically significant variation in the association between Sub-Saharan African genetic similarity and attractiveness ratings. Black interviewers impose stronger penalties onto Black respondents than White interviewers do; while among Black interviewers a 10 percentage point (pp) increase in Sub-Saharan African genetic similarity is associated with an 0.159 SD decrease in attractiveness score and a 5.3 pp increase in the probability of being rated as ‘very attractive’, these relationships attenuate to 0.064 SD and 1.7 pp among White interviewers (but nonetheless remain statistically significant). Similarly, the magnitude of Sub-Saharan African genetic similarity attractiveness penalties is the largest in the South, smaller in the Northeast, and smaller still in the Midwest; finally, in the West, there exists no statistically detectable attractiveness penalty.

Finally, we conduct a simple machine learning analysis to compare the predictive performance of Sub-Saharan African genetic similarity to a range of interviewer-reported survey measures related to race and appearance. In particular, we train prediction models using elastic net regularization, with attractiveness score and ‘very attractive’ as the target variables. Panel A of Figure 3 displays Shapley additive explanations (SHAP) values derived from these two penalized regression models (35); SHAP values decompose a model’s overall predictions into feature-level contributions, with higher average values indicating the most important features. Thus, this approach allows us to quantify the explanatory power of Sub-Saharan African genetic similarity compared to the other variables in our model (skin tone, racial classification, hair color, and eye color). For both the attractiveness score and ‘very attractive’ prediction models, Sub-Saharan African genetic similarity yields the greatest SHAP values of all the features. These findings suggest that, among Black Americans, genetic ancestry captures a broader range of racialized features related to attractiveness evaluations than conventional survey-based measures.

## Discussion

Until now, it has been difficult to distinguish between two competing explanations of White-Black attractiveness disparities: do such gaps arise as a result of observers discretely classifying subjects (and then penalizing those they deem to be Black), or do they instead arise

---

<sup>§§</sup> For instance, when using bivariate linear extrapolation, the predicted Black-White disparity in attractiveness score is over *eight times as large* as the measured Black-White disparity.

from a broader stigmatization of a constellation of physical features associated with Blackness?\*\*\* Our empirical results provide robust evidence of the existence of substantial attractiveness stratification among Black Americans. Notably, interviewers did not have any direct knowledge of a respondent's GSPs – in fact, because the attractiveness ratings were collected between 1996 and 2008, interviewers had likely never even heard of a genetic ancestry test – but they nonetheless systematically penalized those with higher levels of Sub-Saharan African genetic similarity (36). These results imply that, while discrete classification biases may well exist, the collective stigmatization of physical features associated with Blackness is a key part of the story. Importantly, our results do not suggest that there is an objective or universal notion of attractiveness, or that a person's perceived attractiveness is an inevitable result of their biology. Instead, our findings indicate that contemporary American society has developed a conceptualization of beauty which devalues physical features more common in individuals with higher amounts of Sub-Saharan African (rather than European) genetic similarity.

Why do we observe larger attractiveness penalties among Black interviewers and for respondents living in certain geographic regions? One potential explanation of these patterns is the varying degree of social exposure to Black individuals. Most Black Americans live in the South (56%), followed by the Northeast (17%) and Midwest (17%), and finally the West (10%); notably, the South also contains every majority-Black U.S. county (37, 38). And, due to segregated schools, neighborhoods, and social networks, Black Americans tend to – compared to individuals of other races – have a greater number of social interactions and relationships with Black people (39). It may be that the racialized physical features correlated with Sub-Saharan African genetic similarity are especially stigmatized, net of racial classification, in social contexts with a sufficient number of Black individuals. Alternatively, individuals who are exposed to many Black social peers may simply become more accurate at, either implicitly or explicitly, distinguishing between Black individuals with varying amounts of Sub-Saharan African genetic similarity. In addition, it is worth noting that the more extreme slope estimates among Black interviewers appear to result – at least in part – from the fact that they are more likely than interviewers of other races to rate Black respondents with low Sub-Saharan African genetic similarity as 'attractive' or 'very attractive' (see Figure S5). Finally, the historical legacy of slavery in the South, as well as elevated levels of contemporary racial animus (40), could also play an important role.

Our findings also help reconcile conflicting results regarding (small) Black-White attractiveness disparities from studies using survey-based attractiveness ratings, like those in Add Health, and the (large) disparities observed in studies using real-world behavior on dating apps. In particular, our within-race analysis revealed strong penalties of Sub-Saharan African genetic similarity, which, combined with the very large difference in Sub-Saharan African genetic similarity between Black and White Americans, would lead us to expect Black-White attractiveness disparities far larger in magnitude than we, in fact, observe. This suggests an important limitation of survey-based attractiveness measures: they may suffer from social desirability bias (41, 42). That is, reviewers adjust their responses, either consciously or subconsciously, so as not to produce large average differences in attractiveness ratings across

---

\*\*\*See Figure S9 for a directed acyclic graph depicting this theoretical process. A key challenge in answering this question is that the racial boundary which separates individuals as either Black or another race is generally quite clear; thus, there is generally very little variation in the racial classification by others among self-identified Black individuals. In Add Health, for instance, Add Health interviewers classified more than 98% of self-identified Black respondents as Black. By exploring within-group attractiveness variation among a population with little variation in racial classification, our study highlights the existence of a direct relationship between genetically influenced (racialized) physical features and perceived attractiveness.

racial groups. Nonetheless, this correction appears to be relatively coarse; while interviewers are able to almost entirely eliminate unsavory average differences in attractiveness between racial groups, their responses nonetheless still show them penalizing Black Americans with high amounts of Sub-Saharan African genetic similarity. If survey-based measures suffer from this form of social desirability bias, then subsequent analyses will substantially understate the true magnitude of Black-White attractiveness disparities. In addition, these findings illustrate how discrete correction procedures fail to successfully mitigate stratification based on a continuous set of underlying dimensions.

Our results also have substantial methodological implications. The superior predictive power of GSPs compared to other measures of physical features (e.g., skin tone), as well as their desirable measurement properties, highlight their strength as a new tool for social scientists interested in measuring and studying processes of racialization. Notably, we argue GSPs represent a complement to – and not a replacement of – existing survey strategies for quantifying the many continuous dimensions of race. In addition, many studies that lack measures of racialized physical features and racial classification (for instance, the Health and Retirement Study (43) nonetheless have collected genetic data, meaning the use of GSPs may aid the continuing expansion of scholarship on continuous (rather than discrete ) racial stratification.

Finally, our findings add important nuance to the interpretation of genome-wide association study results for complex traits. Consider, for instance, that a given genetic variant is found to have a statistically significant causal effect on a psychiatric trait, like anxiety or depression (44); while this genetic effect might operate strictly through biologically proximal processes within the body – for example, the regulation of neurotransmitters in the brain – it may also operate through biologically distal processes outside of the body: for instance, the social stigmatization of certain individuals based on their genetically influenced physical features (18). In a similar vein, our results highlight that many popular genomic methods that utilize GSPs to identify epistatic effects (10, 11) may, in fact, simply be identifying genetic heterogeneity related to social experiences of racialization .

## **Materials and Methods**

### ***Add Health***

The National Longitudinal Study of Adolescent to Adult Health (Add Health) is a longitudinal survey of a nationally representative sample of 20,745 middle and high school students in the United States (45). The initial wave was fielded in the 1994-1995 school year (Wave I), followed by four additional waves of in-home interviews in 1996 (Wave II), 2001-2002 (Wave III), 2008 (Wave IV), and 2016-2018 (Wave V). At each wave, a rich set of sociodemographic, behavioral, psychosocial, familial, and contextual information was collected. Interviewers each surveyed an average of 25 respondents at each wave, and there was an average of approximately 4 days between each interview (see Figure S10). Approximately 80% of the respondents who participated in Wave IV provided saliva samples and were genotyped using two Illumina platforms – Illumina Human Omni1-Quad BeadChip and Illumina Human Omni-2.5 Quad BeadChip. Rigorous quality control procedures were applied by the Add Health staff at both the SNP-level and the individual-level; in particular, SNPs with call rates < 90%, minor allele frequency < 0.5%, and deviations from Hardy-Weinberg equilibrium ( $p < 5 \times 10^{-5}$ ) were removed, and individuals with call rates < 90% and genetic sex discordance were removed. The final genotype data cover 609,130 SNPs for 9,974 individuals. Table S1 provides detailed descriptive statistics for our Add Health analytic sample, and Table S2 provides descriptive statistics of the Add Health interviewers. The Add Health interviewers of Black respondents tend to be highly educated (36% some college; 52% B.A. or above), and the majority identify as either White (56%) or Black (38%).

### ***Genetic Similarity Proportions***

We use supervised ADMIXTURE (46, 47) with  $K = 4$  to estimate global genetic similarity proportions (GSPs) for each genotyped Add Health respondent. Our reference panels are comprised of the following unrelated individuals from HapMap 3 (32) and the Human Genome Diversity Project (33): 83 Yoruba (Nigeria;  $P^{AFR}$ ), 36 Northern/Western Europeans (Utah;  $P^{EUR}$ ), 137 Han Chinese (Beijing;  $P^{EAS}$ ), and 34 Pima/Maya (Mexico;  $P^{IAM}$ ). We restrict to autosomal SNPs that are present in the Add Health genotype data and our reference panels. After implementing linkage disequilibrium pruning (with a window size of 200kb, a step size of 25, and an  $R^2$  of 0.4) in PLINK1.9 (48), we retain 279,464 SNPs. Because the Add Health data contain siblings and half-siblings, we remove a random respondent from each pair to create a subsample of 9,166 unrelated respondents; all estimates were then projected for the remaining 808 genotyped respondents. Note,  $P^{AFR}$ ,  $P^{EUR}$ ,  $P^{EAS}$ , and  $P^{IAM}$  mechanically sum to one for each individual. Unsupervised ADMIXTURE analysis recovers almost identical GSPs (see Figures S2 and S3 from Zhang and Trejo 2025; (13)) as our supervised ADMIXTURE estimates. Moreover, results from local genetic similarity estimation software, when summed across the genome, are highly comparable to those from ADMIXTURE (see Figures S5 from Zhang and Trejo 2025). Finally, past research using highly similar methods has shown that, in U.S. samples, the resulting GSPs closely correspond to the global information provided by popular genetic ancestry tests (see Figure S14 from Bryc et al. 2015 (49) for a comparison with 23andMe). We focus our main text and supplementary analyses on Black Americans and Hispanic Americans due to the fact that these two racial groups – in contrast to White Americans – exhibit substantial within-group variation in GSPs. In addition, our sample size of Asian Americans is simply too small for rigorous subgroup analysis.

### ***Survey Measures***

**Physical attractiveness:** At Wave I (age 12-21), Waves II (age 13-22), Waves III (age 18-26), and Waves IV (age 24-32), Add Health interviewers were asked to rate the physical attractiveness of each respondent using a Likert scale with five categories (1='very unattractive', 2='unattractive', 3='about average', 4='attractive', and 5='very attractive'). To aid in the interpretation of our attractiveness score variable, we standardize it using the weighted mean and standard deviation ( $\mu = 3.44, \sigma = 0.82$ ) of the full Add Health sample.

**Racial identity:** We construct a categorical variable of the single racial identity that best describes a respondent's racial background using information collected in Wave III. (In the rare event that an individual's Wave III racial identity information is missing, we supplement with racial identity information collected at Waves I and V.) The Wave III Add Health self-reported race measure includes four categories: White, Black, Native American, and Asian/Pacific Islander. We intersect categorical responses from the racial identity question with binary responses to a Hispanicity question to create the following five mutually exclusive racial categories: Non-Hispanic White, Non-Hispanic Black, Non-Hispanic Native American, Non-Hispanic Asian/Pacific Islander, and Hispanic American. Thus, when we refer to 'Black Americans' and 'White Americans', we are describing the group of individuals who

self-identify as Black or White and who do *not* also self-identify as Hispanic. Due to the group's limited sample size ( $N=68$ ), we do not present results regarding self-identified Non-Hispanic Native Americans.

**Physical features:** We rely on interviewer-reported measures from Wave III to create categorical measures of physical features, including skin tone (black, dark brown, medium brown, light brown, and white), hair color (no hair, black, brown, blond, red, grey, and other), and eye color (black, brown, hazel, blue, green, and other).

**Racial classification:** Interviewer's racial classification of the respondent, coded based on their observation alone, was collected in Waves I, III, and IV, and contains the following five categories: White, Black, Native American, Asian/Pacific Islander, and Other.

**Socioeconomic variables:** For our childhood socioeconomic status variable, we use the first principal component of Wave I parental education, parental occupation, household income, and household receipt of public assistance (constructed by Belsky et al. 2018 (50)). For our neighborhood socioeconomic disadvantage variable, individuals are matched to the American Community Survey data of the census tract of their Wave I home address; deciles of five tract-level variables – proportions of female-headed households, individuals living below the poverty threshold, individuals receiving public assistance, adults with less than a high school education, and adults who were unemployed – were totaled for each tract and then standardized within-sample (see Belsky et al. 2019 (51) for more details).

### Regression Analysis

We use multivariate regression analysis to assess the relationship between an individual's genetic similarity proportions and their attractiveness rating. First, we treat attractiveness as a continuous variable and fit the following linear regression model for individual  $i$  observed at age  $j$  by interviewer  $k$ :

$$attractiveness_{ijk} = \delta_j + \gamma_k + \sum_{n=1}^4 (\alpha_n race_{ijk}^n) + \sum_{n=1}^4 \sum_{m=1}^3 (\beta_{n,m} race_{ijk}^n \times P_{ijk}^m) + \sum_{n=1}^4 (race_{ijk}^n \times \mathbf{W}_{ijk} \Phi) + \varepsilon_{ijk}$$

Eq. 1

where  $attractiveness_{ijk}$  represents the interviewer-rated attractiveness score of individual  $i$  at age  $j$  by interviewer  $k$ ,  $\delta_j$  denotes age fixed effects,  $\gamma_k$  denotes interviewer fixed effects,  $race_{ijk}^n$  represents a binary variable for whether individual  $i$  self-identifies as race  $n$ ,  $P_{ijk}^m$  represents the  $m^{\text{th}}$  genetic similarity proportions of individual  $i$ , and  $\mathbf{W}_{ijk}$  is a vector of covariates.  $P^{\text{EUR}}$  is the omitted genetic similarity proportion. The  $\beta_{n,m}$  estimates are our coefficient of interest and capture the relationship between a given genetic similarity proportion  $m$  and attractiveness among individuals who self-identify as race  $n$ .

We also fit analogous logistic regression models, except with the outcome variable instead being *very-attractive* $_{ijk}$ , a dichotomous variable for whether the interviewer rated the respondent as the highest attractiveness category ('very attractive'):

$$\ln \left( \frac{\mathbb{P}(\text{very-attractive}_{ijk} = 1)}{1 - \mathbb{P}(\text{very-attractive}_{ijk} = 1)} \right) = \delta_j + \gamma_k + \sum_{n=1}^4 (\alpha_n race_{ijk}^n) + \sum_{n=1}^4 \sum_{m=1}^3 (\beta_{n,m} race_{ijk}^n \times P_{ijk}^m) + \sum_{n=1}^4 (race_{ijk}^n \times \mathbf{W}_{ijk} \Phi)$$

Eq. 2

For logistic regression models, we report the average marginal effects (AMEs) and estimate their standard errors using fractional weighted bootstrapping (implemented via the *inferences* function in the *marginalEffects* R package).

Importantly, our linear and logistic regression specifications contain no constant term, but we nonetheless utilize omitted categories in both vectors of fixed effects; this allows for the inclusion of dummy variables for all four racial groups – White, Black, Asian/Pacific Islander, and Hispanic – into the regression without introducing multicollinearity. While each interviewer rated an average of 25 total respondents, they interviewed an average of just 5 and 4 Black and Hispanic respondents, respectively; for this reason, we pool fixed effect estimates across races, thereby boosting statistical power and increasing precision. Interviewer identifiers are constructed to be mechanically nested within waves, meaning there is no need to explicitly control for wave fixed effects.

Finally, we decompose the  $\beta_{n,m}$  estimates from Eq. 1 and Eq. 2 by interacting our coefficient of interest with a vector of  $S$  mutually exclusive subgroup variables, which we call  $sub_{ijk}^o$ :

$$attractiveness_{ijk} = \delta_j + \gamma_k + \sum_{n=1}^4 (\alpha_n race_{ijk}^n) + \sum_{n=1}^4 \sum_{m=1}^3 \sum_{o=1}^S (\beta_{n,m,o} race_{ijk}^n \times P_{ijkl}^m \times sub_{ijk}^o) + \varepsilon_{ijk} \quad \text{Eq. 3}$$

$$\ln \left( \frac{\mathbb{P}(very-attractive_{ijk} = 1)}{1 - \mathbb{P}(very-attractive_{ijk} = 1)} \right) = \delta_j + \gamma_k + \sum_{n=1}^4 (\alpha_n race_{ijk}^n) + \sum_{n=1}^4 \sum_{m=1}^3 \sum_{o=1}^S (\beta_{n,m,o} race_{ijk}^n \times P_{ijkl}^m \times sub_{ijk}^o) \quad \text{Eq. 4}$$

The subgroup variables include interviewer race (White, Black, or Other), interviewer gender, interviewer age (under 55 vs. 55 and above), respondent gender, respondent age (i.e., the average age at each Add Health wave, combining Wave I and II), and respondent census region (West, Midwest, South, or Northeast). We focus on decompositions of our unconditional models (i.e., models with no covariates besides gender) to maximize statistical power. For each of our six dimensions, we utilize an omnibus  $F$ -test to determine whether all of the subgroup-specific coefficients are statistically identical. To address concerns regarding multiple hypothesis testing, we implement a 10% Benjamini-Hochberg false discovery rate correction (34).

### Acknowledgments

We are grateful to Dalton Conley, Filiz Garip, Iain Mathieson, Ellis Monk, and Marissa Thompson for helpful comments. This research uses data from Add Health, funded by grant P01 HD31921 (Harris) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Hummer and Aiello) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill. No direct support was received from grant P01 HD31921 for this analysis. Information on obtaining Add Health data is available on the project website.

### Funding

This work has been supported by a grant from the Princeton Data Driven Social Sciences Initiative.

### Author contributions

ST designed the research. BT, LZ, and ST analyzed the data. BT, LZ, and ST wrote the paper.

### Competing interests

Authors declare that they have no competing interests.

### Data and materials availability

All results needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. We utilized the restricted Add Health survey and genotype data, which can be accessed by researchers via application at <https://data.cpc.unc.edu/projects/2/view>.

## References

1. S. F. Reardon, D. Kalogrides, K. Shores, The Geography of Racial/Ethnic Test Score Gaps. *Am. J. Sociol.* **124**, 1164–1221 (2019).
2. E. E. Bruch, M. E. J. Newman, Aspirational pursuit of mates in online dating markets. *Sci. Adv.* **4**, eaap9815 (2018).
3. E. Wrigley-Field, US racial inequality may be as deadly as COVID-19. *Proc. Natl. Acad. Sci.* **117**, 21854–21856 (2020).
4. M. Sen, O. Wasow, Race as a Bundle of Sticks: Designs that Estimate Effects of Seemingly Immutable Characteristics. *Annu. Rev. Polit. Sci.* **19**, 499–522 (2016).
5. E. P. Monk Jr., Inequality without Groups: Contemporary Theories of Categories, Intersectional Typicality, and the Disaggregation of Difference. *Sociol. Theory* **40**, 3–27 (2022).
6. E. Santana, The causal effect of skin color bias in online dating. *Soc. Sci. Res.* **124**, 103076 (2024).
7. M. E. Campbell, V. M. Keith, V. Gonlin, A. R. Carter-Sowell, Is a Picture Worth A Thousand Words? An Experiment Comparing Observer-Based Skin Tone Measures. *Race Soc. Probl.* **12**, 266–278 (2020).
8. L. Hannon, R. DeFina, The reliability of same-race and cross-race skin tone judgments. *Race Soc. Probl.* **12**, 186–194 (2020).
9. National Academies of Sciences Engineering, Medicine, *Using Population Descriptors in Genetics and Genomics Research: A New Framework for an Evolving Field* (The National Academies Press, Washington, DC, 2023; <https://nap.nationalacademies.org/catalog/26902/using-population-descriptors-in-genetics-and-genomics-research-a-new>).
10. G. L. Wojcik, M. Graff, K. K. Nishimura, R. Tao, J. Haessler, C. R. Gignoux, H. M. Highland, Y. M. Patel, E. P. Sorokin, C. L. Avery, G. M. Belbin, S. A. Bien, I. Cheng, S. Cullina, C. J. Hodonsky, Y. Hu, L. M. Huckins, J. Jeff, A. E. Justice, J. M. Kocarnik, U. Lim, B. M. Lin, Y. Lu, S. C. Nelson, S.-S. L. Park, H. Poisner, M. H. Preuss, M. A. Richard, C. Schurmann, V. W. Setiawan, A. Sockell, K. Vahi, M. Verbanck, A. Vishnu, R. W. Walker, K. L. Young, N. Zubair, V. Acuña-Alonso, J. L. Ambite, K. C. Barnes, E. Boerwinkle, E. P. Bottinger, C. D. Bustamante, C. Caberto, S. Canizales-Quinteros, M. P. Conomos, E. Deelman, R. Do, K. Doheny, L. Fernández-Rhodes, M. Fornage, B. Hailu, G. Heiss, B. M. Henn, L. A. Hindorff, R. D. Jackson, C. A. Laurie, C. C. Laurie, Y. Li, D.-Y. Lin, A. Moreno-Estrada, G. Nadkarni, P. J. Norman, L. C. Pooler, A. P. Reiner, J. Romm, C. Sabatti, K. Sandoval, X. Sheng, E. A. Stahl, D. O. Stram, T. A. Thornton, C. L. Wassel, L. R. Wilkens, C. A. Winkler, S. Yoneyama, S. Buyske, C. A. Haiman, C.

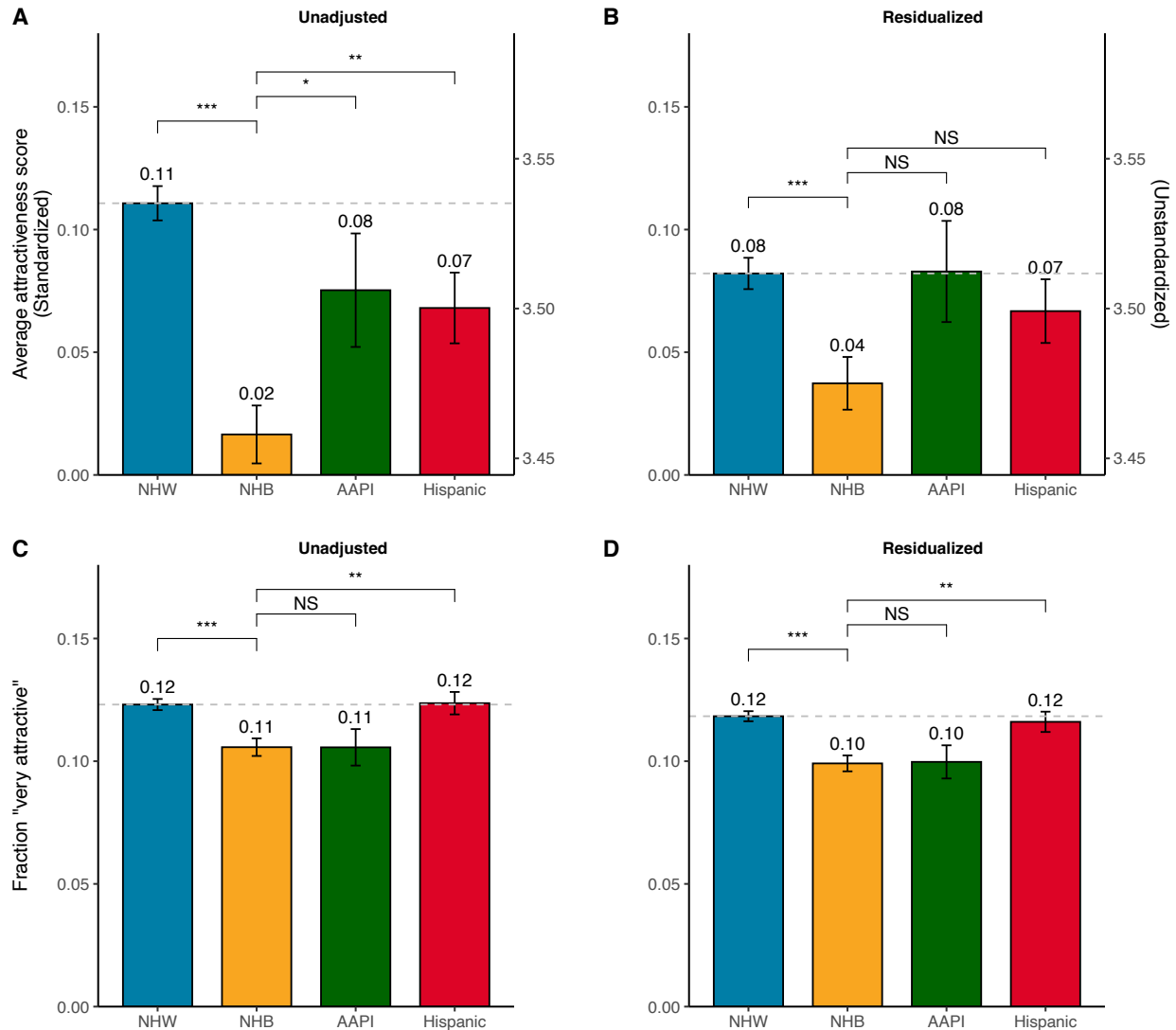
- Kooperberg, L. Le Marchand, R. J. F. Loos, T. C. Matise, K. E. North, U. Peters, E. E. Kenny, C. S. Carlson, Genetic analyses of diverse populations improves discovery for complex traits. *Nature* **570**, 514–518 (2019).
11. R. A. Patel, S. A. Musharoff, J. P. Spence, H. Pimentel, C. Tcheandjieu, H. Mostafavi, N. Sinnott-Armstrong, S. L. Clarke, C. J. Smith, P. P. Durda, K. D. Taylor, R. Tracy, Y. Liu, W. C. Johnson, F. Aguet, K. G. Ardlie, S. Gabriel, J. Smith, D. A. Nickerson, S. S. Rich, J. I. Rotter, P. S. Tsao, T. L. Assimes, J. K. Pritchard, Genetic interactions drive heterogeneity in causal variant effect sizes for gene expression and complex traits. *Am. J. Hum. Genet.* **109**, 1286–1297 (2022).
  12. W. Haak, I. Lazaridis, N. Patterson, N. Rohland, S. Mallick, B. Llamas, G. Brandt, S. Nordenfelt, E. Harney, K. Stewardson, Q. Fu, A. Mittnik, E. Bánffy, C. Economou, M. Francken, S. Friederich, R. G. Pena, F. Hallgren, V. Khartanovich, A. Khokhlov, M. Kunst, P. Kuznetsov, H. Meller, O. Mochalov, V. Moiseyev, N. Nicklisch, S. L. Pichler, R. Risch, M. A. Rojo Guerra, C. Roth, A. Szécsényi-Nagy, J. Wahl, M. Meyer, J. Krause, D. Brown, D. Anthony, A. Cooper, K. W. Alt, D. Reich, Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
  13. L. Zhang, S. Trejo, DNA, Self-Reported Ancestry, and Social Scientific Inquiry. OSF [Preprint] (2025). [https://doi.org/10.31235/osf.io/mdybz\\_v1](https://doi.org/10.31235/osf.io/mdybz_v1).
  14. T. Tan, E. G. Atkinson, Strategies for the Genomic Analysis of Admixed Populations. *Annu. Rev. Biomed. Data Sci.* **6**, 105–127 (2023).
  15. S. R. Browning, R. K. Waples, B. L. Browning, Fast, accurate local ancestry inference with FLARE. *Am. J. Hum. Genet.* **110**, 326–335 (2023).
  16. F. Privé, K. Luu, M. G. B. Blum, J. J. McGrath, B. J. Vilhjálmsson, Efficient toolkit implementing best practices for principal component analysis of population genetic data. *Bioinformatics* **36**, 4449–4457 (2020).
  17. I. Mathieson, A. Scally, What is ancestry? *PLOS Genet.* **16**, e1008624 (2020).
  18. S. Trejo, D. O. Martschenko, *What We Inherit: How New Technologies and Old Myths Are Shaping Our Genomic Future* (Princeton University Press, 2026).
  19. M. Omi, H. Winant, *Racial Formation in the United States* (Routledge, New York, ed. 3, 2014).
  20. D. I. Perrett, K. J. Lee, I. Penton-Voak, D. Rowland, S. Yoshikawa, D. M. Burt, S. P. Henzi, D. L. Castles, S. Akamatsu, Effects of sexual dimorphism on facial attractiveness. *Nature* **394**, 884–887 (1998).
  21. M. Mulford, J. Orbell, C. Shatto, J. Stockard, Physical Attractiveness, Opportunity, and Success in Everyday Exchange. *Am. J. Sociol.* **103**, 1565–1592 (1998).

22. D. Maestriperi, A. Henry, N. Nickels, Explaining financial and prosocial biases in favor of attractive people: Interdisciplinary perspectives from economics, social psychology, and evolutionary psychology. *Behav. Brain Sci.* **40**, e19 (2017).
23. M. M. Jæger, “A Thing of Beauty is a Joy Forever”?: Returns to Physical Attractiveness over the Life Course. *Soc. Forces* **89**, 983–1003 (2011).
24. J. S. Wong, A. M. Penner, Gender and the returns to attractiveness. *Res. Soc. Stratif. Mobil.* **44**, 113–123 (2016).
25. D. S. Hamermesh, J. E. Biddle, Beauty and the Labor Market. National Bureau of Economic Research 4518 [Preprint] (1993). <https://doi.org/10.3386/w4518>.
26. J. M. Fletcher, Beauty vs. brains: Early labor market outcomes of high school graduates. *Econ. Lett.* **105**, 321–325 (2009).
27. J. K. Scholz, K. Sicinski, FACIAL ATTRACTIVENESS AND LIFETIME EARNINGS: EVIDENCE FROM A COHORT STUDY. *Rev. Econ. Stat.* **97**, 14–28 (2015).
28. E. P. Monk, M. Esposito, H. Lee, Beholding Inequality: Race, Gender, and Returns to Physical Attractiveness in the United States<sup>1</sup>. *Am. J. Sociol.*, doi: 10.1086/715141 (2021).
29. C. M. Sheehan, D. S. Hamermesh, Looks and longevity: Do prettier people live longer? *Soc. Sci. Med.* **354**, 117076 (2024).
30. D. S. Hamermesh, J. Abrevaya, Beauty is the promise of happiness? *Eur. Econ. Rev.* **64**, 351–368 (2013).
31. J. Zhan, M. Liu, O. G. B. Garrod, C. Daube, R. A. A. Ince, R. E. Jack, P. G. Schyns, Modeling individual preferences reveals that face beauty is not universally perceived across cultures. *Curr. Biol.* **31**, 2243-2252.e6 (2021).
32. D. M. Altshuler, R. A. Gibbs, L. Peltonen, D. M. Altshuler, R. A. Gibbs, L. Peltonen, E. Dermitzakis, S. F. Schaffner, F. Yu, L. Peltonen, E. Dermitzakis, P. E. Bonnen, D. M. Altshuler, R. A. Gibbs, P. I. W. de Bakker, P. Deloukas, S. B. Gabriel, R. Gwilliam, S. Hunt, M. Inouye, X. Jia, A. Palotie, M. Parkin, P. Whittaker, F. Yu, K. Chang, A. Hawes, L. R. Lewis, Y. Ren, D. Wheeler, R. A. Gibbs, D. Marie Muzny, C. Barnes, K. Darvishi, M. Hurler, J. M. Korn, K. Kristiansson, C. Lee, S. A. McCarroll, J. Nemesh, E. Dermitzakis, A. Keinan, S. B. Montgomery, S. Pollack, A. L. Price, N. Soranzo, P. E. Bonnen, R. A. Gibbs, C. Gonzaga-Jauregui, A. Keinan, A. L. Price, F. Yu, V. Anttila, W. Brodeur, M. J. Daly, S. Leslie, G. McVean, L. Moutsianas, H. Nguyen, S. F. Schaffner, Q. Zhang, M. J. R. Ghorri, R. McGinnis, W. McLaren, S. Pollack, A. L. Price, S. F. Schaffner, F. Takeuchi, S. R. Grossman, I. Shlyakhter, E. B. Hostetter, P. C. Sabeti, C. A. Adebamowo, M. W. Foster, D. R. Gordon, J. Licinio, M. Cristina Manca, P. A. Marshall, I. Matsuda, D. Ngare, V. Ota Wang, D. Reddy, C. N. Rotimi, C. D. Royal, R. R. Sharp, C. Zeng, L. D. Brooks, J. E.

- McEwen, The International HapMap 3 Consortium, Principal investigators, Project coordination leaders, Manuscript writing group, Genotyping and QC, ENCODE 3 sequencing and SNP discovery, Copy number variation typing and analysis, Population analysis, Low frequency variation analysis, Linkage disequilibrium and haplotype sharing analysis, Imputation, Natural selection, Community engagement and sample collection groups, Scientific management, Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58 (2010).
33. A. Bergström, S. A. McCarthy, R. Hui, M. A. Almarri, Q. Ayub, P. Danecek, Y. Chen, S. Felkel, P. Hallast, J. Kamm, others, Insights into human genetic variation and population history from 929 diverse genomes. *Science* **367**, eaay5012 (2020).
  34. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* **57**, 289–300 (1995).
  35. S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **30** (2017).
  36. A. Regalado, 2017 was the year consumer DNA testing blew up, *MIT Technology Review*. <https://www.technologyreview.com/2018/02/12/145676/2017-was-the-year-consumer-dna-testing-blew-up/>.
  37. G. M. and J. S. Passel, Facts About the U.S. Black Population, *Pew Research Center* (2025). <https://www.pewresearch.org/race-and-ethnicity/fact-sheet/facts-about-the-us-black-population/>.
  38. K. Schaeffer, In a rising number of U.S. counties, Hispanic and black Americans are the majority, *Pew Research Center* (2019). <https://www.pewresearch.org/short-reads/2019/11/20/in-a-rising-number-of-u-s-counties-hispanic-and-black-americans-are-the-majority/>.
  39. D. S. Massey, N. A. Denton, *American Apartheid* (Harvard University Press, London, England, 1994).
  40. S. Stephens-Davidowitz, The cost of racial animus on a black candidate: Evidence using Google search data. *J. Public Econ.* **118**, 26–40 (2014).
  41. A. L. Edwards, The social desirability variable in personality assessment and research. (1957).
  42. E. Goffman, *The Presentation of Self in Everyday Life* (Doubleday, 1959).
  43. A. Sonnega, J. D. Faul, M. B. Ofstedal, K. M. Langa, J. W. Phillips, D. R. Weir, Cohort Profile: the Health and Retirement Study (HRS). *Int. J. Epidemiol.* **43**, 576–585 (2014).

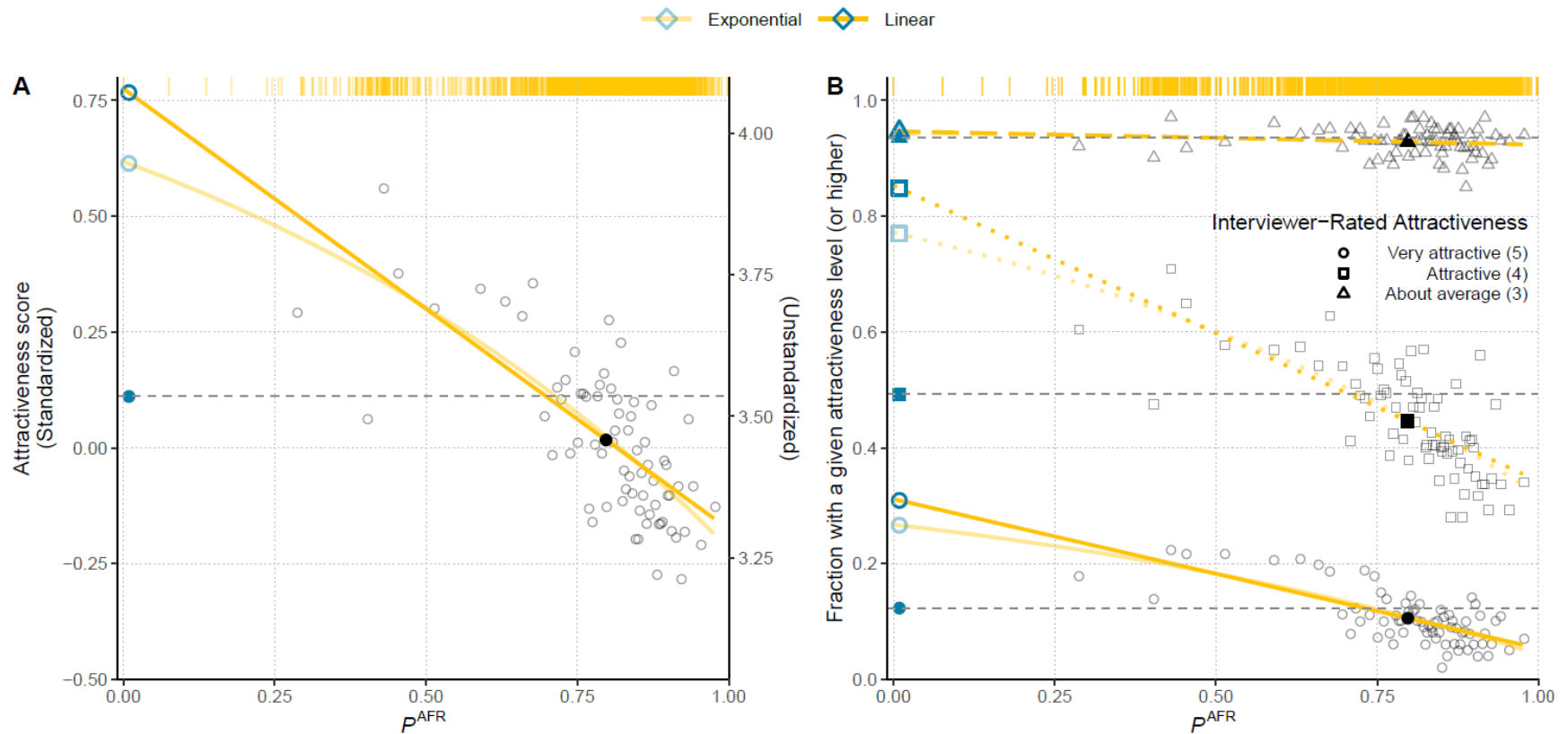
44. T. Tan, H. Jayashankar, J. Guan, S. M. Nehzati, M. Mir, M. Bennett, E. Agerbo, R. Ahlskog, V. Pinto de Andrade Anapaz, B. O. Åsvold, others, Family-GWAS reveals effects of environment and mating on genetic associations. *medRxiv*, 2024–10 (2024).
45. K. M. Harris, C. T. Halpern, E. A. Whitsel, J. M. Hussey, L. A. Killea-Jones, J. Tabor, S. C. Dean, Cohort Profile: The National Longitudinal Study of Adolescent to Adult Health (Add Health). *Int. J. Epidemiol.* **48**, 1415–1415k (2019).
46. D. H. Alexander, K. Lange, Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* **12**, 1–6 (2011).
47. S. S. Shringarpure, C. D. Bustamante, K. Lange, D. H. Alexander, Efficient analysis of large datasets and sex bias with ADMIXTURE. *BMC Bioinformatics* **17**, 218 (2016).
48. C. C. Chang, C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, J. J. Lee, Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, s13742-015-0047–8 (2015).
49. K. Bryc, E. Y. Durand, J. M. Macpherson, D. Reich, J. L. Mountain, The Genetic Ancestry of African Americans, Latinos, and European Americans across the United States. *Am. J. Hum. Genet.* **96**, 37–53 (2015).
50. D. W. Belsky, B. W. Domingue, R. Wedow, L. Arseneault, J. D. Boardman, A. Caspi, D. Conley, J. M. Fletcher, J. Freese, P. Herd, T. E. Moffitt, R. Poulton, K. Sicinski, J. Wertz, K. M. Harris, Genetic analysis of social-class mobility in five longitudinal studies. *Proc. Natl. Acad. Sci.* **115**, E7275–E7284 (2018).
51. D. W. Belsky, A. Caspi, L. Arseneault, D. L. Corcoran, B. W. Domingue, K. M. Harris, R. M. Houts, J. S. Mill, T. E. Moffitt, J. Prinz, K. Sugden, J. Wertz, B. Williams, C. L. Odgers, Genetics and the geography of health, behaviour and attainment. *Nat. Hum. Behav.* **3**, 576–586 (2019).

## Figures and Tables



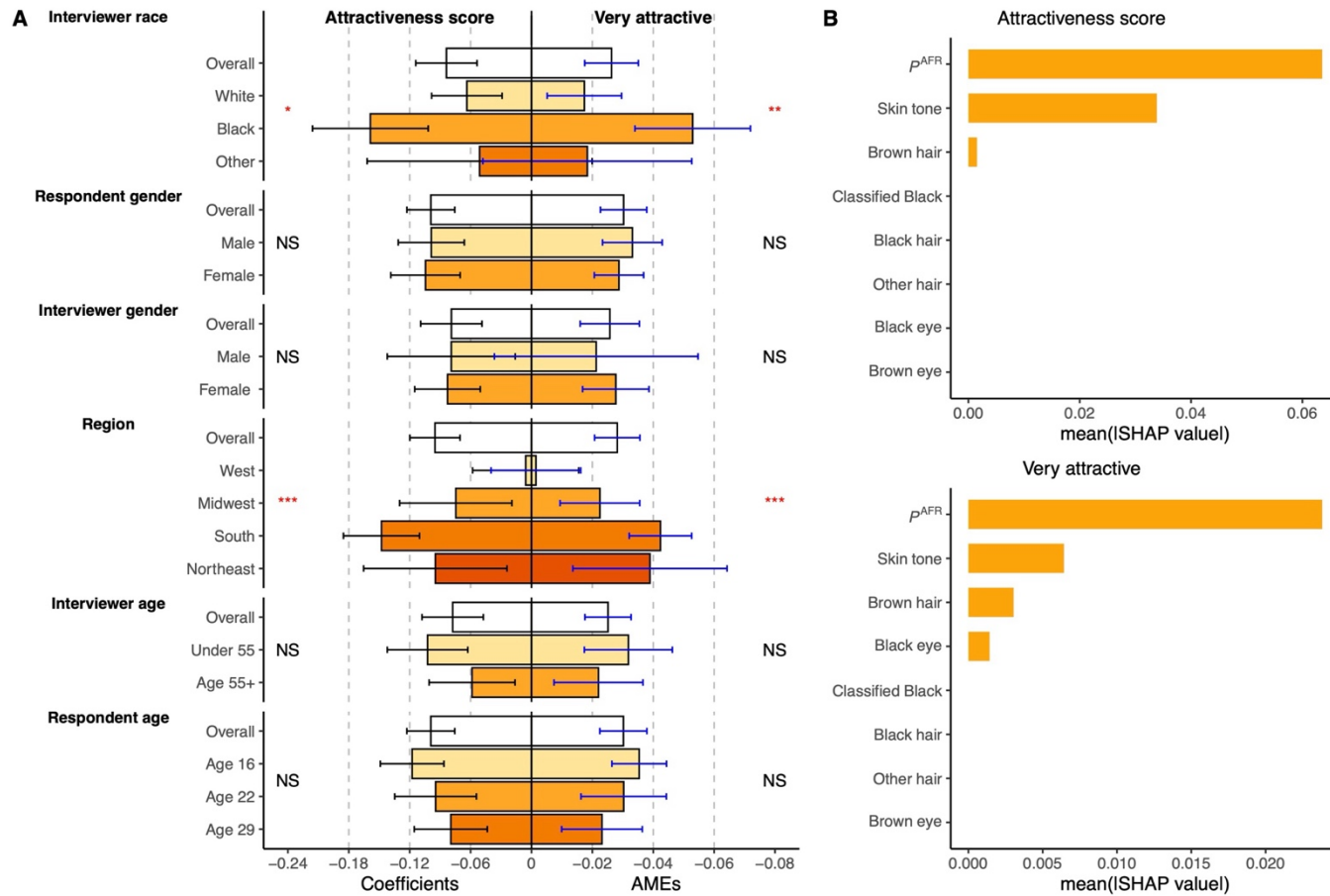
**Figure 1. Racial disparities in interviewer-rated attractiveness**

This figure displays bar graphs created using data from Waves I-IV on 9,902 genotyped respondents of the Add Health study with valid information on racial identity and attractiveness. Self-identified race is collected at Wave III, when the respondents were 18-26 years old. The four mutually exclusive racial categories are: Non-Hispanic White (NHW), Non-Hispanic Black (NHB), Non-Hispanic Asian/Pacific Islander (AAPI), and Hispanic. Panels A and B display average Likert attractiveness score (1-5), which is standardized using the weighted mean and standard deviation of the full Add Health sample. Panels C and D display the fraction of respondents rated in the very highest attractiveness category ('very attractive'). Panels A and C ('Unadjusted') display unadjusted values, whereas Panels B and D ('Residualized') display values that have been statistically adjusted for interviewer fixed effects. Error bars display 95% confidence intervals. Asterisks indicate statistical significance: \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , NS = not significant. See Figure S3 for histograms of the raw physical attractiveness ratings in Waves I-IV and see Figure S4 for an analogous bar chart broken out by respondent-interviewer racial concordance.



**Figure 2. Black Americans with greater Sub-Saharan African genetic similarity are rated as less attractive**

This figure displays binned scatter plots and linear/exponential bivariate regression fit lines using data from Waves I-IV on 2,087 genotyped non-Hispanic Black respondents of the Add Health study; Sub-Saharan African genetic similarity is plotted on the X-axis (with a marginal rug-plot displaying the individual-level distribution), and interviewer-rated attractiveness is plotted on the Y-axis. In Panel A, the attractiveness variable used is the average Likert score (1-5), which is standardized using the weighted mean and standard deviation of the full Add Health sample. In Panel B, the attractiveness variable used is the fraction of respondents rated in a given attractiveness category (or a higher category). Each bin contains approximately 100 respondent-wave observations. The large, solid blue markers display the average attractiveness and genetic similarity of non-Hispanic White Americans, and the large, hollow blue markers display the predicted attractiveness of non-Hispanic White Americans (using their average genetic similarity and extrapolation from bivariate linear/exponential regression). The large, solid black markers display the average attractiveness and genetic similarity of non-Hispanic Black Americans. See Figure S6 for a similar figure which plots mono-racial and multi-racial Black Americans separately.



**Figure 3. Decomposition and prediction analyses**

Panel A displays results from a decomposition analysis on how the relationship between Sub-Saharan African genetic similarity ( $P^{AFR}$ ) and attractiveness ratings among non-Hispanic Black respondents varies across six dimensions: interviewer race, gender, and age and respondent gender, age, and census region. Sample sizes vary across dimensions due to data availability. Beta coefficients for attractiveness score (left side) and average marginal effects for 'very attractive' (AMEs; right side) from linear and logistic regression models, respectively, are displayed; see Equations 3 and 4 in *Materials and Methods*. Statistical significance levels are denoted by asterisks (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ), with "NS" indicating non-significant results. See Tables S5 through S10 for the full set of decomposition results. Panel B presents average absolute SHAP values from two elastic net regularized regression models with attractiveness score and 'very attractive' as the target variables; SHAP values indicate the relative importance of different characteristics in predicting attractiveness ratings.

	Attractiveness Score				Pr ("Very Attractive")			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(5)
$P^{AFR}$ (10 pp)	-0.111*** (0.014)	-0.115*** (0.015)	-0.082*** (0.017)	-0.073*** (0.016)	-0.036*** (0.004)	-0.036*** (0.004)	-0.028*** (0.005)	-0.025*** (0.005)
Age FEs and Interviewer FEs	X	X	X	X	X	X	X	X
Basic Controls		X	X	X		X	X	X
Racial Classification Controls			X	X			X	X
Physical Feature Controls				X				X
Socioeconomic Controls	X	X	X	X	X	X	X	X
$N$ Observations in Subgroup	4628	4628	4628	4628	3434	3434	3434	3434
$N$ Individuals in Subgroup	1557	1557	1557	1557	1526	1526	1526	1526
$N$ Interviewers	760	760	760	760	494	494	494	494

**Table 1. Regression of attractiveness ratings on Sub-Saharan African genetic similarity among Black Americans.**

This table displays results from regressions of interviewer-rated attractiveness on Sub-Saharan African and Indigenous American genetic similarity proportions among samples of non-Hispanic Black Americans and Hispanic Americans. Models 1-4 show beta coefficients from linear regression models with continuous attractiveness score (1-5) as the outcome variable, whereas Models 5-8 show average marginal effects from logistic models with dichotomous ‘very attractive’ (0,1) as the outcome variable. Note, models 5-8 have a smaller sample size because interviewers who rated no respondents as ‘very attractive’ are mechanically dropped from the regression. The Likert attractiveness score variable is standardized using the weighted mean and standard deviation of the full Add Health sample.  $P^{EUR}$  is the omitted genetic similarity category (to eliminate multicollinearity). Among non-Hispanic Black Americans, the standard deviations of  $P^{AFR}$  and  $P^{IAM}$  are 0.13 and 0.02, respectively. See Table S3 for an analogous set of results among Hispanic Americans, and Table S4 for an unstratified version of these results (i.e., pooled across races).

## SUPPLEMENTARY INFORMATION

Leveraging Genomic Data to Document Within-Race  
Attractiveness Penalties Among Black Americans

Taddess, Zhang, & Trejo

# Contents

<b>1</b>	<b>SUPPLEMENTARY FIGURES</b>	<b>3</b>
	Figure S1: Ternary Plots of Genetic Similarity Proportions by Racial Identity . . . . .	3
	Figure S2: Distribution of Genetics Similarity Proportions among Black and Hispanic Americans .	4
	Figure S3: Distribution of Physical Attractiveness Ratings among Add Health Respondents . . . .	5
	Figure S4: Relationship between $P^{\text{AFR}}$ and $P^{\text{IAM}}$ and Attractiveness Among Hispanic Americans .	5
	Figure S5: Attractiveness Ratings by Race and Respondent-Interviewer Racial Concordance . . . .	5
	Figure S6: Relationship between $P^{\text{AFR}}$ and Attractiveness by Interviewer Race . . . . .	5
	Figure S7: Relationship between $P^{\text{AFR}}$ and Attractiveness (Monoracial vs. Multiracial Black) . . .	5
	Figure S8: Comparing Attractiveness Penalties by Skin Tone and by $P^{\text{AFR}}$ among Black Americans	5
	Figure S9: Directed Acyclic Graph . . . . .	11
	Figure S10: Distribution of Days between Interviews among Add Health Interviewers . . . . .	12
<b>2</b>	<b>SUPPLEMENTARY TABLES</b>	<b>13</b>
	Table S1: Descriptive Statistics of Add Health Genotyped Sample by Racial Identity . . . . .	13
	Table S2: Descriptive Statistics of Add Health Interviewers . . . . .	14
	Table S3: Regression Models of Attractiveness on GSPs (Hispanic Americans) . . . . .	15
	Table S4: Regression Models of Attractiveness on GSPs (All Races Pooled) . . . . .	16
	Table S5: Regression Decomposition by Interviewer Race . . . . .	17
	Table S6: Regression Decomposition by Respondent Gender . . . . .	18
	Table S7: Regression Decomposition by Interviewer Gender . . . . .	19
	Table S8: Regression Decomposition by Census Region . . . . .	20
	Table S9: Regression Decomposition by Interviewer Age . . . . .	21
	Table S10: Regression Decomposition by Respondent Age . . . . .	22

# 1 SUPPLEMENTARY FIGURES

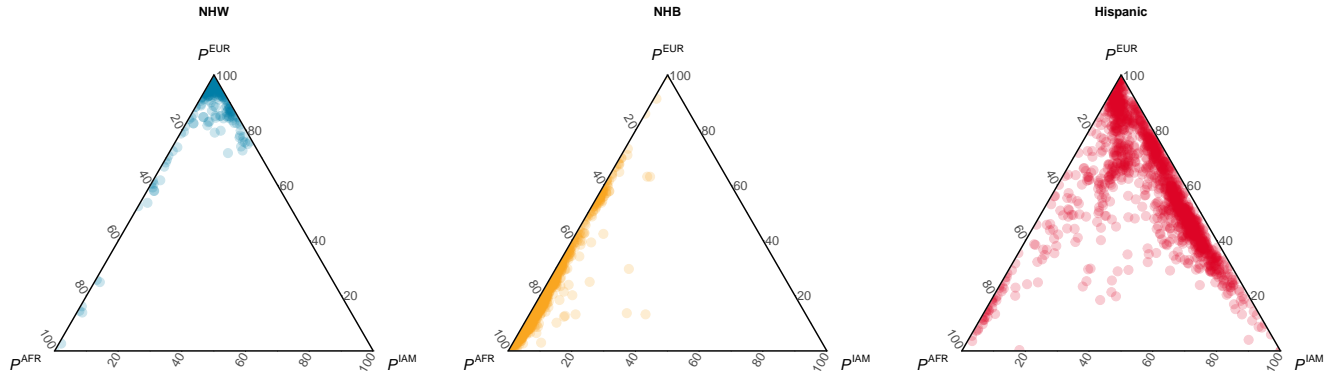


Figure S1: Ternary Plots of Genetic Similarity Proportions by Racial Identity

This figure displays ternary plots of genetic similarity proportion separately for members of three mutually exclusive self-identified racial groups: Non-Hispanic White Americans, Non-Hispanic Black Americans, and Hispanic Americans. Each dot represents an individual, with their position in the ternary plot reflecting their levels of Sub-Saharan African, European, and Indigenous American genetic similarity proportions. The top, bottom left, and bottom right corners of each plot correspond to 100% European, Sub-Saharan African, and Indigenous American genetic similarity, respectively. Note, 77 White respondents, 31 Black respondents, and 76 Hispanic respondents with values of East Asian genetic similarity ( $P^{EAS}$ ) greater than 5% are not displayed. White Americans exhibit little evidence of admixture, with individuals generally having very low levels of both Sub-Saharan African and Indigenous American genetic similarity. Black Americans exhibit evidence of two-way admixture, with substantial variation on the axis between Sub-Saharan African and European genetic similarity and very few individuals having meaningful amounts of Indigenous American genetic similarity. Finally, Hispanic Americans exhibit evidence of three-way admixture, perhaps reflecting the complex demographic history of Latin America.

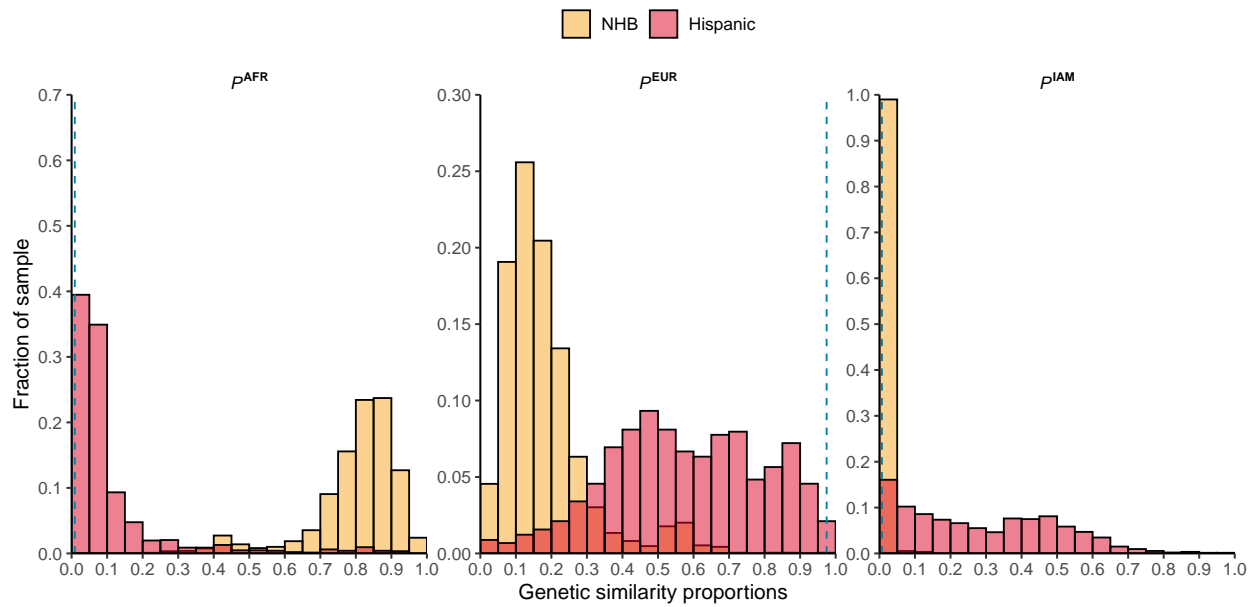


Figure S2: Distribution of Genetics Similarity Proportions among Black and Hispanic Americans

This figure displays histograms of Sub-Saharan African (left panel), European (middle panel), and Indigenous American (right panel) genetic similarity among genotyped respondents of the Add Health study. Histograms of two self-identified racial groups – Non-Hispanic Black Americans in yellow, and Hispanic Americans in red – are displayed separately. The blue dashed line in each panel indicates the average level of the respective genetic similarity proportion among self-identified non-Hispanic White Americans. Self-identified race is collected at Wave III, when the respondents were 18-26 years old; racial categories are mutually exclusive.

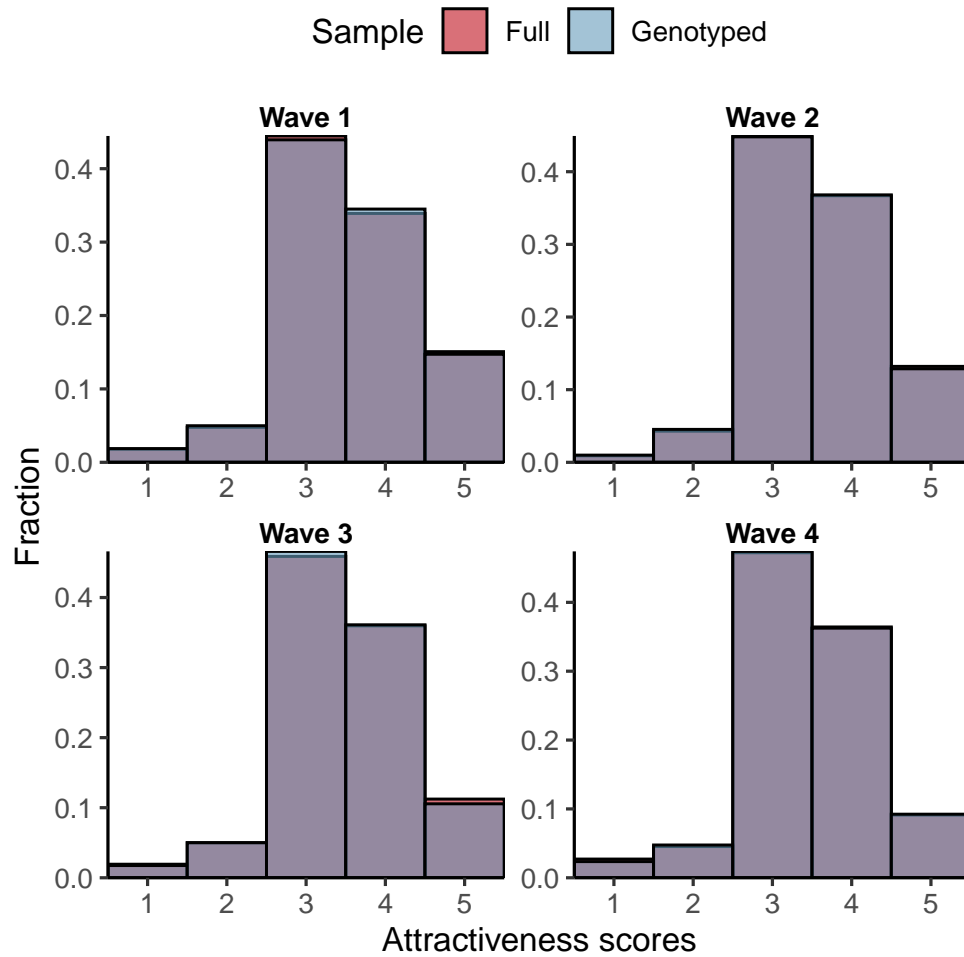


Figure S3: Distribution of Physical Attractiveness Ratings among Add Health Respondents

This figure displays histograms of the physical attractiveness ratings in Waves I–IV of the Add Health study. The blue bars represent the full Add Health sample, whereas the red bars represent the Add Health genotyped subsample. The close correspondence between the two distributions suggests that our analytic sample is largely representative of the overall Add Health population with respect to attractiveness ratings.

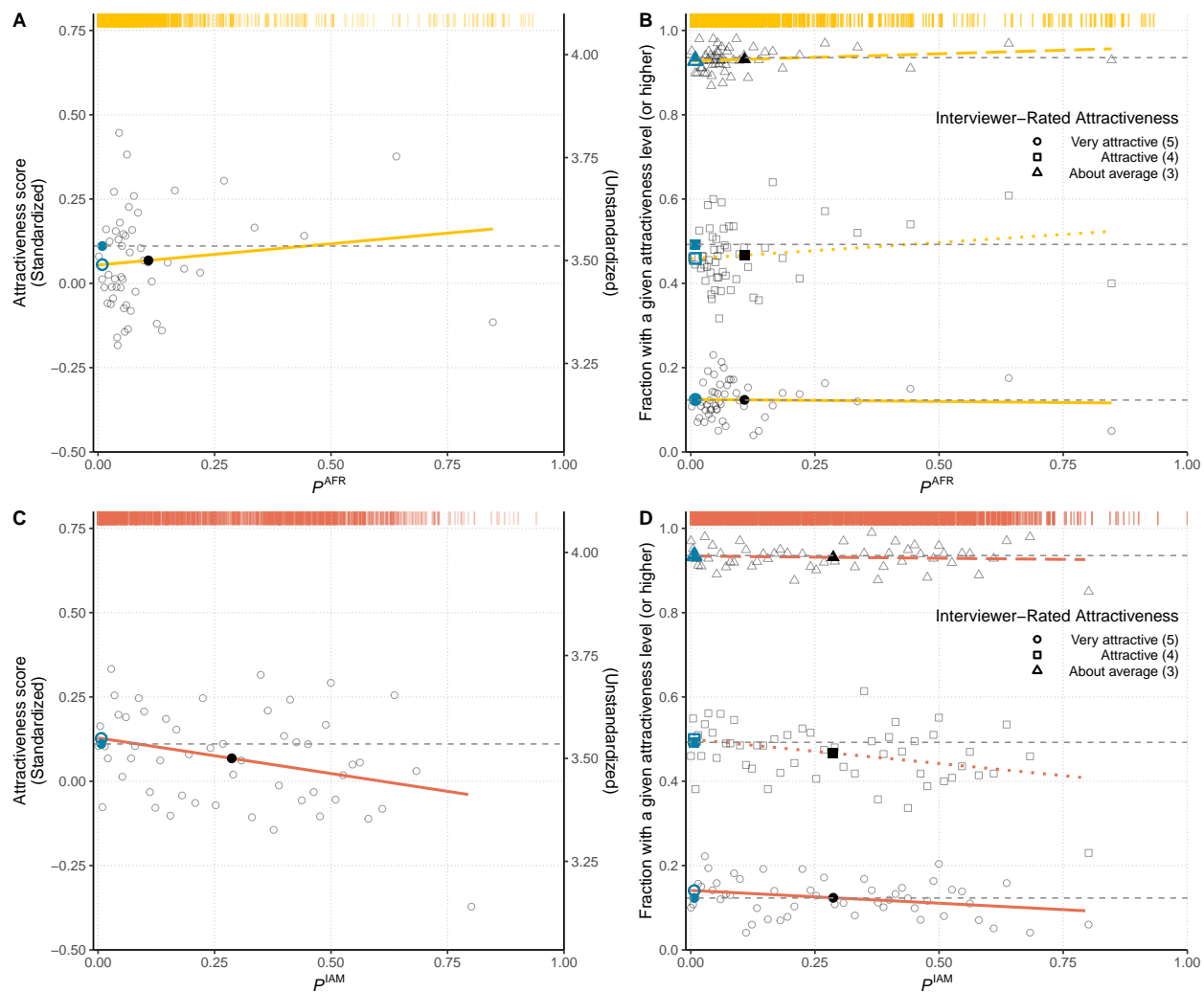


Figure S4: Relationship between  $P^{AFR}$  and  $P^{IAM}$  and Attractiveness Among Hispanic Americans

This figure displays binned scatter plots and quadratic fit regression lines using data from Waves I-IV on 1,469 genotyped Hispanic respondents of the Add Health study. In Panels A and B, Sub-Saharan African genetic similarity is plotted on the X-axis, whereas in Panels C and D, Indigenous American genetic similarity is plotted on the X-axis. In all panels, interviewer-rated attractiveness is plotted on the Y-axis; while the attractiveness variable used in Panels A and C is the average Likert score (1-5), the attractiveness variable used in Panels B and D is the fraction of respondents rated in a given attractiveness category or a higher category. Each bin contains approximately 100 respondent-wave observations. The large, solid blue markers display the average attractiveness and genetic similarity of non-Hispanic White Americans, and the large, hollow blue markers display the predicted attractiveness of non-Hispanic White Americans (using their average genetic similarity). The large, solid black markers display the average attractiveness and genetic similarity of Hispanic Americans. While Panels C and D suggest that, among Hispanic Americans, Indigenous American genetic similarity is negatively associated with attractiveness ratings, this relationship is considerably weaker than the relationship between Sub-Saharan African and attractiveness ratings observed among Black Americans.

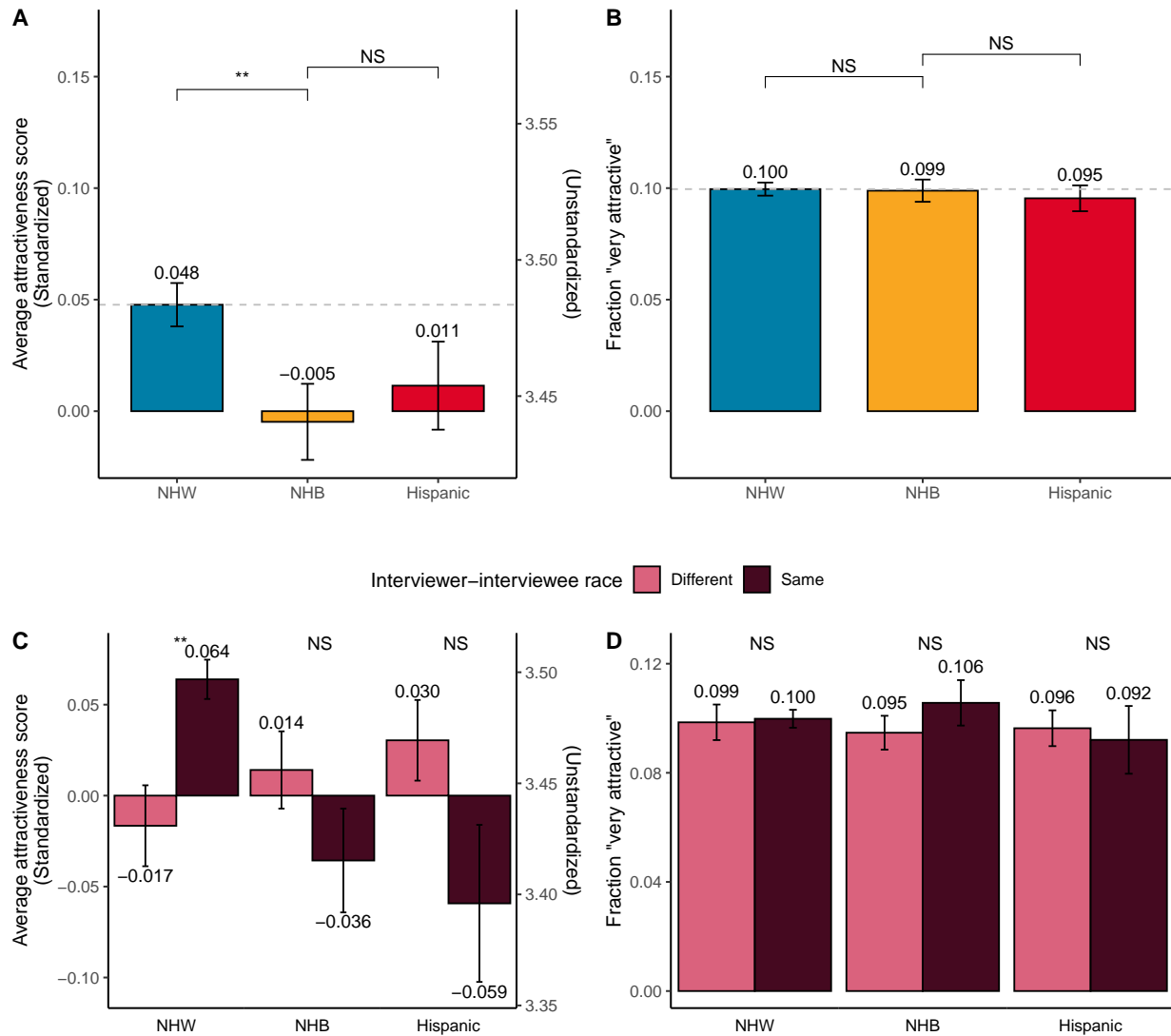


Figure S5: Attractiveness Ratings by Race and Respondent-Interviewer Racial Concordance

This figure displays bar graphs created using data on genotyped respondents of the Add Health study; only Waves III and IV contain information on the race of each interviewer, and therefore only attractiveness ratings from these two Add Health waves are displayed. Self-identified race is collected at Wave III, when the respondents were 18-26 years old; racial categories are mutually exclusive. Panels A and C display average Likert attractiveness score (1-5), which is standardized using the weighted mean and standard deviation of the full Add Health sample. Panels B and D display the fraction of respondents rated in the very highest attractiveness category ('very attractive'). Panels A and B show the average attractiveness ratings for each of three mutually exclusive self-identified racial groups: Non-Hispanic White Americans, Non-Hispanic Black Americans, and Hispanic Americans. Panels C and D further break these outcomes out by whether the respondent and interviewer are of the same race. White respondents receive slightly higher attractiveness scores (but not 'very attractive' ratings) from White interviewers than from interviewers of other races; in contrast, there is no statistically significant difference in ratings between same-race and different-race interviewers for Black and Hispanic respondents. Error bars display 95% confidence intervals. Asterisks indicate statistical significance: \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , NS = not significant.

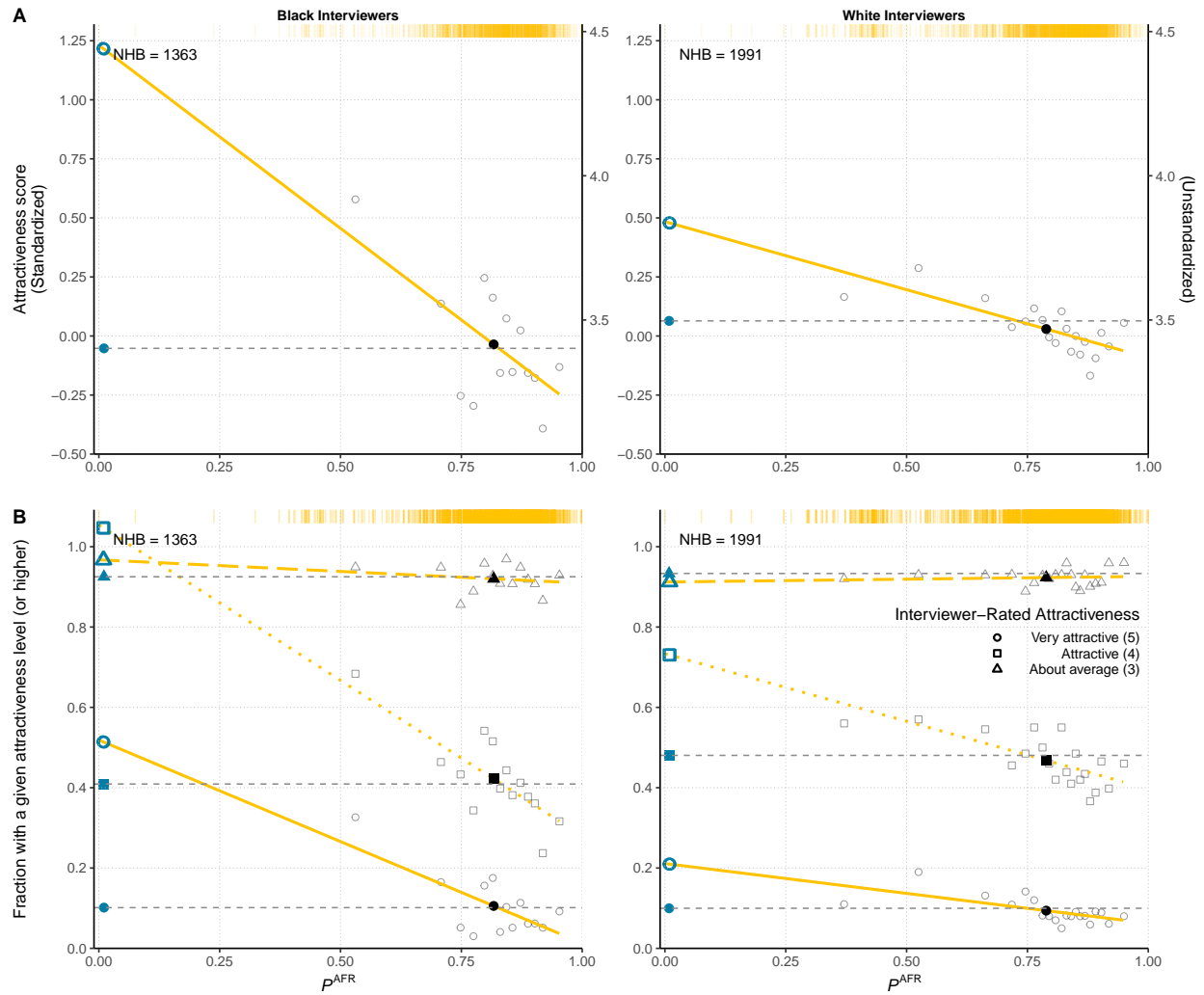


Figure S6: Relationship between  $P^{AFR}$  and Attractiveness by Interviewer Race

This figure displays binned scatter plots and quadratic fit regression lines using data on genotyped non-Hispanic Black respondents of the Add Health study; only Waves III and IV contain information on the race of each interviewer, and therefore only attractiveness ratings from these two Add Health waves are displayed. This figure, unlike Figure 2, displays the relationship between Sub-Saharan African genetic similarity (X-axis) and interviewer-rated attractiveness (Y-axis) separately for non-Hispanic Black interviewers and non-Hispanic White interviewers. The attractiveness variable used in Panel A is the average Likert score (1-5), whereas the attractiveness variable used in Panel B is the fraction of respondents rated in a given attractiveness category or a higher category. Each bin contains approximately 100 observations. The large, solid blue markers display the average attractiveness and genetic similarity of non-Hispanic White Americans, and the large, hollow blue markers display the predicted attractiveness of non-Hispanic White Americans (using their average genetic similarity). The large, solid black markers display the average attractiveness and genetic similarity of non-Hispanic Black Americans. While the attractiveness penalties associated with Sub-Saharan African genetic similarity are largest for Black interviewers, they are nonetheless also present for White interviewers; importantly, the larger slope estimates among Black interviewers appear to, at least in part, result from the fact that they are more likely than interviewers of other races to rate Black respondents with low Sub-Saharan African genetic similarity as ‘attractive’ or ‘very attractive’.

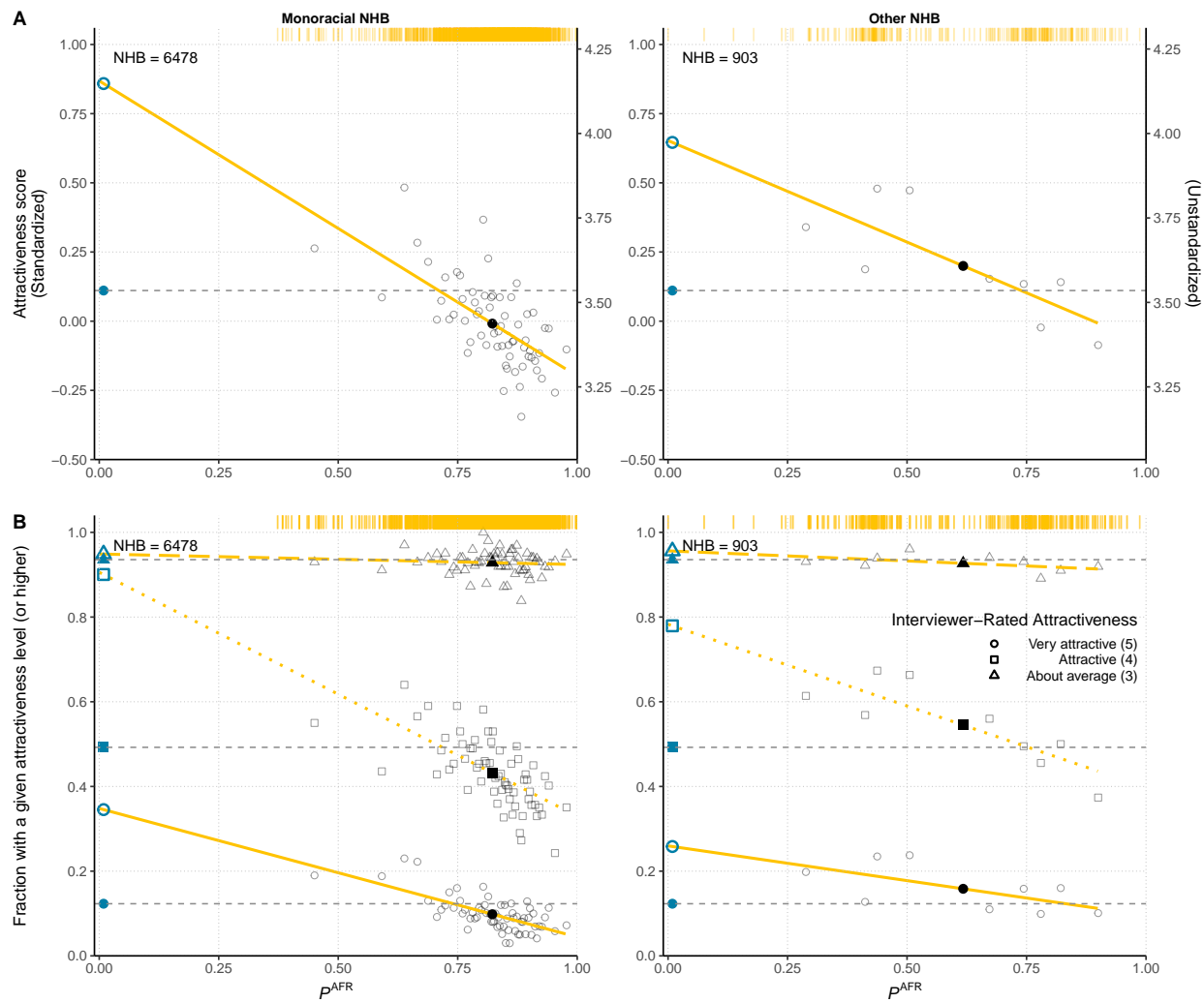


Figure S7: Relationship between  $P^{AFR}$  and Attractiveness (Monoracial vs. Multiracial Black)

This figure displays binned scatter plots and quadratic fit regression lines using data on genotyped non-Hispanic Black respondents of the Add Health study; only Waves III and IV contain information on the race of each interviewer, and therefore only attractiveness ratings from these two Add Health waves are displayed. This figure, unlike Figure 2, displays the relationship between Sub-Saharan African genetic similarity (X-axis) and interviewer-rated attractiveness (Y-axis) separately for Black respondents who identify exclusively as Black (i.e., monoracial) and for Black respondents who identify as Black and an additional race/ethnicity (multiracial). The attractiveness variable used in Panel A is the average Likert score (1-5), whereas the attractiveness variable used in Panel B is the fraction of respondents rated in a given attractiveness category or a higher category. Each bin contains approximately 100 observations. The large, solid blue markers display the average attractiveness and genetic similarity of non-Hispanic White Americans, and the large, hollow blue markers display the predicted attractiveness of non-Hispanic White Americans (using their average genetic similarity). The large, solid black markers display the average attractiveness and genetic similarity of non-Hispanic Black Americans. The negative association between Sub-Saharan African genetic similarity and attractiveness is similar for monoracial and multiracial Black Americans. This indicates that observed attractiveness penalties are not merely artifacts of ambiguity in racial identification/classification but rather likely reflect consistent stigmatization of physical features associated with Sub-Saharan African genetic similarity.

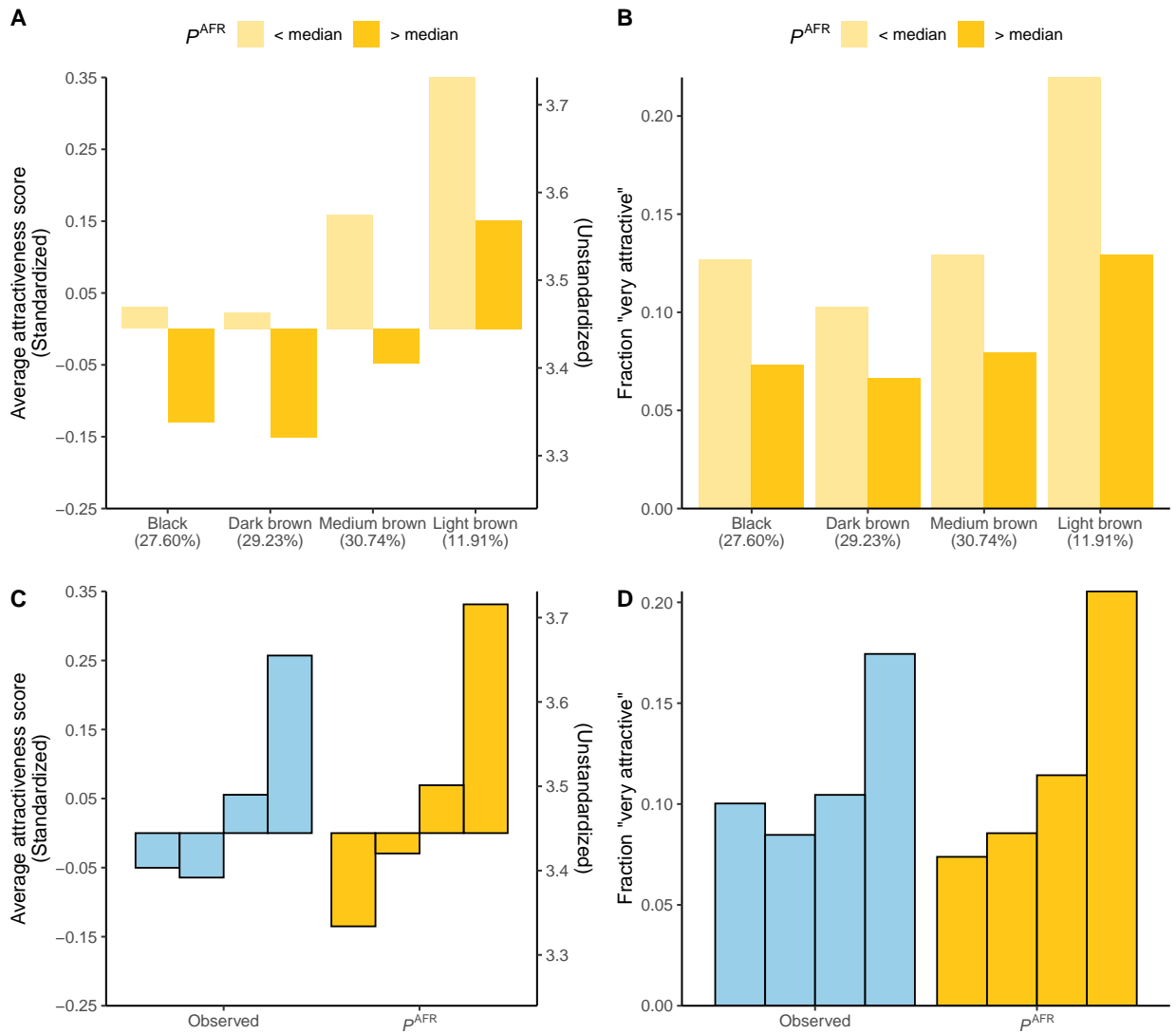


Figure S8: Comparing Attractiveness Penalties by Skin Tone and by  $P^{AFR}$  among Black Americans

This figure displays data on genotyped non-Hispanic Black respondents from Waves I-IV of the Add Health Study. Panels A and B show the average attractiveness ratings by interviewer-rated skin tone. Within each skin tone category, respondents are further divided into two equally sized subgroups at the median value of Sub-Saharan African genetic similarity in a given skin tone category. Panels C and D compare the skin tone attractiveness gradient to Sub-Saharan African genetic similarity attractiveness gradient; four groups of Sub-Saharan African genetic similarity are constructed by ranking the respondents in ascending order of  $P^{AFR}$  and dividing them into four groups using the same quantiles that separate each successive category in the observed distribution of skin tone. Panels A and C display average Likert attractiveness score (1-5), whereas Panels B and D display the fraction of respondents rated in the very highest attractiveness category ('very attractive'). Black respondents rated as white-skinned ( $N = 9$ ) are excluded from the figure. Even among Black Americans with the same interviewer-rated skin tone, those with higher Sub-Saharan African genetic similarity proportions receive lower attractiveness ratings. In addition, skin tone groups display a weaker attractiveness gradient than equally sized groups based on Sub-Saharan African genetic similarity.

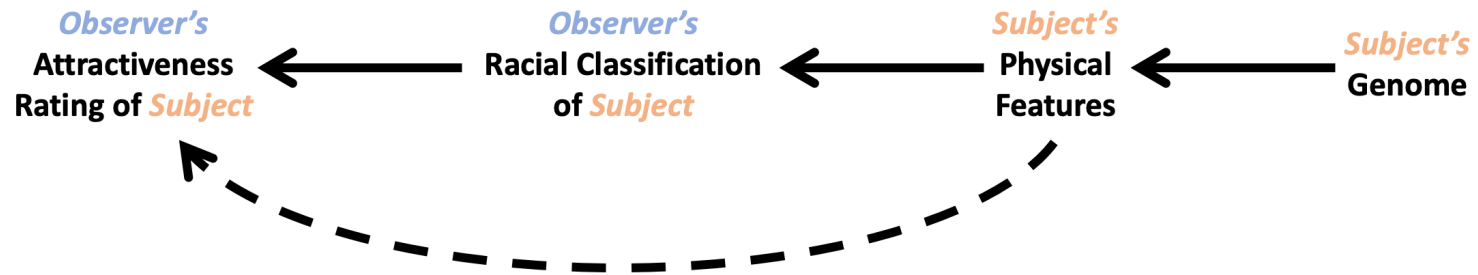


Figure S9: Directed Acyclic Graph

This directed acyclic graph (DAG) illustrates theoretical causal pathways linking genetic information to attractiveness ratings. The DAG depicts four key nodes: a subject's genome, a subject's physical features, an observer's racial classification of the subject, and an observer's attractiveness rating of the subject. A person's genome influences their eventual physical features via biological development. In turn, a person's physical features may influence their perceived attractiveness [i] indirectly via their racial classification by others and [ii] directly (if those features are culturally and/or aesthetically valued). Our study, by exploring within-group attractiveness variation among a population with little variation in racial classification, highlights the existence of a direct relationship between genetically influenced racialized physical features and perceived attractiveness (indicated by the dashed arrow above).

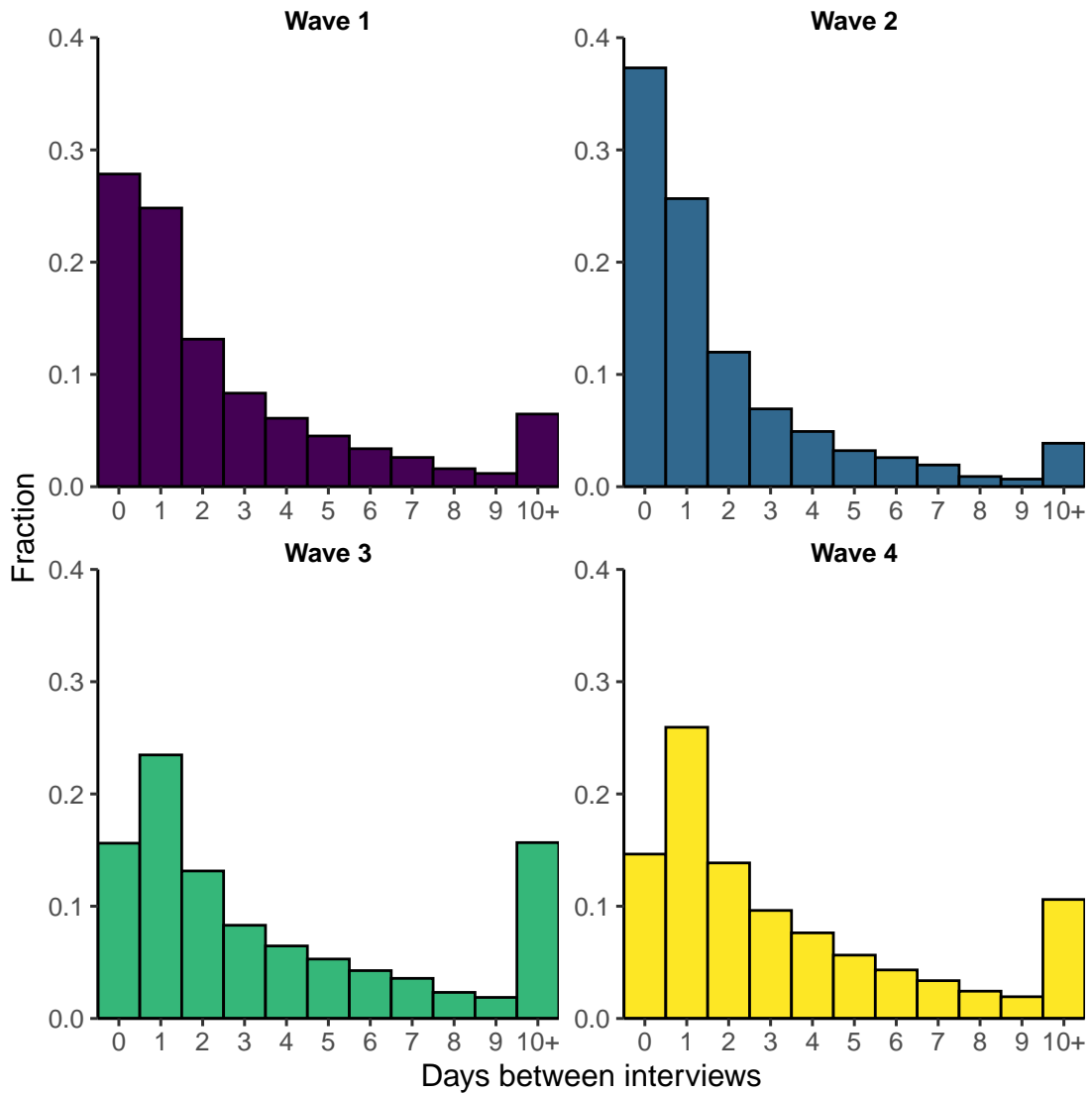


Figure S10: Distribution of Days between Interviews among Add Health Interviewers

This figure displays histograms of the number of days between respondent interviews conducted by the same interviewer across Waves I-IV of the Add Health study. Interviewers rarely conducted more than one interview within a single day, and sometimes more than a week passed between two interviews.

## 2 SUPPLEMENTARY TABLES

Characteristic	Non-Hispanic Black			Hispanic			Other Races		
	N	Mean / %	SD	N	Mean / %	SD	N	Mean / %	SD
$P^{AFR}$	2087	0.80	0.13	1469	0.11	0.15	6415	0.01	0.04
$P^{EUR}$	2087	0.18	0.12	1469	0.58	0.22	6415	0.90	0.25
$P^{EAS}$	2087	0.01	0.04	1469	0.03	0.11	6415	0.08	0.24
$P^{IAM}$	2087	0.01	0.02	1469	0.29	0.21	6415	0.01	0.05
Attractiveness score	7381	0.02	1.01	1469	0.07	1.04	6415	0.11	1.01
"Very attractive"	7381	10.57%		1469	12.36%		6415	12.12%	
Gender	2087			1469			6415		
Male		44.47%			50.99%			47.67%	
Female		55.53%			49.01%			52.33%	
Birth year	2087	1978.98	1.76	1469	1978.60	1.73	6415	1978.95	1.74
Educational attainment	2087			1469			6415		
Less than high school		10.97%			10.28%			7.19%	
High school		17.39%			21.10%			15.98%	
Some college		48.92%			48.26%			43.73%	
Bachelor's or above		22.71%			20.35%			33.11%	
Skin tone	1721			1201			5450		
Black		27.60%			1.00%			0.22%	
Dark brown		29.23%			2.66%			0.42%	
Medium brown		30.74%			11.32%			1.96%	
Light brown		11.91%			36.30%			6.97%	
White		0.52%			48.71%			90.42%	
Hair color	1721			1201			5450		
No hair		1.16%			1.50%			0.73%	
Black		79.78%			46.13%			13.93%	
Brown		16.27%			43.80%			53.60%	
Blond		0.35%			5.08%			26.11%	
Red		1.16%			1.58%			4.59%	
Grey		0.00%			0.00%			0.00%	
Other		1.28%			1.92%			1.05%	
Eye color	1721			1201			5450		
Black		16.91%			8.66%			2.42%	
Brown		79.90%			74.94%			40.99%	
Hazel		2.38%			8.33%			17.83%	
Blue		0.06%			3.50%			27.54%	
Green		0.41%			3.08%			7.49%	
Other		0.35%			1.50%			3.72%	
W1 social origins score	1907	-0.37	1.46	1347	-0.59	1.41	6157	0.29	1.17
W1 neighborhood disadvantage	2064	34.33	11.31	1459	29.27	10.49	6334	22.18	9.71

Table S1: Descriptive Statistics of Add Health Genotyped Sample by Racial Identity

Means, proportions, and standard deviations of genetic similarity proportions, attractiveness ratings, demographic characteristics, and physical features among genotyped members of the Add Health Study. Descriptive statistics are displayed separately for non-Hispanic Black Americans, Hispanic Americans, and Non-Hispanic Other Races. Sample sizes vary slightly across measures due to missing data. For the attractiveness score and 'very attractive', the reported  $N$  values reflect the total number of ratings across Waves I-IV. Genetic similarity proportions range from 0 to 1, with  $P^{AFR}$  indicating Sub-Saharan African genetic similarity,  $P^{EUR}$  indicating European genetic similarity,  $P^{EAS}$  indicating East Asian genetic similarity, and  $P^{IAM}$  indicating Indigenous American genetic similarity.

<b>Characteristic</b>	<b>N = 667</b>	
	<b>Mean (SD) / %</b>	
	<b>Unweighted / Weighted</b>	
Number of interviewees	25 (20)	
Number of Non-Hispanic Black interviewees	5 (8)	
Interviewer's age	51 (11)	52 (11)
Female interviewer	83%	83%
Interviewer's race		
Non-Hispanic White	62%	56%
Non-Hispanic Black	26%	38%
Non-Hispanic Native American	1%	1%
Non-Hispanic Asian/Pacific Islander	1%	0%
Hispanic	9%	5%
Other	1%	1%
Interviewer's educational attainment		
High school or less	11%	12%
Some college	43%	36%
Bachelor's degree or above	47%	52%

Table S2: Descriptive Statistics of Add Health Interviewers

Means, proportions, and standard deviations of selected demographic characteristics of Wave III and IV Add Health interviewers. Weighted column displays descriptive statistics where each interviewer is weighted by the number of Non-Hispanic Black respondents that they interviewed.

	Attractiveness Score				Pr("Very Attractive")			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$P^{AFR}$ (10 pp)	0.016 (0.013)	0.002 (0.017)	0.008 (0.017)	0.010 (0.017)	-0.003 (0.005)	-0.009 (0.007)	-0.007 (0.007)	-0.007 (0.007)
$P^{IAM}$ (10 pp)	-0.027* (0.010)	-0.027* (0.011)	-0.022+ (0.012)	-0.010 (0.012)	-0.013** (0.004)	-0.012** (0.004)	-0.009+ (0.005)	-0.006 (0.005)
Age FEs and Interviewer FEs	X	X	X	X	X	X	X	X
Basic Controls		X	X	X		X	X	X
Racial Classification Controls			X	X			X	X
Physical Feature Controls				X				X
Socioeconomic Controls	X	X	X	X	X	X	X	X
N Observations in Subgroup	3,254	3,254	3,254	3,254	2,333	2,333	2,333	2,333
N Individuals in Subgroup	1,095	1,095	1,095	1,095	1,068	1,068	1,068	1,068
N Interviewers	727	727	727	727	487	487	487	487

Table S3: Regression Models of Attractiveness on GSPs (Hispanic Americans)

This table displays results from regressions of interviewer-rated attractiveness on Sub-Saharan African and Indigenous American genetic similarity proportions among Hispanic Americans. The top panel shows beta coefficients from linear regression models with Likert attractiveness score (1-5) as the outcome variable, whereas the bottom panel displays average marginal effects from logistic models with a ‘very attractive’ (0,1) as the outcome variable. The attractiveness score variable is standardized using the weighted mean and standard deviation of the full Add Health sample.  $P^{EUR}$  is the omitted genetic similarity category (to eliminate multicollinearity). Note, Among Hispanic Americans, the standard deviations of  $P^{AFR}$  and  $P^{IAM}$  are 0.15 and 0.21, respectively.

	Attractiveness Score				Pr("Very Attractive")			
	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
$P^{AFR}$ (10 pp)	-0.010*** (0.002)	-0.042*** (0.008)	-0.035*** (0.009)	-0.021* (0.010)	-0.005*** (0.001)	-0.015*** (0.003)	-0.016*** (0.004)	-0.013*** (0.004)
$P^{EAS}$ (10 pp)	-0.003 (0.003)	-0.001 (0.008)	-0.005 (0.009)	0.001 (0.009)	-0.004** (0.001)	-0.004 (0.003)	-0.006+ (0.003)	-0.004 (0.003)
$P^{IAM}$ (10 pp)	-0.017** (0.005)	-0.015* (0.006)	-0.023** (0.007)	-0.001 (0.008)	-0.005** (0.002)	-0.003 (0.002)	-0.007* (0.003)	-0.001 (0.003)
Basic Controls	X	X	X	X	X	X	X	X
Racial Classification Controls		X	X	X		X	X	X
Physical Feature Controls			X	X			X	X
Socioeconomic Controls				X				X
Age FEs and Interviewer FEs	X	X	X	X	X	X	X	X
Num. Obs.	35,508	27,998	24,852	23,285	27,194	21,755	18,859	17,429
Num. Inds.	9,971	9,971	8,372	7,843	9,837	9,745	8,216	7,678
Num. Age	20	20	20	20	20	20	20	20
Num. Inter. × Waves	1,564	1,218	1,211	1,207	1,041	814	783	765
$R^2$	0.181	0.171	0.182	0.194	—	—	—	—
$R^2$ Within	0.009	0.010	0.016	0.026	—	—	—	—
Pseudo $R^2$	—	—	—	—	0.156	0.153	0.154	0.162

Table S4: Regression Models of Attractiveness on GSPs (All Races Pooled)

This table displays results from regressions of interviewer-rated attractiveness on genetic similarity proportions in the full sample of genotyped Add Health respondents. That is, in this set of models, the association between each genetic similarity proportion and attractive rating is *not* estimated separately for each self-identified racial group but instead pooled across groups. While Models 1-4 on the left show beta coefficients from linear regression models with Likert attractiveness score (1-5) as the outcome variable, Models 1-4 on the right show average marginal effects from logistic models with a ‘very attractive’ (0,1) as the outcome variable. The attractiveness score variable is standardized using the weighted mean and standard deviation of the full Add Health sample.  $P^{EUR}$  is the omitted genetic similarity category (to eliminate multicollinearity).

	Non-Hispanic Black				Hispanic			
	White Int.	Black Int.	Other Int.	P-value	White Int.	Hisp. Int.	Other Int.	P-value
<b>Attractiveness Score</b>								
$P^{AFR}$ (10 pp)	-0.064*** (0.018)	-0.159*** (0.029)	-0.051 (0.057)	0.015 (0.044)	0.023 (0.017)	-0.053+ (0.029)	0.017 (0.023)	0.066 (0.167)
$P^{IAM}$ (10 pp)	-0.159 (0.262)	0.080 (0.151)	0.179 (0.504)	0.694 (0.832)	-0.037* (0.014)	-0.053* (0.023)	-0.022 (0.021)	0.604 (0.839)
N Respondents	1519	1143	227		1035	489	587	
N Interviewers	272	144	58		271	60	133	
N Observations	1991	1363	238		1357	543	667	
<b>Pr("Very Attractive")</b>								
$P^{AFR}$ (10 pp)	-0.017** (0.006)	-0.053*** (0.010)	-0.018 (0.018)	0.002 (0.006)	0.000 (0.008)	-0.056 (0.043)	-0.014 (0.017)	0.140 (0.210)
$P^{IAM}$ (10 pp)	0.020 (0.067)	-0.025 (0.058)	0.110 (0.353)	0.614 (0.832)	-0.013* (0.006)	-0.016 (0.012)	-0.016+ (0.009)	0.919 (0.926)
N Respondents	1312	901	175		877	364	395	
N Interviewers	195	82	37		200	33	80	
N Observations	1638	1015	180		1085	397	426	

Table S5: Regression Decomposition by Interviewer Race

This table displays results from stratified regressions of interviewer-rated attractiveness on Sub-Saharan African and Indigenous American genetic similarity proportions among samples of non-Hispanic Black Americans and Hispanic Americans. The top panel shows beta coefficients from linear regression models with Likert attractiveness score (1-5) as the outcome variable, whereas the bottom panel displays average marginal effects from logistic models with a ‘very attractive’ (0,1) as the outcome variable. The attractiveness score variable is standardized using the weighted mean and standard deviation of the full Add Health sample.  $P^{EUR}$  is the omitted genetic similarity category (to eliminate multicollinearity). Note, these regressions are estimated on only Waves III and IV, as Waves I and II did not collect information on the interviewer race. We utilize an omnibus  $F$ -test to determine whether all of the subgroup-specific coefficients are statistically identical. Raw  $p$ -values are displayed without parentheses above, and adjusted  $p$ -values (using a Benjamini- Hochberg false discovery rate correction) are displayed in parentheses below.

	Non-Hispanic Black			Hispanic		
	Male Resp.	Female Resp.	P-value	Male Resp.	Female Resp.	P-value
<b>Attractiveness Score</b>						
$P^{AFR}$ (10 pp)	-0.099*** (0.017)	-0.104*** (0.017)	0.809 (0.882)	0.034* (0.016)	-0.021 (0.015)	0.009 (0.112)
$P^{IAM}$ (10 pp)	-0.015 (0.093)	-0.116 (0.173)	0.605 (0.832)	-0.040*** (0.011)	-0.020 (0.013)	0.210 (0.839)
N Respondents	928	1159		749	720	
N Interviewers	816	865		772	758	
N Observations	3239	4142		2630	2546	
<b>Pr("Very Attractive")</b>						
$P^{AFR}$ (10 pp)	-0.033*** (0.005)	-0.029*** (0.004)	0.517 (0.689)	0.004 (0.006)	-0.013* (0.006)	0.038 (0.153)
$P^{IAM}$ (10 pp)	0.004 (0.061)	0.000 (0.034)	0.938 (0.938)	-0.019*** (0.006)	-0.010* (0.005)	0.179 (0.839)
N Respondents	915	1143		740	713	
N Interviewers	566	606		534	548	
N Observations	2378	3212		1902	1984	

Table S6: Regression Decomposition by Respondent Gender

This table displays results from stratified regressions of interviewer-rated attractiveness on Sub-Saharan African and Indigenous American genetic similarity proportions among samples of non-Hispanic Black Americans and Hispanic Americans. The top panel shows beta coefficients from linear regression models with Likert attractiveness score (1-5) as the outcome variable, whereas the bottom panel displays average marginal effects from logistic models with a ‘very attractive’ (0,1) as the outcome variable. The attractiveness score variable is standardized using the weighted mean and standard deviation of the full Add Health sample.  $P^{EUR}$  is the omitted genetic similarity category (to eliminate multicollinearity). We utilize an omnibus  $F$ -test to determine whether all of the subgroup-specific coefficients are statistically identical. Raw  $p$ -values are displayed without parentheses above, and adjusted  $p$ -values (using a Benjamini- Hochberg false discovery rate correction) are displayed in parentheses below.

	Non-Hispanic Black			Hispanic		
	Male Int.	Female Int.	P-value	Male Int.	Female Int.	P-value
<b>Attractiveness Score</b>						
$P^{AFR}$ (10 pp)	-0.079*	-0.083***	0.916	-0.043	0.015	0.070
	(0.032)	(0.017)	(0.916)	(0.029)	(0.015)	(0.167)
$P^{IAM}$ (10 pp)	0.094	0.050	0.866	-0.050+	-0.035**	0.577
	(0.184)	(0.197)	(0.938)	(0.026)	(0.012)	(0.839)
N Respondents	542	1933		402	1376	
N Interviewers	82	397		79	389	
N Observations	608	3000		441	2152	
<b>Pr("Very Attractive")</b>						
$P^{AFR}$ (10 pp)	-0.021	-0.028***	0.657	-0.025	-0.009	0.304
	(0.017)	(0.006)	(0.788)	(0.023)	(0.007)	(0.365)
$P^{IAM}$ (10 pp)	0.052	0.014	0.633	-0.010	-0.017**	0.489
	(0.168)	(0.041)	(0.832)	(0.013)	(0.006)	(0.839)
N Respondents	365	1756		233	1203	
N Interviewers	42	275		44	272	
N Observations	396	2446		244	1688	

Table S7: Regression Decomposition by Interviewer Gender

This table displays results from stratified regressions of interviewer-rated attractiveness on Sub-Saharan African and Indigenous American genetic similarity proportions among samples of non-Hispanic Black Americans and Hispanic Americans. The top panel shows beta coefficients from linear regression models with Likert attractiveness score (1-5) as the outcome variable, whereas the bottom panel displays average marginal effects from logistic models with a ‘very attractive’ (0,1) as the outcome variable. The attractiveness score variable is standardized using the weighted mean and standard deviation of the full Add Health sample.  $P^{EUR}$  is the omitted genetic similarity category (to eliminate multicollinearity). Note, these regressions are estimated on only Waves III and IV, as Waves I and II did not collect information on the interviewer gender. We utilize an omnibus  $F$ -test to determine whether all of the subgroup-specific coefficients are statistically identical. Raw  $p$ -values are displayed without parentheses above, and adjusted  $p$ -values (using a Benjamini-Hochberg false discovery rate correction) are displayed in parentheses below.

	Non-Hispanic Black					Hispanic				
	West	Midwest	South	Northeast	P-value	West	Midwest	South	Northeast	P-value
<b>Attractiveness Score</b>										
$P^{\text{AFR}}$ (10 pp)	-0.006 (0.027)	-0.075** (0.028)	-0.148*** (0.019)	-0.095** (0.036)	0.000 (0.001)	0.013 (0.032)	0.071* (0.030)	-0.015 (0.015)	0.013 (0.026)	0.083 (0.167)
$P^{\text{IAM}}$ (10 pp)	-0.285 (0.178)	-0.104 (0.065)	0.032 (0.122)	0.508 (0.336)	0.145 (0.832)	-0.024 (0.015)	-0.010 (0.041)	-0.034* (0.014)	-0.026 (0.025)	0.926 (0.926)
N Respondents	282	367	1236	133		580	109	548	192	
N Interviewers	282	235	528	166		366	183	372	188	
N Observations	995	1304	4384	446		2034	379	1939	672	
<b>Pr("Very Attractive")</b>										
$P^{\text{AFR}}$ (10 pp)	-0.001 (0.008)	-0.022*** (0.007)	-0.042*** (0.005)	-0.039** (0.013)	0.000 (0.001)	0.004 (0.011)	0.015 (0.011)	-0.016** (0.006)	-0.006 (0.018)	0.024 (0.145)
$P^{\text{IAM}}$ (10 pp)	-0.003 (0.094)	-0.014 (0.056)	-0.027 (0.043)	0.124 (0.260)	0.508 (0.832)	-0.007 (0.005)	-0.019 (0.016)	-0.019** (0.006)	-0.004 (0.012)	0.364 (0.839)
N Respondents	276	363	1224	126		576	108	545	184	
N Interviewers	190	176	371	98		248	135	274	115	
N Observations	730	1018	3339	273		1567	292	1460	411	

Table S8: Regression Decomposition by Census Region

This table displays results from stratified regressions of interviewer-rated attractiveness on Sub-Saharan African and Indigenous American genetic similarity proportions among samples of non-Hispanic Black Americans and Hispanic Americans. The top panel shows beta coefficients from linear regression models with Likert attractiveness score (1-5) as the outcome variable, whereas the bottom panel displays average marginal effects from logistic models with a ‘very attractive’ (0,1) as the outcome variable. The attractiveness score variable is standardized using the weighted mean and standard deviation of the full Add Health sample.  $P^{\text{EUR}}$  is the omitted genetic similarity category (to eliminate multicollinearity). We utilize an omnibus  $F$ -test to determine whether all of the subgroup-specific coefficients are statistically identical. Raw  $p$ -values are displayed without parentheses above, and adjusted  $p$ -values (using a Benjamini- Hochberg false discovery rate correction) are displayed in parentheses below.

	Non-Hispanic Black			Hispanic		
	Int. under 55	Int. aged 55+	P-value	Int. under 55	Int. aged 55+	P-value
<b>Attractiveness Score</b>						
$P^{AFR}$ (10 pp)	-0.102*** (0.020)	-0.059** (0.022)	0.122 (0.293)	0.011 (0.017)	-0.011 (0.020)	0.372 (0.406)
$P^{IAM}$ (10 pp)	0.186 (0.157)	-0.082 (0.215)	0.328 (0.832)	-0.035* (0.015)	-0.043** (0.015)	0.699 (0.839)
N Respondents	1535	1291		1020	987	
N Interviewers	281	194		262	197	
N Observations	2014	1549		1342	1209	
<b>Pr("Very Attractive")</b>						
$P^{AFR}$ (10 pp)	-0.032*** (0.007)	-0.022** (0.007)	0.221 (0.374)	-0.017+ (0.010)	-0.004 (0.008)	0.227 (0.303)
$P^{IAM}$ (10 pp)	0.043 (0.055)	0.006 (0.055)	0.521 (0.832)	-0.017* (0.008)	-0.014* (0.006)	0.681 (0.839)
N Respondents	1310	1061		829	795	
N Interviewers	186	132		180	133	
N Observations	1613	1227		994	935	

Table S9: Regression Decomposition by Interviewer Age

This table displays results from stratified regressions of interviewer-rated attractiveness on Sub-Saharan African and Indigenous American genetic similarity proportions among samples of non-Hispanic Black Americans and Hispanic Americans. The top panel shows beta coefficients from linear regression models with Likert attractiveness score (1-5) as the outcome variable, whereas the bottom panel displays average marginal effects from logistic models with a ‘very attractive’ (0,1) as the outcome variable. The attractiveness score variable is standardized using the weighted mean and standard deviation of the full Add Health sample.  $P^{EUR}$  is the omitted genetic similarity category (to eliminate multicollinearity). Note, these regressions are estimated on only Waves III and IV, as Waves I and II did not collect information on the interviewer age. We utilize an omnibus  $F$ -test to determine whether all of the subgroup-specific coefficients are statistically identical. Raw  $p$ -values are displayed without parentheses above, and adjusted  $p$ -values (using a Benjamini- Hochberg false discovery rate correction) are displayed in parentheses below.

	Non-Hispanic Black				Hispanic			
	Age 16	Age 22	Age 29	P-value	Age 16	Age 22	Age 29	P-value
<b>Attractiveness Score</b>								
$P^{AFR}$ (10 pp)	-0.118*** (0.016)	-0.094*** (0.021)	-0.080*** (0.018)	0.190 (0.374)	0.006 (0.016)	-0.008 (0.019)	0.013 (0.017)	0.692 (0.692)
$P^{IAM}$ (10 pp)	-0.161+ (0.094)	0.062 (0.200)	-0.020 (0.133)	0.400 (0.832)	-0.020+ (0.012)	-0.036* (0.016)	-0.040** (0.012)	0.397 (0.839)
N Respondents	2063	1718	2078		1449	1196	1466	
N Interviewers	555	267	229		517	250	231	
N Observations	3585	1718	2078		2514	1196	1466	
<b>Pr("Very Attractive")</b>								
$P^{AFR}$ (10 pp)	-0.035*** (0.005)	-0.030*** (0.007)	-0.023*** (0.007)	0.249 (0.374)	0.002 (0.005)	-0.017+ (0.010)	-0.007 (0.010)	0.132 (0.210)
$P^{IAM}$ (10 pp)	-0.017 (0.045)	0.058 (0.078)	-0.032 (0.049)	0.228 (0.832)	-0.011* (0.005)	-0.016+ (0.008)	-0.018* (0.009)	0.643 (0.839)
N Respondents	1723	1328	1695		1256	960	1037	
N Interviewers	381	179	151		364	177	150	
N Observations	2567	1328	1695		1889	960	1037	

Table S10: Regression Decomposition by Respondent Age

This table displays results from stratified regressions of interviewer-rated attractiveness on Sub-Saharan African and Indigenous American genetic similarity proportions among samples of non-Hispanic Black Americans and Hispanic Americans. The top panel shows beta coefficients from linear regression models with Likert attractiveness score (1-5) as the outcome variable, whereas the bottom panel displays average marginal effects from logistic models with a ‘very attractive’ (0,1) as the outcome variable. The attractiveness score variable is standardized using the weighted mean and standard deviation of the full Add Health sample.  $P^{EUR}$  is the omitted genetic similarity category (to eliminate multicollinearity). We utilize an omnibus  $F$ -test to determine whether all of the subgroup-specific coefficients are statistically identical. Raw  $p$ -values are displayed without parentheses above, and adjusted  $p$ -values (using a Benjamini- Hochberg false discovery rate correction) are displayed in parentheses below.