



THE UNIVERSITY OF CHICAGO

AI-DRIVEN JUDGMENTS: VISUALIZING OCCUPATIONAL
STEREOTYPES FROM FACIAL FEATURES

By
Zimo Ma

May, 2025

A paper submitted in partial fulfillment of the requirements for
the Master of Arts degree in the Master of Arts in
Computational Social Science

Faculty Advisor: Alexander Todorov

Preceptor: Henry Dambanemuya

Abstract

Physical appearance, particularly facial attractiveness, plays a significant but context-sensitive role in workplace evaluations. While the “what is beautiful is good” stereotype suggests a universal beauty premium effect in organizational settings, recent research has shown that additional facial traits, such as trustworthiness, competence, gender, and age, also shape judgments, and their effects vary by occupational context. This study investigates how both appearance- and competence-based visual traits influence perceived job suitability across client-facing and back-office roles in two industries: finance and retail. A total of 376 participants rated 817 hyper-realistic synthetic faces, generated using StyleGAN2, on their suitability for four roles: Financial Advisor, Risk Analyst, Retail Sales Associate, and Inventory Manager. Predictive models trained on these ratings were used to estimate trait scores of perceived job suitability and facial traits for novel faces and analyze their associations. The client-facing Retail Sales Associate role was strongly associated with attractiveness-based traits, while both back-office roles and the client-facing Financial Advisor prioritized competence-based traits, with appearance-based traits showing weaker or negative correlations. These findings suggest that the influence of facial traits on perceived job suitability depends more on the industry context than on whether a role involves client-facing interactions. The findings challenge the notion of a uniform beauty premium and emphasizes the need for bias-conscious hiring practices and theoretical frameworks that account for contextual variability in impression formation.

Keywords: facial judgment, visual stereotypes, attractiveness bias, job suitability, hiring decisions, occupational evaluation

1 Introduction

In professional evaluations, decisions regarding who gets hired, promoted, or trusted are not always based solely on qualifications. A growing body of research shows that facial appearance, particularly physical attractiveness, plays a powerful role in shaping impressions and occupational judgments. The “what is beautiful is good” stereotype, first identified by Dion et al. (1972), suggests that attractive individuals are naturally assumed to be more competent, trustworthy, socially skilled, and have better marital life. Subsequent studies have confirmed that attractive candidates receive more callbacks in the interview, higher salaries, and more favorable evaluations across a variety of roles and industries (Chiu & Babcock, 2002; Hamermesh & Biddle, 1994; Ruffle & Shtudiner, 2015).

However, recent work suggests that attractiveness is only one of many facial cues that

influence professional evaluations. People also make rapid inferences about trustworthiness, competence, and dominance based on facial features—impressions that may complement or override attractiveness effects depending on the context (Oh et al., 2023; Todorov, 2012). Moreover, these judgments cannot be applied uniformly. Factors such as the evaluator and candidate’s gender, the candidate’s age, and the visibility of the role can significantly shape how appearance cues are interpreted (Agthe et al., 2011; Korthase & Trenholme, 1982; Timming et al., 2017).

This study examines how facial impressions, including attractiveness and other inferred traits, affect perceptions of job suitability across different occupational contexts. Specifically, it compares evaluations of facial images in client-facing versus back-office roles, within two industries that differ in gender representation: *Finance* and *retail*. These contrasts allow us to examine how visual expectations vary not only by job function (client-facing versus back-office) but also by the gendered norms of the industry.

This leads to two core research questions:

- (1) What are the visual stereotypes associated with client-facing versus back-office occupations?
- (2) What role does facial attractiveness play within these stereotypes?

To address these questions, participants were randomly assigned to evaluate one of four occupational roles, *Financial Advisor*, *Risk Analyst*, *Retail Sales Associate*, or *Inventory Manager*, by rating the job suitability of 817 hyper-realistic faces generated via StyleGAN2 (Peterson et al., 2022). Predictive models trained on human ratings were then applied to a new set of synthetic faces to estimate suitability scores of the four job roles. Attribute scores, including attractiveness, trustworthiness, dominance, competence, and age, were calculated for each novel face using a predictive model trained on human responses in Peterson et al. (2022). Correlating these predicted attributes with predicted suitability allowed an evaluation of appearance-based expectations regarding job function and industry context.

This approach contributes to a more detailed understanding of how appearance stereotypes operate in hiring scenarios. Theoretically, it challenges simplified notions of a universal “beauty premium” by showing how different visual traits are emphasized across job types. Practically, it offers insights into how organizations can design fairer hiring processes that mitigate implicit bias based on facial characteristics.

2 Literature Review

In organizational settings, visual stereotypes, which are assumptions formed based on visible traits at the surface level, can influence how candidates are perceived and judged. These judgments are often made rapidly and unconsciously and in fact influence employ-

ment, promotion, and interpersonal perceptions (Dion et al., 1972; Judge & Cable, 2004; Ruffle & Shtudiner, 2015; Solomon, 2022). Research in psychology and organizational behavior has shown that attributes such as age, gender, attire, facial expressions, and trustworthiness can significantly bias professional judgments (Agthe et al., 2011; Davis & Lennon, 1988; Finkelstein et al., 2015; Frost, 2014; Oh et al., 2023). These visible characteristics often act as cognitive shortcuts, guiding impressions before any substantive information about an individual’s ability or qualifications is known.

2.1 Facial Attractiveness and the “What is Beautiful is Good” Stereotype

Among all visual stereotypes, physical attractiveness stands out as one of the most influential yet complex factors in professional evaluations. As first proposed by Dion et al. (1972), the “what is beautiful is good” stereotype suggests the halo effect that it is the natural tendency that people have to extrapolate from the attractive traits possessed by an individual, and assume they must also possess other desirable traits as well. Specifically, it linked physical attractiveness to perceived social and professional competence, such as being more competent, more likely to have successful careers, having a better marriage life, and so on (Dion et al., 1972).

Building on foundational research by Dion et al. (1972), subsequent studies have tried to understand how these perceptions influence success in professional settings and found out that the notion that more attractive individuals are consistently perceived as more capable, confident, and successful (Chiu & Babcock, 2002; Hamermesh & Biddle, 1994; Heineck, 2005; Ruffle & Shtudiner, 2015). For example, Hamermesh and Biddle (1994) explored the economic advantages of beauty and examined how looks affect job candidates’ earnings. They found that attractive individuals enjoy a “beauty premium” in terms of earnings (Hamermesh & Biddle, 1994). Similarly, Chiu and Babcock (2002) investigated the effect of facial attractiveness on interview selection chances, and it concluded that attractiveness has the most significant influence on hiring decisions and the probability of getting an interview. The impact of attractiveness even overshadowed work-related factors, such as work-related skills, work experience, GPA, and public examination results (Chiu & Babcock, 2002). Moreover, Chiu and Babcock (2002) found that perceived attractiveness was positively correlated with ratings of work-related skills and professional experience, and it suggested that human resource professionals may infer higher competence and stronger work experience from more attractive candidates, independent of actual qualifications.

2.2 Beyond Attractiveness: Gender, Social Perceptions and Facial Judgments

Although physical attractiveness is a dominant visual cue in social evaluations, recent work has highlighted that it is not the only factor shape perceptions of professional success. Gender also plays a significant role in modulating how attractiveness is interpreted. For example, Ruffle and Shtudiner (2015) found that attractive male candidates received significantly more call-backs than their less attractive counterparts, while attractive female candidates often experienced penalties for their appearance.

In addition, recent studies highlight that attractiveness-related evaluations are highly conditional, depending on the gender of both the evaluator and the target. Agthe et al. (2010) and Agthe et al. (2011) demonstrated that the evaluators tended to rate same-sex attractive individuals more negatively, while the opposite-sex evaluators generally favored attractive candidates. These findings presented the importance of the gender dynamics in shaping how attractiveness is interpreted in professional contexts and point to the need for broader consideration of other appearance-based traits beyond attractiveness alone.

Moreover, other literature suggests that people spontaneously infer a wide range of social traits from faces, such as trustworthiness, competence, dominance, and emotional valence, which also play critical roles in occupational judgments (Todorov, 2012).

In response to the halo effect of “what is good is beautiful” attractiveness bias, Oh et al. (2023) developed data-driven facial models to disentangle perceived trustworthiness from attractiveness. By subtracting or orthogonalizing attractiveness-related information from trustworthiness representations, results showed that trustworthiness judgments can be driven independently by other cues, especially facial expressions related to approachability and positive emotion. Faces manipulated to appear more trustworthy without the attractiveness cues were consistently rated as more trustworthy, approachable, and happier by both human raters and machine learning algorithms, which suggests that these emotional signals are foundational to trustworthiness perception.

Furthermore, Oh et al. (2019) demonstrated that impressions of competence and dominance can be learned and shaped through social-conceptual associations rather than fixed visual cues. Their findings emphasized that facial features become associated with certain trait judgments through exposure and cultural learning, meaning that competence judgments are not merely downstream from attractiveness but can be guided by domain-specific stereotype structures.

Additionally, Oh et al. (2020) explored the variability of face trait impressions across global populations. They found that facial inferences about trustworthiness, competence, and dominance are underpinned by a shared latent structure but are differentially weighted across cultures, with certain traits (e.g., warmth vs. agency) prioritized depending on the

local social ecology. This suggests that the trait dimensions inferred from faces may interact with cultural, professional, and contextual expectations to shape employment judgments, instead of purely attractiveness.

2.3 The Role of Perceived Age in Appearance-Based Judgments

Age is also a prominent visual cue that can trigger deeply rooted stereotypes in professional settings. Although age is strongly and negatively correlated with perceived attractiveness (Korthase & Trenholme, 1982), age-related bias is not merely a byproduct of declining aesthetic appeal. Rather, perceptions of age activate distinct social stereotypes that signal reduced adaptability, declining productivity, or higher employment cost—regardless of an individual’s actual performance or qualifications. These biases may be reinforced by cultural and evolutionary associations that equate youth with vitality, fertility, and capacity for innovation (Solomon, 2022).

Empirical studies confirm the organizational consequences of such bias. Older workers are frequently perceived as less adaptable, less productive, or more costly to employ, even when their qualifications are equal or superior to those of younger colleagues (Finkelstein et al., 1995, 2015; Ho et al., 1999). A meta-analysis by Bal et al. (2011) supports this view, showing that although older workers are sometimes perceived as more dependable, negative stereotypes about flexibility and trainability remain prevalent in professional evaluations. In addition, Finkelstein et al. (1995) found that younger raters tended to evaluate older candidates less favorably, especially when no job-relevant information was available. Similarly, in the Hong Kong labor market, older workers faced higher unemployment, longer job searches, fewer promotions, and diminished training opportunities, despite cultural norms that emphasize respect for the elderly (Ho et al., 1999, 2000). These patterns reflect a global tendency to undervalue older workers due to perceived costs, rather than demonstrated limitations.

However, age is not always a disadvantage. Zhao et al. (2021), in a meta-analysis of over 65,000 entrepreneurs, found that age is positively associated with entrepreneurial success, particularly among women. Frieze et al. (1990) also observed that older male professionals earned more, even when controlling for prior experience. These findings suggest that, while age-related stereotypes are pervasive, their real-world implications vary by context—such as job function, gender, and industry norms.

2.4 Industry Contexts and Functional Roles in Appearance Evaluation

While previous research has established the general influence of physical appearance on professional evaluations, fewer studies have examined how these expectations vary across

job functions and industries. Scholars have suggested that certain roles, particularly those involving direct client interaction, are thought to impose stronger aesthetic demands than internally focused, analytical roles (Timming et al., 2017). For example, Timming et al. (2017) found that body art such as tattoos or piercings negatively affected hireability for customer-facing positions but had minimal impact on back-office roles. These findings suggest that appearance-based standards may be shaped by both occupational function and role expectations.

This perspective aligns with broader social and industry norms. In retail, hospitality, and beauty industries, for example, front-line employees are often expected to maintain a polished, attractive appearance that aligns with brand image and customer expectations (Solomon, 2022; Warhurst et al., 2009). In contrast, roles in finance, engineering, or data analysis are typically more internally focused, where greater emphasis is placed on credentials, technical expertise, and reliability over visual appeal. Such variation suggests that the impact of facial attractiveness, as well as other visual cues, may be shaped by both job function and the normative expectations embedded within industry culture (Rivera, 2012).

2.5 Research Questions and Hypotheses

While previous research has demonstrated that visual cues such as age, gender, trustworthiness, and attractiveness influence workplace judgments, fewer studies have systematically examined how these biases interact with structural features of employment—namely, job function and industry context. Existing work has suggested that client-facing roles, particularly in appearance-oriented sectors such as retail or hospitality, may impose higher expectations regarding appearance, while technical or back-office roles in fields like finance or engineering often prioritize competence over aesthetics (Rivera, 2012; Timming et al., 2017; Warhurst et al., 2009). Yet the precise role that facial attractiveness plays in these differentiated expectations remains unclear.

This ambiguity gives rise to two central research questions: (1) What are the visual stereotypes associated with client-facing versus back-office roles? and (2) How does facial attractiveness contribute to these stereotypes? The first question is motivated by evidence that employers often infer job-relevant psychological traits—such as sociability, competence, or professionalism—based on facial appearance, particularly when roles require interpersonal interaction (Oh et al., 2023; Peterson et al., 2022; Rivera, 2012; Todorov, 2012; Warhurst et al., 2009). The second question draws from prior findings that attractiveness can yield either benefits or penalties depending on job visibility, evaluator gender, and industry norms. (Agthe et al., 2010; Agthe et al., 2011; Buss & Haselton, 2005; Chiu & Babcock, 2002; Dion et al., 1972; Finkelstein et al., 2015; Timming et al., 2017; Warhurst et al., 2009)

To address these questions, the present study evaluates human judgments of hyper-realistic facial images to assess perceived suitability for different occupational roles. Based on the literature reviewed above, there were two hypotheses that this study targeted.

H1: Facial attributes associated with appearance and aesthetics, such as *attractive*, *well-groomed*, *familiar*, *memorable*, *age*, and *trustworthy*, will be more strongly associated with perceived suitability for client-facing roles than for back-office roles.

H2: Facial attributes not directly related to attractiveness, such as *smart* and *dominant*, will be more strongly associated with perceived suitability for back-office roles than for client-facing roles.

Drawing inspiration from the modeling approach developed by Peterson et al. (2022), this study trained predictive models to estimate job suitability based on facial features. These models also enabled systematic manipulation of facial cues in the latent space, providing insight into how visual traits relate to occupational judgments. Additionally, predicted trait scores from Peterson et al. (2022)’s model were used to examine correlations between facial impressions and job suitability.

3 Methodology

3.1 Methods

3.1.1 Stimuli

In the present study, a set of synthetic, but hyper-realistic facial images were adapted from the stimuli used by Peterson et al. (2022). The original dataset comprised 1,004 synthetic face images generated using StyleGAN2, a generative adversarial network trained to produce high-fidelity human faces. These images were conditioned on a latent representation derived from a 512-dimensional vector, which was broadcast across the 18 layers of the StyleGAN2 generator, following the latent space structure described in Peterson et al. (2022). To maintain high visual quality and diversity, the images were carefully curated to exclude artifacts.

Given the study’s focus on job suitability evaluations, facial images depicting children and those involving sunglasses were excluded from the original dataset. As a result, the final stimulus set included 817 facial images of adults.

To ensure that participants could complete the experiment, including rating all assigned images, within approximately 20 minutes, the entire set of 817 images was randomly divided into six subsets, each containing roughly 136 images. Within each subset, an additional 30 images were randomly sampled from the 136 selected faces to be repeated for the purpose of assessing test-retest reliability. These repeated images were added back to the subset,

resulting in six final groups of approximately 166 images each.

3.1.2 Job Conditions

This study included four job roles selected from two major industries—finance and retail—to examine how facial appearance interacts with occupational expectations and social stereotypes in hiring judgments. These two industries were selected for both their functional contrast and their differing gender associations in public perception. Finance is often perceived as a male-dominated industry, commonly associated with high-status, analytical, and risk-oriented work. Prior research has found that finance roles systematically favor male applicants (Kline et al., 2024). In contrast, retail roles tend to be viewed as more associated with women (Kline et al., 2022).

For each industry, one client-facing and one back-office position were selected to further vary the level of interpersonal demands and public visibility in the role. The client-facing roles included Financial Advisor (finance) and Retail Sales Associate (retail), while the back-office roles were Risk Analyst (finance) and Inventory Manager (retail) (Indeed Employer Content Team, 2025a, 2025b, 2025c, 2025d). These roles were identified and defined based on publicly available job descriptions from Indeed.com and were selected to reflect realistic hiring contexts that vary in both function and the social impressions they might elicit based on facial appearance. Indeed is a widely used job search engine that offers job descriptions across a broad range of industries.

3.1.3 Participants

Participants were recruited via Prolific, an established online research platform that supports high-quality data collection from diverse and reliable participant pools. The required sample size was determined by the study design, which specified that each of the 817 facial images should be rated by at least 20 unique participants across each of the four job conditions. Given that each participant rated 166 images, the minimum required sample size was calculated as: $(817 \text{ images} \times 4 \text{ job conditions} \times 20 \text{ ratings}) \div 166 \text{ images per participant} = 394 \text{ participants}$. This recruitment target ensured adequate coverage across all images and experimental conditions. The participants were compensated according to Prolific’s fair pay guidelines; Based on an estimated task duration of approximately 20 minutes, each participant received \$4.00 for their participation.

Test-Re-Test Reliability. To assess response consistency, each participant rated 30 repeated images. The test-retest correlation for each participant by correlating their ratings of the original and repeated images was calculated. Across all 424 participants who completed the study, the mean test-retest correlation was $r = 0.657$ ($SD = 0.260$), with values

ranging from $r = -0.317$ to $r = 1$. When categorizing by four job conditions, the average test-retest correlations were highest for *Retail Sales Associate* ($r = 0.713$, $SD = 0.237$), followed by *Financial Advisor* ($r = 0.670$, $SD = 0.228$), *Inventory Manager* ($r = 0.624$, $SD = 0.313$), and *Risk Analyst* ($r = 0.615$, $SD = 0.249$).

To ensure data reliability in subsequent analyses, participants with zero or negative test-retest correlations were excluded. After this exclusion, 376 participants remained in the final analytic sample, with a slightly higher overall average test-retest correlation of $r = 0.681$ ($SD = 0.222$). At the job-condition level, mean correlations increased across all roles, with *Retail Sales Associate* showing the highest reliability ($r = 0.732$, $SD = 0.197$), followed by *Financial Advisor* ($r = 0.685$, $SD = 0.203$), *Inventory Manager* ($r = 0.659$, $SD = 0.274$), and *Risk Analyst* ($r = 0.645$, $SD = 0.205$).

Among these 375 participants, 228 identified as female, 126 as male, 7 as non-binary or third gender, and 1 chose not to disclose their gender. The most common age group was 25–34 years old (98), followed by 35–44 (80), 45–54 (66), 18–24 (58), 55–64 (38), and 65 or above (22).

In terms of educational attainment, the majority held a bachelor’s degree (139), followed by those with some college education (123), graduate degrees (48), high school diplomas (45), and associate degrees (4). Regarding ethnicity, the sample was predominantly White or Caucasian (220), with additional representation from Black or African American (62), Asian (37), Hispanic or Latino (23), individuals of multiple ethnicities (13), American Indian or Alaskan Native (6), and Pacific Islander (1).

3.1.4 Design and Procedure

The survey was administered using Qualtrics, an online platform for building and distributing web-based experiments. Upon accessing the study, participants first viewed an informed consent form outlining the purpose of the study, participation requirements, confidentiality assurances, and compensation details. Only participants who provided consent were allowed to proceed.

Each participant was randomly assigned to one of four job conditions, corresponding to one of the four target roles (*Financial Advisor*, *Risk Analyst*, *Retail Sales Associate*, or *Inventory Manager*). The assigned condition included both the job title and a detailed job description derived from real-world postings on Indeed.com. To ensure comprehension of the assigned role, participants completed a manipulation check, in which they were asked to identify a job title that correctly matched the description they had just read. If they answered incorrectly on the first attempt, the job description was presented again, followed by a second manipulation check. Participants who answered incorrectly on the second attempt were screened out and did not proceed with the remainder of the survey.

Participants who passed the manipulation check were then randomly assigned to one of six stimuli groups, each containing 166 hyper-realistic facial images. Images were presented one at a time in random order. Beneath each image, participants responded to the prompt: “To what extent do you think this person is suitable for a [job position] role?” using a sliding scale ranging from 0 (not at all suitable) to 100 (extremely suitable). Participants were required to provide a response for each image before proceeding to the next.

Upon completing all image ratings, participants answered four demographic questions concerning their gender, age, education level, and ethnicity.

3.2 Data

3.2.1 Data Processing for Modeling

The final analytic sample included responses from 376 participants, each of whom rated 166 hyper-realistic facial images under one of four assigned job conditions: *Financial Advisor*, *Risk Analyst*, *Retail Sales Associate*, or *Inventory Manager*. Prior to modeling, irrelevant metadata and administrative variables were removed, and each participant’s ratings were matched to corresponding image identifiers.

Mean suitability scores were then calculated for each of the 817 unique facial stimuli within each job condition. To facilitate downstream modeling, the dataset was reshaped into long format, where each row represented a stimulus–condition pair and included the raw average suitability rating and a standardized z-score that was computed within each condition.

3.3 Modeling and Synthetic Face Analysis

To quantify how facial features contribute to perceived job suitability across different occupational contexts, a series of predictive models were trained using the aggregated dataset of human judgments described in the previous section. Each image was represented by a 512-dimensional latent vector extracted from a pretrained StyleGAN2 generator (as used in Peterson et al. (2022)) and paired with its average suitability rating under one of four job roles.

For each job role, *Financial Advisor*, *Risk Analyst*, *Retail Sales Associate*, and *Inventory Manager*, a separate regularized linear regression model was fitted to predict human suitability ratings from the latent features. Model training used 10-fold cross-validation to search for optimal values of the regularization parameter. Performance was evaluated using cross-validated R^2 scores, which reflect the proportion of variance in human ratings explained by the model. Specifically, R^2 values were 0.59 for *Retail Sales Associate*, 0.47 for *Risk Analyst*, 0.37 for *Financial Advisor*, and 0.11 for *Inventory Manager*.

To examine the generalizability of these models beyond the originally rated faces, the same StyleGAN2 generator was used to synthesize 600 novel facial images (Peterson et al., 2022). These images were created by randomly sampling from the latent space of the model (Peterson et al., 2022).

The newly generated images were manually screened to remove those that contained visual artifacts, appeared unrealistic, depicted children, or featured occlusions such as sunglasses—features that were excluded from the training set. After filtering, 537 images remained and were deemed suitable for inference. The latent values for these synthetic faces were then used to predict job suitability ratings using the models described above. Additionally, predicted values for the 34 facial attributes (e.g., attractiveness, dominance, health, race) were obtained for these images using the psychological trait encoder developed by Peterson et al. (2022).

This resulted in a dataset containing 38 predicted variables per face, including 34 psychological traits and 4 job-related suitability scores. The full correlation heatmap of all 38 attributes is presented in Figure 3 in the Appendix. Rather than analyzing all attributes, only a set of attractiveness- and competence-based traits was selected to test the hypotheses. For Hypothesis 1, attractiveness-based attributes including *attractive*, *well-groomed*, *familiar*, *memorable*, *age*, *trustworthy*, *gender*, *happy*, and *outgoing* were selected. For Hypothesis 2, *smart* and *dominant* were selected as competence-related attributes, along with the four job suitability scores: *retail sales associate*, *financial advisor*, *inventory manager*, and *risk analyst*.

Pearson correlation coefficients were then computed between the selected facial attributes and the predicted job suitability scores to explore appearance-based patterns of occupational fit. To assess the structural similarity of trait–suitability relationships across roles, a correlation-of-correlations analysis was also conducted by comparing the correlations between different job roles.

4 Results

4.1 Pre-Model: Descriptive Distribution of Suitability Ratings

Based on the survey responses, the mean suitability score assigned to each facial image was calculated for each of the four job conditions. Figure 1 presents kernel density plots visualizing the distribution of these scores across two industries (Finance vs. Retail) and by job role (Client-facing vs. Back-office).

Within the finance industry, the distributions for *Financial Advisor* and *Risk Analyst* were tightly clustered and highly similar, with a slightly higher mean suitability observed for the client-facing role of *Financial Advisor* (55.97 vs. 55.25, respectively).

By contrast, the retail industry showed a more pronounced separation. The client-facing role of *Retail Sales Associate* received a higher mean suitability score (61.4) than the back-office role of *Inventory Manager* (56.12), indicating a greater divergence in perceived fit between customer-facing and back-end positions in retail.

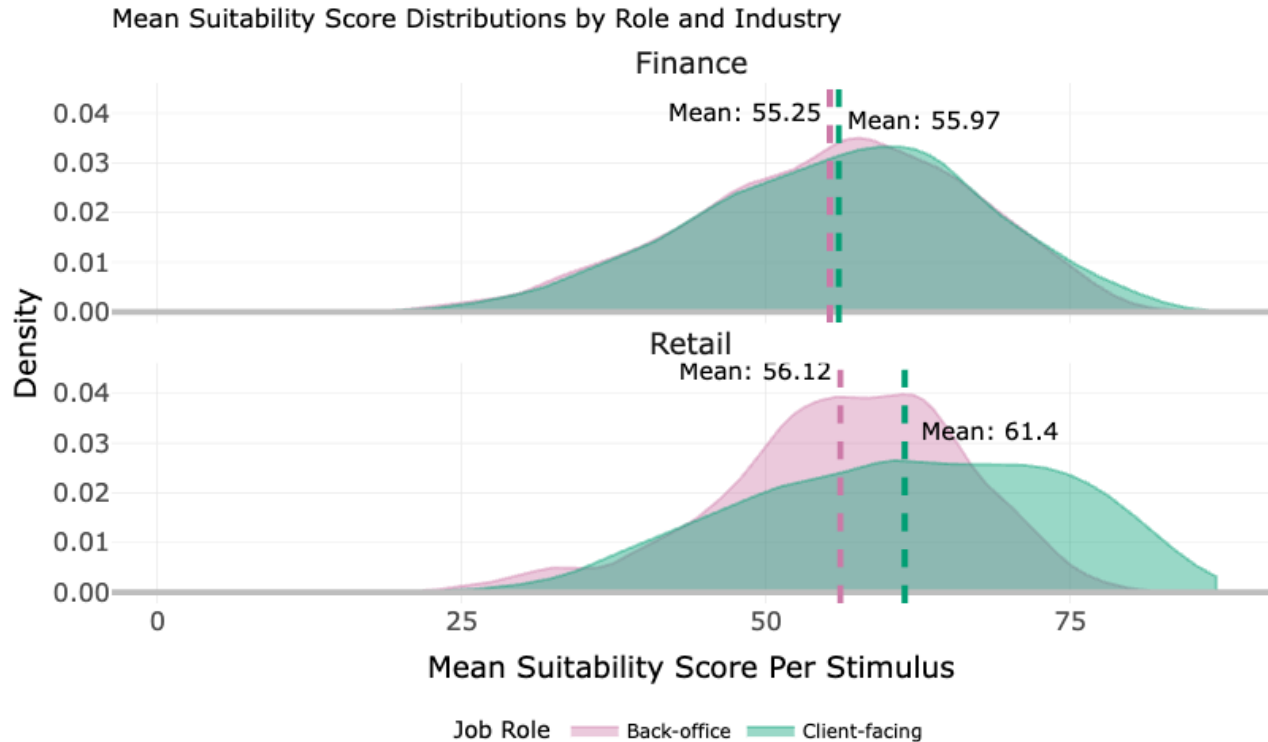


Figure 1: Distribution of mean suitability scores across job roles. Kernel density plots were shown to display the distribution of the per-face average ratings by role and industry. Dashed vertical lines indicate group means for client-facing and back-office roles within each industry.

4.2 Model: Visualizing Latent Suitability Representations

To provide a more intuitive understanding of how the model operationalizes job suitability across roles, facial transformations were generated by systematically manipulating the StyleGAN2 latent vectors along the predicted suitability dimension. For each job condition, *Retail Sales Associate*, *Financial Advisor*, *Inventory Manager*, and *Risk Analyst*, a subset of the StyleGAN2-generated faces were transformed across seven standard deviation points (−6 to +6) based on the predictive model’s estimated suitability scores.

Figures 4, 5, 6, and 7 in Appendix present example visualizations of the same two male and two female faces per job condition. Each row illustrates how the same synthetic face changes as its predicted suitability increases or decreases along the latent space. These visualizations offer an interpretable reference for how the model encodes occupational fit in appearance-based terms.

Visual inspection of the manipulated faces revealed several interesting effects. Notably, high predicted suitability for the client-facing *Retail Sales Associate* role tended to yield facial expressions that appeared more expressive, youthful, and socially engaging. In contrast, transformations along the suitability dimension for back-office roles such as *Risk Analyst* emphasized features commonly associated with competence, maturity, and authority, such as increased age. These qualitative transformations visualized the human judgments of higher or lower suitability regarding these four jobs. These patterns are further examined through correlation analyses between inferred traits and predicted suitability scores.

4.3 Post-Model: Trait–Suitability Correlations Across Job Roles

While the pre-model analysis presented descriptive patterns based on participants’ responses, the following analysis turns to the model-predicted data to examine how specific facial attributes relate to perceived job suitability across different roles. Correlation heatmaps were constructed to evaluate the relationship between selected facial traits and the predicted suitability scores within each job condition: *Financial Advisor*, *Risk Analyst*, *Retail Sales Associate*, and *Inventory Manager*.

As shown in Figure 2, each heatmap displays Pearson correlation coefficients between a subset of facial traits and the predicted suitability scores. Within each panel, traits are ordered from top to bottom by the strength of their correlation with the specific job role, which supports clearer comparisons of which attributes are most influential in each context.

For the client-facing *Retail Sales Associate* role, strong positive correlations were observed with many appearance-oriented traits such as *trustworthy* ($r = .84$), *attractive* ($r = .81$), *outgoing* ($r = .75$), *happy* ($r = .70$), *well-groomed* ($r = .69$), and *memorable* ($r = .52$). Traits negatively associated with this role included *dominant* ($r = -.78$), *gender* ($r = -.79$), and *age* ($r = -.59$). The predicted values for *smart* ($r = .029$) were not correlated with this role.

For the client-facing *Financial Advisor* role, positive correlations were found with *smart* ($r = .84$), *age* ($r = .66$), and *familiar* ($r = .45$), while negative correlations were observed with *memorable* ($r = -.41$) and *attractive* ($r = -.28$).

For the back-office *Inventory Manager* role, positive associations included *smart* ($r = .76$), *age* ($r = .60$), *dominant* ($r = .52$), *gender* ($r = .43$), and *familiar* ($r = .38$). The role was negatively correlated with *memorable* ($r = -.50$) and *attractive* ($r = -.37$).

For the back-office *Risk Analyst* role, strong positive correlations were seen with *smart* ($r = .85$), *age* ($r = .70$), *dominant* ($r = .53$), *gender* ($r = .46$), and *familiar* ($r = .33$). Negative associations included *memorable* ($r = -.58$) and *attractive* ($r = -.46$).

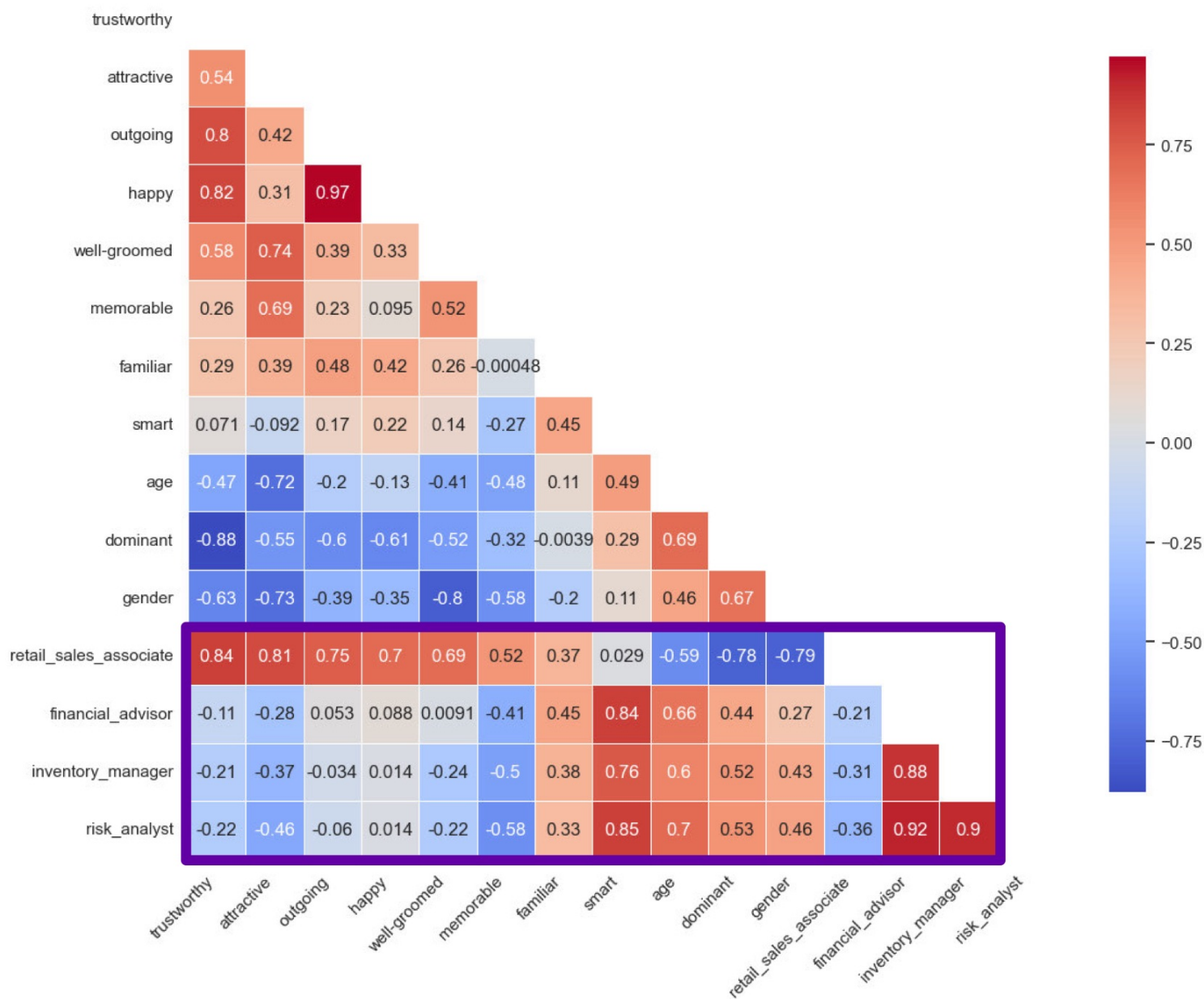


Figure 2: Correlation heatmaps showing the relationship between selected facial attributes and predicted job suitability scores across four job roles.

To quantify the degree of similarity in trait–suitability structures across roles, Spearman correlations were computed between the vectors of Pearson correlation coefficients for each pair of job conditions. The three roles of *Financial Advisor*, *Inventory Manager*, and *Risk Analyst* exhibited highly similar correlation profiles, with all pairwise comparisons yielding $\rho = 0.96$. This result suggests a consistent underlying trait structure for suitability judgments across these positions, which emphasizes competence-related impressions such as perceived smartness and dominance. In addition, the consistently high correlations with age across these roles suggest a preference for faces perceived as more experienced, reinforcing the preference over competence.

By contrast, the *Retail Sales Associate* role displayed a markedly different pattern. It showed stronger positive associations with appearance-based and socially expressive traits, including *attractiveness* ($r = .81$), *trustworthiness* ($r = .84$), *outgoingness* ($r = .75$), and *happiness* ($r = .70$). At the same time, it exhibited strong negative correlations with *dominance* and *age*, traits positively associated with the other three roles. These differences indicate that suitability for this client-facing retail role is more closely tied to perceived warmth and approachability, rather than analytical or authoritative characteristics. This divergence highlights how aesthetic and emotional expressiveness may play a disproportionately role in shaping judgments in highly visible, customer-oriented contexts.

5 Discussion

The present study set out to investigate how facial attributes—particularly attractiveness and other inferred social traits—shape perceptions of job suitability across client-facing and back-office roles in two distinct industries. Two hypotheses were proposed: (H1) that aesthetic and warmth-related traits would be more strongly associated with client-facing roles, and (H2) that competence-related traits would be more strongly associated with back-office roles.

The findings offer partial support for H1. As expected, the client-facing role of *Retail Sales Associate* was strongly associated with appearance-oriented traits such as *attractiveness*, *well-groomed*, *trustworthy*, and *happy*, with several correlations exceeding ($r = .70$). This aligns with the literature suggesting that interpersonal expectation heightens the relevance of aesthetic and emotionally expressive features (Timming et al., 2017; Warhurst et al., 2009).

However, the other client-facing role, *Financial Advisor*, did not show this pattern. Instead, it aligned more closely with the two back-office roles with significantly high correlation, all of which exhibited strong correlations with competence-related traits such as *smart*, *age*, and *dominant*, and modest negative correlations with *attractiveness*. This con-

vergence suggests a stable trait schema linked to analytical and operational job demands. In this regard, the findings provide strong support for H2, as both two back-office roles in two industries emphasized competence-related traits and shared a consistent evaluative structure. However, more broadly, the inclusion of *Financial Advisor* in this cluster indicates that job function alone may not fully explain appearance-based suitability judgments; instead, industry norms and functional expectations appear to play a more determinative role.

A similar pattern emerged in the correlations with perceived gender. The role of *Retail Sales Associate* exhibited a strong negative correlation with the gender dimension, which indicated a strong visual association with feminine features. This aligns with prior findings that some retail positions tend to favor female candidates or feminine-coded traits (Kline et al., 2022). In contrast, the other retail role—*Inventory Manager*—was positively correlated with perceived gender, much like the two finance roles, indicating a moderate preference for masculine traits. These patterns suggest that gendered expectations are also not solely driven by industry, but also by the functional nature of the role and its alignment with stereotypical traits of competence, authority, and sociability.

Before turning to the broader theoretical implications, it is worth noting that the predictive performance of the models varied across roles. In particular, the model trained to estimate suitability for the *Inventory Manager* role yielded a relatively low R^2 score of 0.11, compared to 0.59 for *Retail Sales Associate*, 0.47 for *Risk Analyst*, and 0.37 for *Financial Advisor*. However, the test–retest reliability of the *Inventory Manager* role was not significantly different from the correlations of the other three roles. This suggests that while participants were consistent in their individual evaluations of *Inventory Manager*, their judgments were highly variable across individuals. This implies a lack of shared visual stereotype for this role that resulted in weaker model performance. It further underscores that, beyond job function and facial traits, other contextual or idiosyncratic factors may influence perceived suitability for certain roles.

These findings expand on prior research by bringing the effect of appearance-based stereotypes and the function of job roles all together and demonstrating that appearance-based stereotypes are shaped by both job functions and industry-specific expectations. While earlier work has emphasized the beauty premium in different industries contexts (Agthe et al., 2010; Agthe et al., 2011; Chiu & Babcock, 2002; Hamermesh & Biddle, 1994), this study shows that the stereotype does not uniformly apply across all interpersonal roles. Instead, traits like competence and maturity may be more salient in professional services (e.g., finance), even in roles that involve client interaction.

The convergence of trait–suitability correlations for *Financial Advisor*, *Risk Analyst*, and *Inventory Manager* suggests a shared evaluative structure across finance and analytical

retail roles. This finding underscores the importance of incorporating both role function and industrial culture when theorizing about visual stereotypes in hiring.

Limitations and Future Directions

Several limitations should be noted. First, the discussion above mentioned the uneven model performance across four job roles, with the cross-validated R^2 for *Inventory Manager* was substantially lower than for the other three positions. This might be because the selected job roles may not fully represent the broader diversity of positions within each industry or job category. For example, *Retail Sale Associate* and *Inventory Manager* capture only a narrow slice of retail roles, while *Financial Advisor* and *Risk Analyst* reflect just a subset of finance occupations. Similarly, the specific client-facing and back-office roles selected in this study may not generalize to other roles of the same job type within each industry. For instance, the interpersonal demands of a Financial Advisor may differ substantially from those of other client-facing roles in finance.

Second, the study focused exclusively on the retail and finance industries for a clear functional contrast and manageable scope. However, this limited industry scope restricts the generalizability of the findings, as appearance-based stereotypes may manifest differently in other industries such as healthcare, education, or technology.

Finally, while this study used advanced modeling and synthetic imagery to control for visual content, actual hiring decisions occur in more social and interactive contexts. Combining behavioral or interview-based simulations could provide additional valid insights into how facial impressions affect suitability judgments.

6 Conclusion

This study investigated how facial appearance influences perceived job suitability across different occupational roles and industries. By combining human judgment data with predictive models on synthetic faces, the analysis revealed that visual cues shaping occupational evaluations are nuanced and context dependent. While the “what is beautiful is good” attractiveness stereotype held true for certain client-facing roles, most notably the *Retail Sales Associate*, it did not generalize across all client-facing positions. Although the *Financial Advisor* is considered client-facing, it shared more in common with the back office roles in terms of its emphasis on competence, dominance, maturity and masculinity rather than aesthetic appeal. This divergence suggests that the role of facial attractiveness and other perceived social traits in professional judgments may be better explained by industry-specific norms than by functional job categories alone.

In general, the findings challenge the notion of a universal beauty premium. They highlight that other traits inferred from faces, including smartness, trustworthiness, gender, age, and even industry expectation, can shape hiring perceptions in ways that go beyond attractiveness alone. By modeling and visualizing these visual stereotypes, the study advances our understanding of how visual impressions affect perceived job fit.

These insights have both theoretical and practical implications. Theoretically, they offer diverse perspectives for studying visual stereotypes in hiring. From a practical thinking, they emphasize the need for organizations to critically assess how appearance-based expectations can reinforce bias, especially in roles where visual presentation is emphasized. This study also draws organizations' attention to the potential for implicit visual biases to disadvantage certain groups and underscores the importance of designing evaluation processes that promote diversity, equity, and inclusion through more objective hiring practices.

Data and Code Availability Statement

The survey dataset and preprocessing scripts used in this study are available at [this GitHub repository](#). This includes participant responses, demographic summaries, and all code used for data cleaning and formatting.

The facial image generation and trait modeling components rely on code and pretrained models developed by Peterson et al. (2022), which are publicly available at [this GitHub repository](#). For additional details on the modeling architecture and trait prediction methods, see Peterson et al. (2022).

References

- Agthe, M., Spörrle, M., & Maner, J. K. (2010). Don't hate me because I'm beautiful: Anti-attractiveness bias in organizational evaluation and decision making. *Journal of Experimental Social Psychology, 46*(6). <https://doi.org/10.1016/j.jesp.2010.05.007>
- Agthe, M., Spörrle, M., & Maner, J. K. (2011). Does being attractive always help? positive and negative effects of attractiveness on social decision making. *Personality and Social Psychology Bulletin, 37*(8). <https://doi.org/10.1177/0146167211410355>
- Bal, A. C., Reiss, A. E. B., Rudolph, C. W., & Baltes, B. B. (2011). Examining positive and negative perceptions of older workers: A meta-analysis. *The Journals of Gerontology: Series B, 66B*(6), 687–698. <https://doi.org/10.1093/geronb/gbr056>
- Buss, D. M., & Haselton, M. (2005). The evolution of jealousy. <https://doi.org/10.1016/j.tics.2005.09.006>
- Chiu, R. K., & Babcock, R. D. (2002). The relative importance of facial attractiveness and gender in Hong Kong selection decisions. *International Journal of Human Resource Management, 13*(1). <https://doi.org/10.1080/09585190110092857>
- Davis, L. L., & Lennon, S. J. (1988). Social cognition and the study of clothing and human-behavior. *Social Behavior and Personality, 16*, 175–186.
- Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*(3). <https://doi.org/10.1037/h0033731>
- Finkelstein, L. M., Burke, M. J., & Raju, M. S. (1995). Age discrimination in simulated employment contexts: An integrative analysis. *Journal of Applied Psychology, 80*(6), 652. <https://doi.org/10.1037/0021-9010.80.6.652>
- Finkelstein, L. M., King, E. B., & Voyles, E. C. (2015). Age meta-stereotyping and cross-age workplace interactions: A meta view of age stereotypes at work. *Work, Aging and Retirement, 1*(1), 26–40. <https://doi.org/10.1093/workar/wau002>
- Frieze, I. H., Olson, J. E., & Good, D. C. (1990). Perceived and Actual Discrimination in the Salaries of Male and Female Managers. *Journal of Applied Social Psychology, 20*(1). <https://doi.org/10.1111/j.1559-1816.1990.tb00377.x>
- Frost, P. (2014). The Puzzle of European Hair, Eye, and Skin Color. *Advances in Anthropology, 04*(02). <https://doi.org/10.4236/aa.2014.42011>
- Hamermesh, D. S., & Biddle, J. E. (1994). Beauty and the Labour Market. *The American Economic Review, 84*(5).
- Heineck, G. (2005). Up in the skies? The relationship between body height and earnings in Germany. <https://doi.org/10.1111/j.1467-9914.2005.00302.x>
- Ho, L. S., Wei, S. D., & Voon, J. P. (1999). Age discrimination in the labour market: Fact or fiction? *International Employment Relations Review, 5*, 43–52.

- Ho, L. S., Wei, X. D., & Voon, J. P. (2000). Are old workers disadvantaged in the hong kong labour market? *Asian Economic Journal*, *14*(3), 283–300. <https://doi.org/10.1111/1467-8381.00112>
- Indeed Employer Content Team. (2025a). *Financial advisor job description*. Retrieved April 21, 2025, from <https://www.indeed.com/hire/job-description/financial-advisor?co=US>
- Indeed Employer Content Team. (2025b). *Inventory manager job description*. Retrieved April 21, 2025, from <https://www.indeed.com/hire/job-description/inventory-manager?co=US>
- Indeed Employer Content Team. (2025c). *Retail sales associate job description*. Retrieved April 21, 2025, from <https://www.indeed.com/hire/job-description/retail-sales-associate?co=US>
- Indeed Employer Content Team. (2025d). *Risk analyst job description*. Retrieved April 21, 2025, from <https://www.indeed.com/hire/job-description/risk-analyst>
- Judge, T. A., & Cable, D. M. (2004). The effect of physical height on workplace success and income: Preliminary test of a theoretical model. *Journal of Applied Psychology*, *89*(3). <https://doi.org/10.1037/0021-9010.89.3.428>
- Kline, P., Rose, E. K., & Walters, C. R. (2022). Systemic discrimination among large u.s. employers. *The Quarterly Journal of Economics*, *137*(4), 1963–2036. <https://doi.org/10.1093/qje/qjac024>
- Kline, P., Rose, E. K., & Walters, C. R. (2024). A discrimination report card. *American Economic Review*, *114*(8), 2472–2525. <https://doi.org/10.1257/aer.20230700>
- Korthase, K. M., & Trenholme, I. (1982). Perceived age and perceived physical attractiveness. *Perceptual and Motor Skills*, *54*(3-suppl), 1251–1254. <https://doi.org/10.2466/pms.1982.54.3c.1251>
- Oh, D., Buck, E. A., & Todorov, A. (2019). Revealing hidden gender biases in competence impressions of faces. *Psychological Science*, *30*(1), 65–79. <https://doi.org/10.1177/0956797618813092>
- Oh, D., Shafir, E., & Todorov, A. (2020). Economic status cues from clothes affect perceived competence from faces. *Nature Human Behaviour*, *4*(3), 287–293. <https://doi.org/10.1038/s41562-019-0782-4>
- Oh, D., Wedel, N., Labbree, B., & Todorov, A. (2023). Trustworthiness judgments without the halo effect: A data-driven computational modeling approach. *Perception*, *52*(8), 590–607. <https://doi.org/10.1177/03010066231178489>
- Peterson, J. C., Uddenberg, S., Griffiths, T. L., Todorov, A., & Suchow, J. W. (2022). Deep models of superficial face judgments. *Proceedings of the National Academy of*

-
- Sciences of the United States of America*, 119(17). <https://doi.org/10.1073/pnas.2115228119>
- Rivera, L. A. (2012). Hiring as cultural matching: The case of elite professional service firms. *American Sociological Review*, 77(6), 999–1022. <https://doi.org/10.1177/0003122412463213>
- Ruffle, B. J., & Shtudiner, Z. (2015). Are good-looking people more employable? *Management Science*, 61(8). <https://doi.org/10.1287/mnsc.2014.1927>
- Solomon, M. R. (2022). *Consumer Behavior Buying, Having, and Being Thirteenth Edition* (Vol. 53). Pearson.
- Timming, A. R., Nickson, D., Re, D., & Perrett, D. (2017). What do you think of my ink? assessing the effects of body art on employment chances. *Human Resource Management*, 56(1), 133–149. <https://doi.org/10.1002/hrm.21770>
- Todorov, A. (2012). The social perception of faces. In S. T. Fiske & C. N. Macrae (Eds.), *The sage handbook of social cognition* (pp. 248–264). SAGE Publications Ltd. <https://doi.org/10.4135/9781446247631.n6>
- Warhurst, C., van den Broek, D., Hall, R., & Nickson, D. (2009). Lookism: The new frontier of employment discrimination? *Journal of Industrial Relations*, 51(1), 131–136. <https://doi.org/10.1177/0022185608096802>
- Zhao, H., O'Connor, G., Wu, J., & Lumpkin, G. T. (2021). Age and entrepreneurial career success: A review and a meta-analysis. *Journal of Business Venturing*, 36(1). <https://doi.org/10.1016/j.jbusvent.2020.106007>

Appendix

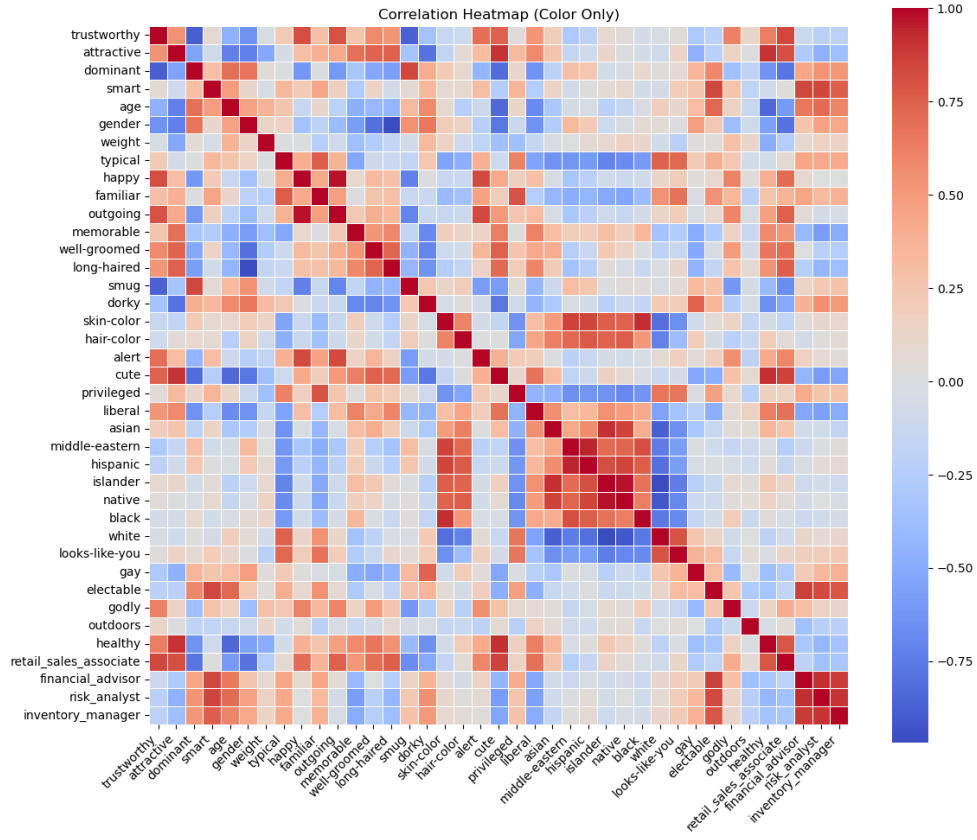


Figure 3: Correlation Matrix. It presents the heatmap of the correlations between four job attributes and 34 facial attributes Peterson et al. (2022).

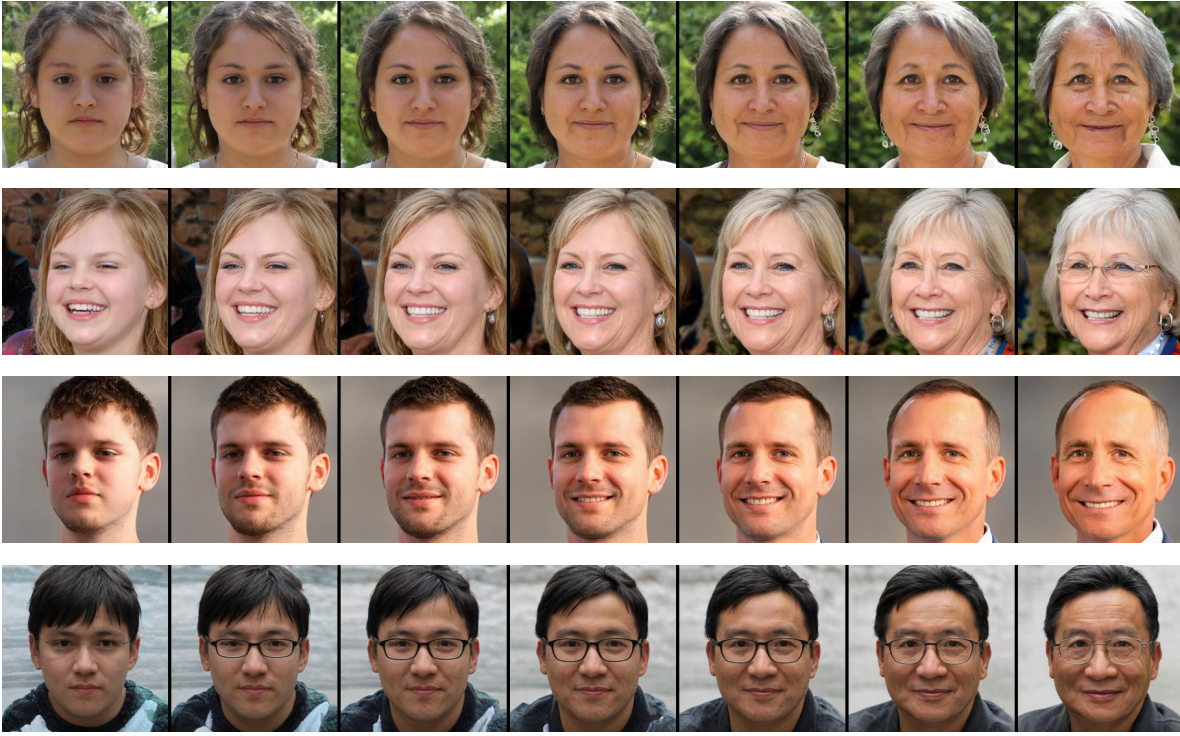


Figure 4: Financial Advisor - client-facing role in Finance

Two male and two female hyper-realistic faces generated by StyleGAN2, transformed along the predicted suitability dimension for the role of Financial Advisor. Images represent -6 , -4 , -2 , 0 , $+2$, $+4$, and $+6$ standard deviations.

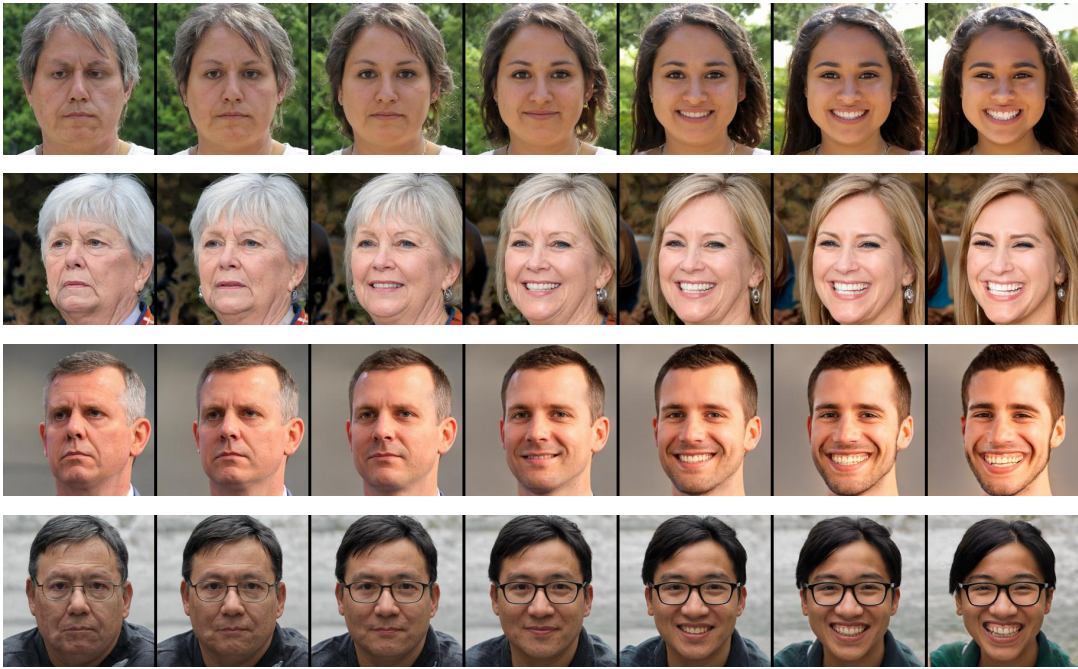


Figure 5: Retail Sale Associate - client-facing role in Retail

Two male and two female hyper-realistic faces generated by StyleGAN2, transformed along the predicted suitability dimension for the role of Retail Sale Associate. Images represent -6 , -4 , -2 , 0 , $+2$, $+4$, and $+6$ standard deviations.

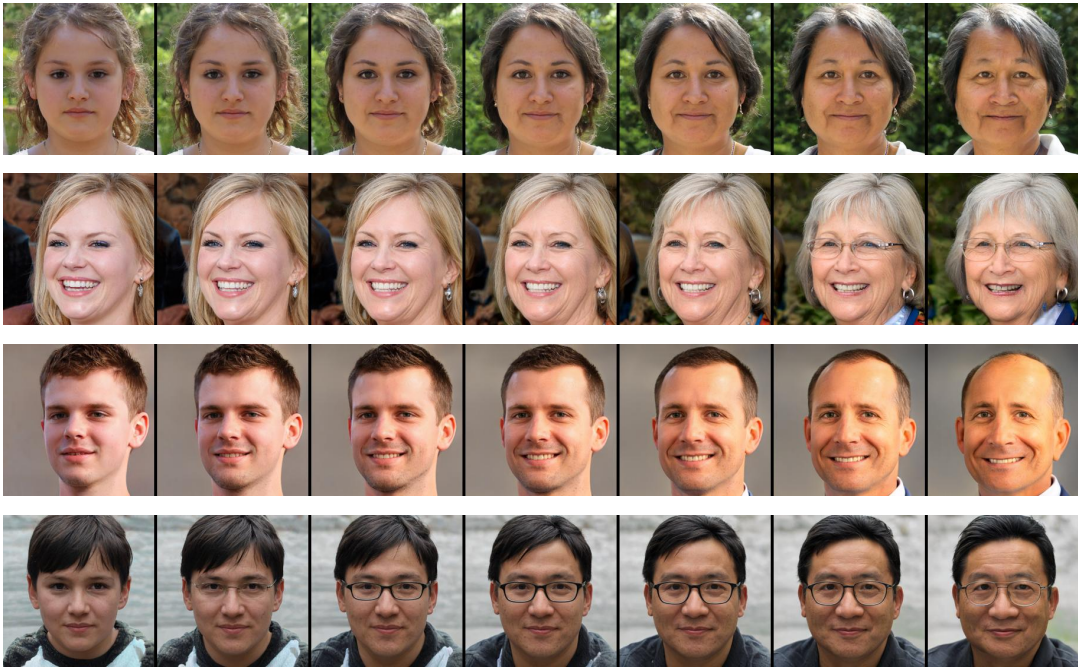


Figure 6: Risk Analyst - back-office role in Finance

Two male and two female hyper-realistic faces generated by StyleGAN2, transformed along the predicted suitability dimension for the role of Risk Analyst. Images represent -6 , -4 , -2 , 0 , $+2$, $+4$, and $+6$ standard deviations.

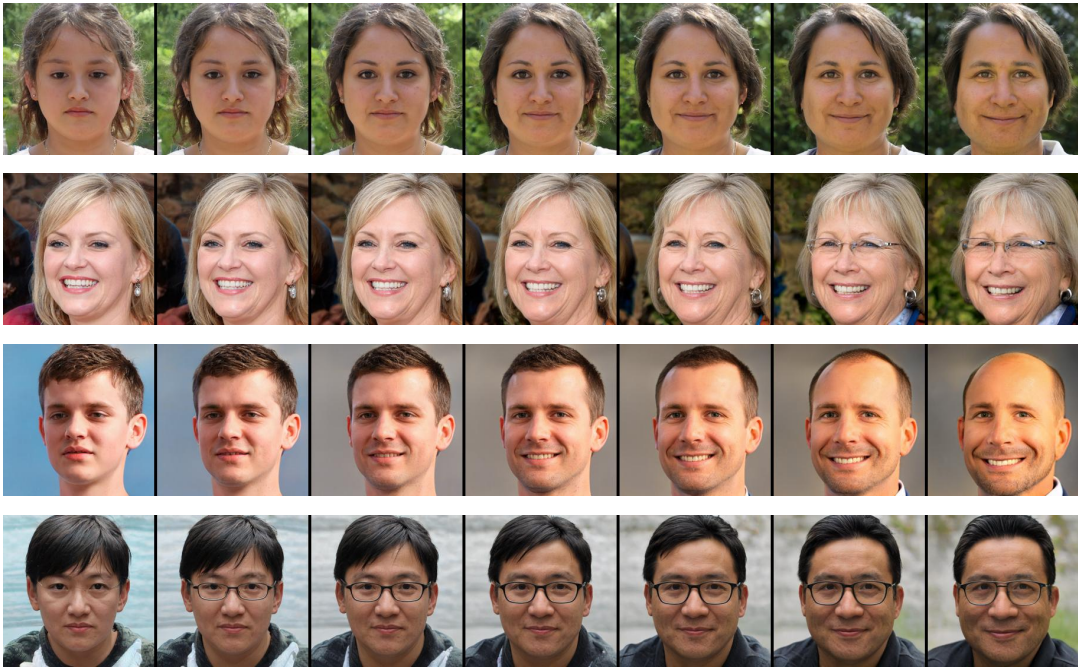


Figure 7: Inventory Manager - back-office role in Retail

Two male and two female hyper-realistic faces generated by StyleGAN2, transformed along the predicted suitability dimension for the role of Inventory Manager. Images represent -6 , -4 , -2 , 0 , $+2$, $+4$, and $+6$ standard deviations.