

Prejudices of the Selection Process: The Interaction of Attractiveness, Phenotype, and
Qualifications in Women

Samantha R. Stone

Tufts University

Author Note

Samantha R. Stone, Department of Psychology, Tufts University.

Thanks should be given to Keith Maddox for advising the experiment, Joe DeBold, Sam Sommers, and to the rest of the Tufts Social Cognition Lab for all of their invaluable insights.

Correspondence concerning this article should be addressed to Samantha R. Stone, Department of Psychology, Tufts University, 490 Boston Avenue, Medford, Massachusetts 02155. Email: Samantha.stone@tufts.edu

Abstract

Previous research has found that unattractive and black individuals are discriminated against in the job selection process. In the form of two online surveys, the current experiment examined the combined effects of racial phenotype, attractiveness, and qualifications on participants' impressions of female candidates. Each experiment included a questionnaire and 12 mock-up LinkedIn profiles of two women from each category: Attractive white, unattractive white, attractive light-skinned black, unattractive light-skinned black, attractive dark-skinned black, and unattractive dark-skinned black. Experiment 1 aimed to investigate, within the Tufts University undergraduate population ($N = 67$), the interaction between racial phenotype and attractiveness. Experiment 2 explored, in a more general sample ($N = 188$), the interaction between racial phenotype, attractiveness, and level of qualification. Consistent with hypotheses, analyses revealed significant interactions for both studies; however the results were not in the predicted directions. This implies that while the three variables individually affect employers' judgments, when combined they are more effective.

Keywords: racial phenotypicality bias, physical attractiveness, prejudice in selection

Prejudices of the Selection Process: The Interaction of Attractiveness, Phenotype, and Qualifications in Women

Each day, hundreds of people go on interviews. There are millions of articles and books about mastering and preparing for the interview. Despite how many books a person buys or how many hours he or she spends researching online, there are still so many factors within the interview that can affect the outcome. A person may be the most qualified for the job, but may not be offered the position because someone else had a connection or a more aggressive approach. Unfortunately, discrimination against physical and racial characteristics is also a huge determinant during recruitment. With that in mind, people would greatly benefit by understanding the decision making process. However, to my knowledge, there are few studies that explore the combined effects of attractiveness, phenotype, and qualification during the selection process. The few experiments that have studied discrimination have often considered racial phenotype and attractiveness as one variable (Harrison & Thomas, 2009). And, the cross-sectional studies that do consider phenotype often do not include data for white individuals (Keith & Herring, 1991; Hill, 2000; Hunter, 2002). Furthermore, the few studies that explore the effects of racial phenotype and attractiveness do not do so in an interviewing setting (Parks & Kennedy, 2007; DeMeis & Turner, 1978; Ritts, Patterson, & Tubbs, 1992).

There are various factors that affect interviewer attitudes, such as racial phenotype, attractiveness, and qualifications. Previous findings suggest that the darker a person's skin color, the more likely he or she is regarded as incompetent (Harrison & Thomas, 2009). Furthermore, past studies have found that attractiveness is beneficial in most circumstances (Dion, Berscheid, & Walster, 1972; Benson, Karabenick, Lerner, 1976; Chaiken, 1979). It is generally accepted that when people have impressive credentials, they are considered more capable. The present

experiments explored the interaction of attractiveness, racial phenotype, and qualifications to examine their impact on perceived competence. The mentioned studies and findings will be discussed in upcoming sections of the paper.

Attractiveness

The role of attractiveness is supported by the “what is beautiful is good” theory in which attractive individuals are generally thought to have many other positive attributes (Dion et al., 1972). Research has found supporting evidence for this theory. Overall, attractive individuals are perceived as happier, more sociable, and more successful on average (Chaiken, 1979). As a result, others are more likely to help an attractive stranger that they’ve never met over an unattractive one (Benson, Karabenick, Lerner, 1976).

However, other research has found that attractiveness only helps when the assessor and candidate are of opposite sex. One study found that college students rate attractive members of the opposite sex more highly (Anderson & Nida, 2006). Also, in another study where participants took part in a mock negotiation scenario, female interviewers gave attractive male interviewees higher status job packages than average looking males and vice versa (Senior, Thomson, Badger, & Butler, 2007). However, despite such compelling evidence for the “what is beautiful is good” theory, additional research has yielded opposing theories when examining its interacting with gender. One such study concluded that when the person being evaluated is of the same sex as the interviewer, attractiveness can hurt rather than help (Agthe, Sporrle, & Maner, 2011).

Furthermore, other research examining the effects of attractiveness in Marketing suggests that attractiveness may be most influential when selling a physical product. For example, attractive salesmen often have the highest performance when consumers are forced to make snap

judgments (Ahearne, Gruenb, & Jarvis, 1996). Similarly, prior research has shown that attractive spokespeople are more persuasive when advertising a product or service (Caballero & Pride, 1984; Kahle & Homer, 1985). Additionally, consumers perceive attractive salesmen more favorably, are more responsive to their sales pitches, and purchase more from them than they do from unattractive salesmen (DeShields, Kara, & Kaynak, 1996). Finally, employee's physical attractiveness, displayed positive emotion, and helpfulness all positively influenced customer satisfaction (Keh, Ren, Hill, & Li, 2013). While prior research has also shown that attractiveness can be impactful by being linked to a product, studies have also found it to be nominally effective without such an association. For example, one such study found a small, positive effect on the impact of facial attractiveness on and judges' expertise about decisions in pre-selective recruitment (Dubois & Pansu, 2004). The study concludes that while attractiveness is taken into account, it surely isn't the most predominant factor during the selection process.

While evidence points to an overall positive correlation between attractiveness and perceived ability, there are strong implications that the interaction of gender and attractiveness can have a differing effect when combined. This demonstrated effect is most likely not limited to just gender and attractiveness.

Racial phenotype

Like attractiveness, it is common knowledge that racial minorities are often discriminated against, including in work settings. Although there has been extensive research on racial discrimination, former research on the correlation of racial phenotype and success has mostly been cross-sectional. Furthermore, the investigative studies only included information about differing black phenotypes and excluded data for white individuals. While they demonstrate that a lighter phenotype within the black race is advantageous, the studies don't allow comparison

with white individuals. Such a gap in analyses prevents us from seeing the data relative to other ethnic groups.

One study examined the success of deceased black men of either Black and White or Black decent (Hill, 2000). The findings reveal that subjects who identified as mulatto were of a higher socioeconomic class than those that identified as black. Similarly, another study drew the same conclusions when comparing the phenotype of black women (Hunter, 2002). The National Survey data showed that women of lighter skin tended to have a higher educational level, greater income earnings, and higher spousal status. Furthermore, data from the National Survey of Black Americans shows that a lighter skin tone increases the likelihood of educational attainment, personal income, socioeconomic status, and overall success (Keith & Herring, 1991). The studies suggest the importance of skin color in life outcome. Evidence shows that success is more strongly correlated with skin color than the upbringing or background of family.

While several studies are able to demonstrate the effects of racial phenotype-based discrimination, there are few to my knowledge that explore the causal effect. One such found that preferential treatment in the job-selection process is given to light over dark-skinned Blacks (Harrison & Thomas, 2009). In the between experiment, participants were given one of six photos and the same resume. The experimenters used a questionnaire about the resume to measure the participants' impressions. Findings show a negative correlation between darker phenotypes and perceived competence.

Like unattractive individuals, the hurdle of success for dark-skinned Blacks begins with the interview process, so it is important to acknowledge the prejudices from the beginning to improve onward.

Attractiveness and Race

One such study explored the impact of race, physical attractiveness, and gender on education majors' and teachers' perceptions of student competence (Parks & Kennedy, 2007). Participants rated photos having read short, neutral narratives that contained one mildly positive fact and one mildly negative fact about the student. They found that people rated unattractive, black students as the least competent. Findings suggest that physical appearance can often override credentials.

Similarly, another study assessed the effects of attractiveness, race, and dialect on teachers' evaluations (DeMeis & Turner, 1978). Teachers significantly rated unattractive, black students with "Black English" dialects as less competent than their classmates. The results showed that teachers across subjects had consistent ratings. The study suggests that these children's academic failures can be attributed to their superficial appearances rather than their actual academic performance. Furthermore, another study found that teachers treat attractive students more favorably than the rest of their peers (Ritts, Patterson, & Tubbs, 1992). Analyses show that teachers regard their attractive students are more intelligent, having great academic potential, and more socially capable.

After conducting the literature review, I have concluded that a person's qualifications will be the main predictor of their assessment. Consistent with common knowledge and assumption, studies have found that employee's displayed ability positively correlates with overall satisfaction (Keh et al., 2013). Research exploring the effects of attractiveness has yielded conflicting results. While most evidence points to a strong positive correlation between high regard and attractiveness, some studies suggest that there is a gender influence. In regards to racial phenotype, the cross-sectional studies consistently prove a negative correlation between a darker skin color and success. Experiments that have assessed the effects of both appearance and

phenotype have found an even more negative interaction. The studies cited in this sections show that, when combined, unattractiveness and a darker phenotype yield an even harsher rating.

Taking such reliable findings into account, I believed that the results would be transferrable to adult women in a professional setting.

Current Study

Two experiments aimed to investigate the combined effects of phenotype and attractiveness on perceived competence. In the form of two online surveys, both experiments showed participants 12 different LinkedIn profiles and asked to rate the candidates' ability. The primary prediction for the first experiment was that there would be an interaction between racial phenotype and attractiveness. When the candidates were attractive, I expected ratings to be higher for lighter-skinned women compared to darker-skinned women when candidates were of equal qualification. The experiment explored the interaction between attractiveness and racial phenotype. For the second experiment, the primary prediction was a three-way interaction between racial phenotype, attractiveness, and qualifications. I believed that women with better qualifications would be rated the highest regardless of attractiveness or racial phenotype. However, I anticipated the attractive, lighter-skinned women within the same qualification group will be rated more favorably. The experiment explored the interaction between attractiveness, phenotype, and qualification. Attractiveness, racial phenotype, and qualifications have a significant effect on perceived competence.

General Method

Overview

The experiments were in the form of two online surveys developed through Qualtrics (www.tufts.qualtrics.com), an online research software. In both experiments, I measured

participants' attitudes toward the candidates. For both experiments, there were three levels of phenotype: white, light-skinned black, and dark-skinned black; and there were two levels of attractiveness: attractive and unattractive. The second experiment had an additional variable, qualification, which had two levels: high and low. The survey presented all of the combinations of the three variables for a total of 12 profiles. It was a repeated measures design in which participants saw different profiles across conditions.

Stimuli

Pretest 1

The stimuli for my experiment were chosen after two rounds of pretesting (Figure 1). The first Pretest included 40 faces collected from Yahoo personals and Google. The first Pretest asked participants to use a 10-point scale to rate overall attractiveness, classification of ethnicity, and typicality of her skin tone for her ethnicity. There were 18 participants, 12 of them were female and 6 male $M_{age} = 20.94$, $SD = .42$, $Min = 20$, $Max = 21$. Of the 18 participants, 12 of them were White, 3 were Asian, one Black, one Hispanic, and one Multiracial. All of the participants were from a Tufts Social Psychology course and received extra credit for fulfilling the survey.

Pretest 2

The second Pretest included 22 of the faces from the initial Pretest by selecting faces of comparable phenotypic and attractiveness ratings. The survey asked participants to use a 10-point scale to rate the women's typicality of her skin tone for her ethnicity, general level of afrocentricity, age, level of smiling, and level of competency. The Pretest also had a component in which participants used a 10-point scale to rate statements which would later be compiled to create references. After all of the participants rated the statements, I calculated the means for each characteristic. The statements included positive characteristics like "worked nonstop" and

“lived and breathed work.” In addition, it also included negative traits like “didn’t go the extra mile” and “late to meetings.” After all of the traits were rated, I created the references: the positive references included two positive statements and one negative statement with the average ratings adding up to 20.5; the negative references included two negative statements and one positive statement with the average ratings adding up to 16.75; the neutral references included one positive and negative statement with the average ratings adding up to 12.5. There were 8 participants, 7 of them were female and 1 was male ($M_{age} = 20.5$, $SD = .71$, $Min = 20$, $Max = 22$). Of the 8 participants, 4 of them were White, 2 of them were Black, and 2 of them were Multiracial. All of the participants were Research Assistants in the Tufts Social Psychology lab.

Studies 1&2

My experiment consisted of one survey, which contained 12 mockup LinkedIn pages (see Figure 2). There were two profiles for each of the following groups: attractive white $M_{attract.} = 7.00$, $M_{phen.} = 4.00$, $M_{afro.} = 2.50$, $M_{age.} = 25$, $M_{smile.} = 8.13$, $M_{comp.} = 7.07$, attractive light-skinned black $M_{attract.} = 6.78$, $M_{phen.} = 6.07$, $M_{afro.} = 6.95$, $M_{age.} = 30.5$, $M_{smile.} = 7.38$, $M_{comp.} = 7.94$, attractive dark-skinned black $M_{attract.} = 6.83$, $M_{phen.} = 7.15$, $M_{afro.} = 7.82$, $M_{age.} = 27$, $M_{smile.} = 9.53$, $M_{comp.} = 8.19$, unattractive white $M_{attract.} = 3.26$, $M_{phen.} = 4.57$, $M_{afro.} = 2.88$, $M_{age.} = 34$, $M_{smile.} = 7.76$, $M_{comp.} = 7.44$, unattractive light- skinned black $M_{attract.} = 3.20$, $M_{phen.} = 5.57$, $M_{afro.} = 7.32$, $M_{age.} = 28$, $M_{smile.} = 7.85$, $M_{comp.} = 7.00$, and unattractive dark- skinned black $M_{attract.} = 3.26$, $M_{phen.} = 8.19$, $M_{afro.} = 9.01$, $M_{age.} = 28.5$, $M_{smile.} = 8.32$, $M_{comp.} = 7.19$. Each profile contained a picture of a woman, her major, and a reference from an employer. Her contact information, University she attended, and name a position of the reference were all crossed out. The instructions informed participants that the information was crossed out to ensure the candidates’ privacy. To ensure competency was being directly measured, each profile

consisted of four questions, which were answered using a 7-point scale. For the questions, “How likely would you be to hire the woman above?,” “How competent do you think she is?,” “How skilled do you think she would be in a work setting?,” and “Would you consider her knowledgeable?” the scale ranged from 1=*Not at all* to 7=*Very*. Additionally gender, age, ethnicity, and level of education were requested. Participants were also welcome to share any thoughts they had about the study and asked to rate what most influenced their judgments.

Procedure

I used an online questionnaire through the Tufts Qualtrics website. I manipulated the mock-up LinkedIn profiles to see the effects on employer attitudes. I included the same pictures and profiles in each condition to increase internal validity. In each condition, each photo was assigned a new LinkedIn page and participants were randomly assigned to one of the three conditions to ensure internal validity. The survey consisted of 12 pages each containing one LinkedIn profile and four questions, a page for inquiries and comments, and another page asking demographic questions.

Experiment 1

Experiment 1 explored only the impact of physical appearance on viewers’ judgments. By only having neutral references, participants are forced to assess candidates based off of their pictures. Experiment 1 shows the basic evaluations made when people judge others without any diagnostic context.

Method

Design

The experiment was a 3x2 repeated measures design in the form of an online survey. There were three levels of racial phenotype: white, light-skinned black, and dark-skinned black.

There were two levels of attractiveness: attractive and unattractive. Attractiveness and racial phenotype served as within-study independent variable. I mixed and matched the faces with each of the profiles so each condition had different pairings. For example, in Condition 1, one of the attractive white women was placed on a different profile than in Conditions 2 and 3. Such pairings and distributions were made to increase validity.

Participants

Participants were from a convenience sample that came from various Tufts Psychology classes, which they registered through Tufts Sona. They received credit toward their class in exchange for their time. The participants consisted of 40 females and 27 males ($M_{age} = 19.70$, $SD = 1.52$, $Min = 17$, $Max = 26$), 41 of the participants were White, 8 Asian, 8 Multiracial, 5 Black, 4 Hispanic, one Eastern Indian, and one Native American. Of the 67 participants, 22 of them were assigned to Condition 1, 23 to Condition 2, and 23 to Condition 3. Out of the 22 people assigned to Condition 1, 8 of them were males and 14 females. From the 23 participants in Condition 2, 11 of them were male and 12 of them were female. Of the 22 assigned to Condition 3, 8 of them were males and 14 females.

Procedure

I recruited all of my participants through Tufts Sona and granted them credit toward their class. After signing printed consent forms, participants began the survey, which informed them that they could exit the site at any time if they wished to do so. After completion, the survey showed a debrief form, which explained the nature of my study. The survey was in accordance with all of the Institutional Review Board standards.

Results

I predicted an interaction between attractiveness and racial phenotype on perceived competence. For my experiment, I conducted a repeated measures ANOVA with an alpha level of .05 and a confidence interval of 95%. The four questions asked to assess candidates' ability were reliable (Cronbach's $\alpha = .94$). There was a significant main effect for phenotype, $F(2,66) = 2.80, p < .05$. Dark-skinned black women ($M = 4.22, SE = 0.10, 95\% CI = [4.01, 4.42]$) had higher evaluations than the white ($M = 4.08, SE = 0.10, 95\% CI = [3.87, 4.29]$) and light-skinned black women ($M = 4.10, SE = 0.10, 95\% CI = [3.87, 4.33]$). However, there was no significant main effect for attractiveness, $F(1,66) = 2.11, p > 0.05$. Participants did not rate attractive candidates ($M = 4.17, SE = 0.099, 95\% CI = [3.96, 4.34]$) and unattractive candidates ($M = 4.09, SE = 0.10, 95\% CI = [3.88, 4.29]$) statistically differently. There was a significant interaction between attractiveness and phenotype, $F(2,66) = 8.49, p = .01$ (see Figure 3). Attractiveness was beneficial for the black women whereas the white women benefited more when unattractive. Participants' gender or race did not have a significant interaction. The data supports my hypothesis that phenotype and attractiveness have a significant effect on viewers' attitudes. However, the interaction was not in the predicted direction in that attractiveness and a lighter skin tone did not consistently yield higher ratings.

Discussion

There was a significant interaction between phenotype and attractiveness in Study 1. The data do not support my hypothesis that a lighter phenotype and higher level of attractiveness positively affect employer attitudes. In fact, the results show that a darker phenotype is actually beneficial and that the effects of attractiveness depend on the women's racial phenotype. Participants rated the black women more highly when attractive whereas they gave white women higher ratings when unattractive.

Such an opposing outcome could be attributed to the transparency of Experiment 1. All of the participants from Experiment 1 were Tufts students who were taking a Psychology course. It is entirely possible that they were skeptical about the nature of the study since most students are aware that a majority of Tufts studies are assessing discrimination, particularly looking at race. Furthermore, since the references were neutral, it may have been too obvious that the statements were fabricated. If the participants had such knowledge, they could have been overcompensating by giving the darker-skinned black women higher ratings.

Experiment 2

Experiment 2 was designed to extend the results of Experiment 1. First, instead of only exploring the effects of attractiveness and racial phenotype, Experiment 2 adds a third variable, level of qualification. Second, Experiment 2 was open to a wider range participant sample to generalize findings. Such additions make the findings more applicable to a real world setting.

Method

Design

The study was a 3x2x2 repeated measures design in the form of an online survey. There were three levels of racial phenotype: white, light-skinned black, and dark-skinned black. There were two levels of qualifications: high and low. There were two levels of attractiveness: attractive and unattractive. I mixed and matched the faces with each of the profiles so each condition had different pairings. For example, in Condition 1, one of the attractive white women was placed on a positive profile whereas the second attractive white woman was placed on a negative profile. In Condition 2, the two women switched profiles. Such pairings and distributions were made to increase validity by ensuring each classification of women would be placed with most of the profiles.

Participants

Participants were recruited via email and Social Media from a convenience sample and consisted of 130 females, 53 males, and 3 who did not identify ($M_{age} = 37.34$, $SD = 18.30$, $Min = 18$, $Max = 88$). There was a wide range in participants' educational level: 2 of them were in High School, 24 attended some college, 4 attended a 2-year college, 42 attended a 4-year college, 25 were at a Masters level, 3 had pursued a Doctoral degree, and 6 pursued a Professional Degree. However, there was not much variance in terms of ethnicity: 162 of the participants were White, 3 were Black, 20 were Asian, 4 Hispanic, and one Multiracial. Of the 186 participants, 49 of them were assigned to Condition 1, 52 to Condition 2, 33 to Condition 3, and 52 to Condition 4. Out of the 49 people assigned to Condition 1, 17 of them were males and 30 females. From the 52 participants in Condition 2, 16 of them were male and 36 of them were female. Of the 33 assigned to Condition 3, 12 of them were males and 20 females. And, out of the 51 people assigned to Condition 4, 8 of them were males and 43 females.

Materials

The same materials were used as in Experiment 1 except with the manipulation of the favorability of the reference. Unlike Experiment 1 which included neutral references, Experiment 2 included two levels of references: positive and negative. Other than the references, the same faces, majors, layout, and questions were used.

Procedure

I used an online questionnaire through the Tufts Qualtrics website. I manipulated the mock-up LinkedIn profiles to see the effects on potential employer attitudes. I included the same pictures and profiles in each condition to increase internal validity. In each condition, each photo was assigned a new LinkedIn page, so it would be paired with a negative profile half the time

and a positive profile the other half. To ensure random assignment the survey then automatically assigned participants to one of four conditions. The survey consisted of 12 pages each containing one LinkedIn profile and four questions, a page for inquiries and comments, and another page asking demographic questions.

I recruited 14 participants from the Mechanical Turk website (www.MechanicalTurk.com), and I compensated them 25 cents. Mechanical Turk is a crowdsourcing Internet Marketplace owned by Amazon, which enables “workers” to complete tasks for “requesters” for a small fee. The rest of the participants were recruited via email and social networking websites, and they were not offered any compensation for taking the study. Once participants opened the link to the survey, the site presented participants with a consent form informed participants that they could exit the site at any time if they wished to do so. After completion, the survey showed a debriefing form, which explained the nature of my study. The survey was in accordance with all of the Intuitional Review Board standards.

Results

I predicted a three way interaction of attractiveness, qualifications, and phenotype on perceived competence. For my experiment, I conducted a Repeated Measures ANOVA with an alpha level of .05 and a confidence interval of 95%. The four questions were statistically comparable to each other (Cronbach’s $\alpha = .97$). The levels of qualification were statistically different, $F(1,187) = 184.25, p < 0.01$. The profiles with higher qualifications ($M = 4.26, SE = 0.10, 95\% CI = [4.05, 4.47]$) had higher ratings than those with lower qualifications ($M = 3.26, SE = 0.10, 95\% CI = [3.07, 3.45]$). Candidates with positive profiles were perceived as more competent than those with negative profiles. Phenotype also had a significant main effect, $F(2,187) = 34.72, p < 0.01$. White women ($M = 3.54, SE = 0.10, 95\% CI = [3.34, 3.74]$) were

not perceived as competent as lighter-skinned black women ($M = 4.05$, $SE = 0.10$, 95 % $CI = [3.85, 4.24]$) and dark-skinned black women ($M = 3.70$, $SE = 0.10$, 95 % $CI = [3.5, 3.89]$).

Lighter-skinned black women were regarded the most highly and then darker-skinned black and white women respectively. However, there was no significant effect for attractiveness, $F(1,187) = 1.20$, $p > 0.05$. Attractive candidates ($M = 3.7$, $SE = 0.099$, 95 % $CI = [3.54, 3.93]$) and unattractive candidates ($M = 3.79$, $SE = 0.10$, 95 % $CI = [3.0, 3.98]$) were rated similarly. On its own, attractiveness did not play an important role in the decision making process. Overall, there was a significant three-way interaction of attractiveness, phenotype, and qualifications, $F(2,187) = 20.72$, $p < 0.01$ (see Figure 4). The data support my hypothesis in that phenotype, attractiveness, and qualifications have a significant interaction on viewers' attitudes. However, it was not in accordance with my predictions that lighter skin and attractiveness would positively affect perceived ability.

Discussion

The combined effects of racial phenotype, attractiveness, and qualifications had a significant interaction. When assessing the candidates, participants gave most weight to qualifications and then considered the attractiveness and racial phenotype of the person. Attractiveness was beneficial for Whites and negative for dark-skinned Blacks overall. As a function of valence, light-skinned Blacks rated much higher than the darker-skinned Blacks and Whites. When given a negative reference, the darker-skinned Blacks and Whites scored very similarly. However, when given a positive reference, darker-skinned Blacks rated a bit higher than the Whites, but still significantly less than the light-skinned Blacks. Overall, attractiveness helped for white women and was disadvantageous for the black women. However, attractiveness did not matter for the light-skinned black women with positive references or for the white women with negative

references. There was no significant interaction with participants' gender, though on average, males gave the women higher ratings. There also was no significant interaction with participants' race.

General Discussion

Primary Hypothesis

Given the previous research, it was predicted that there would be a positive correlation between lighter skin tone and attractiveness and perceived competence. However, while there were significant interactions, they were not in the predicted direction. In both Experiments 1 & 2, black women received higher ratings than white women. Additionally, the findings of both experiments show attractiveness is not consistently helpful.

The findings for Experiment 1 were not in accordance with my literature review on phenotype. Previous research has consistently found that lighter skin color is positively correlated with competency (Hill, 2000; Hunter (2002); Keith & Herring (1991). However, in Experiment 1, darker-skinned black women were rated the highest among the three groups. While it is difficult to attribute a cause for such differing data, it might be reasonable to assume the goal of the study was transparent. Since the statements were all neutral, participants may have developed suspicion. Furthermore most of the subjects had participated in other Tufts Psychology studies before and were aware that many labs study race. Such factors might have contributed to such discrepancies in results.

On the other hand, the findings for Experiment 2 are in accordance with those studies that previously investigated the effects of phenotype on perceived competence. Previous cross-sectional studies all found that light-skinned Blacks were more successful than those with darker-skin (Hill, 2000; Hunter, 2002; Keith & Herring, 1991). Additionally, experiments that

explored the relationship between perception and race found that judgments were more lenient for light-skinned blacks (Harrison & Thomas, 2009). Such studies concluded that lighter-skinned Blacks were consistently given better ratings than darker-skinned Blacks. As expected, Experiment 2 found that, on average, light-skinned black women were rated more competent than dark-skinned black women. However, Experiment 2 also found that light-skinned black individuals had higher ratings than white individuals. While the cross-sectional studies only compared the phenotypes of Blacks, disregarding Whites, it is surprising that a medium skin tone scored the highest. Overall, Experiment 2 concluded light-skinned Blacks were rated the most highly within their qualification condition regardless of attractiveness.

Additionally, the findings for both Experiments 1&2 conflict previous research regarding the “what’s beautiful is good” phenomenon. For Experiment 1, the results are in accordance with such a belief when isolating the results of the black women. However, attractiveness is actually disadvantageous for white women as Experiment 1 demonstrates. On the other hand, the findings for Experiment 2 show that attractiveness is helpful only for white women. With inconsistent effects of attractiveness, the findings contradict the studies discussed earlier which found that attractive women were overall considered better (Ahearne Gruenb, & Jarvis, 1999; Benson, Karabenick, & Lerner, 1979; Dion, Berscheid, & Walster, 1972). However, one such study found that attractiveness helps, but its effects are marginally small (Dubois & Pansu, 2004). An explanation for the different results might be that the small effect found in the previous study was overridden by the qualification and phenotype variables.

Furthermore, the results for Experiments 1&2 conflict with previous research that explores the interaction of attractiveness and phenotype. Unlike other research, which found that people typically rated the photos of attractive, white students higher than their unattractive, black

counterparts, neither of my studies resulted in attractive, white women with the highest ratings (Parks and Kennedy, 2007). Additionally, my experiment contradicts the DeMeis (1978) study, which compared race, attractiveness and dialect to measure teachers' responses. The DeMeis study found that teachers gave most weight to race and attractiveness while evaluating their students whereas the results of Experiment 2 show that evaluators give most attention to qualification. One possible explanation for the differing results from the DeMeis study is that I did not include a dialect variable. Furthermore, the DeMeis and Parks study explored the judgments of children whereas I only looked into the effects on women.

Finally, it is also possible that other aspects of facial appearance besides physical attractiveness may have played a role. For example, prior research has found that inferences of power taken from photos of top lawyers significantly correlated with their level of success (Rule & Ambady, 2008). Furthermore, a study which compiled photos of West Point students found that facial dominance was strongly related with promotions and rankings (Mazur & Mueller, 1996). It is possible that a "competency" feature influenced the outcome of the experiment when considering the results of the Pretest data. The Pretest asked participants very straightforwardly to rate the women's perceived competence. It is hard to compare the numbers since the experiments used a 7-point scale and the Pretest used a 10-point scale, but the overall trend of the Pretest is consistent with Experiment 1's results (see Figure 5).

Limitations and Future Directions

The design used minimized threats to internal validity. The present experiment used stimuli that were pretested to assure consistency of attractiveness, phenotype, and overall appearance and quality of the photo. Two women were used for each group to ensure a representative sample of the classification they were signifying. For example, the survey

included two attractive, white women, two unattractive white women, two attractive light-skinned women, etc. Furthermore, in each condition in Experiment 2, one woman from each of the groups was assigned to a negative profile, and the other was assigned to a positive profile. In Study 1, each woman was assigned a different profile each condition. Additionally, the present study increased the internal validity by having ensuring random assignment. The program automatically decided which the order in which the profiles appeared and randomly assigned participants to one of conditions. I eliminated potential biases and environmental confounds by having randomization and consistency among stimuli groups.

However, my stimuli across groups were not all rated statistically the same. The women in the unattractive dark-skinned black group were significantly darker than the women in the attractive dark-skinned black group. However, I do not believe this to be an issue because the Pretest ratings for each classification show that participants could still tell these factors apart. Additionally, the women in the attractive light-skinned black category were classified as looking much younger than the rest of the women. Finally, the attractive, dark-skinned black women were rated as smiling much more than the rest of the women. A lack of generalization across groups may have affected my results. However, it is unlikely that such discrepancies served as confounds because, within the Pretest, participants rated the faces' competency and there was no correlation between smiling, age, and competency.

Additionally, in Experiment 2 two of the Conditions in the Qualtrics survey were not formatted correctly. In condition three, participants saw two LinkedIn profiles of attractive, dark-skinned black women with positive references, but they did not see any profiles of attractive, dark-skinned black women with negative references. Also, in the fourth condition, participants saw two LinkedIn profiles of attractive white and light-skinned black women with positive

references. They did not see profiles of attractive light-skinned and dark-skinned black women with negative references. For the instances where participants saw two profiles from the same category, I averaged their two answers together and inputted it as one response. Overall, these variances among Conditions and profiles could have affected my data since certain groups could be given more weight statistically. Furthermore, the majors assigned to the LinkedIn profiles were not pretested to ensure they were generally the same. While these differences could have affected my data, they most likely did not because most participants said at the end of the survey that they did not take major into account when rating the candidates.

Despite the present experiments' many findings, there are several ways in which future research can expand on these results. For future replication, the stimuli could be reproduced to be standardized across levels, in which levels of attractiveness and skin tone are statistically equivalent throughout conditions. This can be accomplished by performing a larger pretest with more than 40 faces or by making the stimuli digitally. Such tactics would yield stronger stimuli with less variance among levels. Additionally, the study should be repeated with uniform Conditions, in which each category of women is paired with both a positive and negative profile. While the gaps in my Conditions most likely did not play a huge role, it is important that participants see one LinkedIn profile from each of the 12 categories in order to ensure validity. Furthermore, it would be interesting to see the study replicated as a between subjects design instead of the performed within subjects design. Such a design would lessen the likelihood of transparency.

Conclusion

My findings for Experiment 1 show a main effect for attractiveness and racial phenotype. The combination of attractiveness and a darker skin tone yielded the highest ratings. White

women received the highest scores when they were unattractive. For the second experiment, the results show that attractiveness, phenotype, and qualifications in combination have significant effects on people's judgment during the interviewing process. Qualifications had the most weight on a person's decision, but attractiveness and phenotype affected them as well. Overall, attractiveness helps white candidates whereas it can be disadvantageous for dark-skinned Blacks and somewhat irrelevant for light-skinned black women. Unattractive, dark-skinned women and attractive white women were rated similarly, but positive qualifications tended to weigh more for the unattractive, dark-skinned black candidates. My results strongly indicate that the three variables interact to significantly affect employer responses. My findings can serve as a foundation for future research on improving the interviewing selection process. This information is highly valuable for those involved in Human Resources or the *Psychology and Marketing* journal. While no one can completely monitor subconscious thoughts or beliefs, acknowledging such prejudices is a good start.

References

- Agthe, M., Sporrle, M., & Maner, J. (2011). Does Being Attractive Always Help? Positive and Negative Effects of Attractiveness on Social Decision Making. *Personality Social Psychology Bulletin*, 37(8), 1042-1054. doi:10.1177/0146167211410355
- Ahearne, M., Gruenb, T., & Jarvis, B.(1999). If Looks Could Sell: Moderation and Mediation of the Attractiveness Effect on Salesperson Performance. *International Journal of Research in Marketing*, 16, 269-284. doi: 10.1016/s0167-8116(99)00014-2
- Anderson, R. & Nida, S. (2006). Effect of Physical Attractiveness on Opposite- and Same-Sex Evaluations. *Journal of Personality*, 46(3), 401-413. doi: 10.1111/j.1467-6494.1978.tb01008
- Benson, P., Karabenick, S., Lerner, R. (1976). Pretty Pleases: The Effects of Physical Attractiveness, Race, and Sex on Receiving Help. *Journal of Experimental Psychology*, 12, 409-415. doi: 10.1016/0022-1031(76)90073-1
- Caballero, M.J. & Pride, W.M. (1984). Selected Effects of Salesperson Sex and Attractiveness in Direct Mail Advertisements. *Journal of Marketing*, 48(1), 94-100. doi: 10.2307/1251315
- Chaiken, S. (1979). Communicator Physical Attractiveness and Persuasion. *Journal of Personality and Social Psychology*, 37(8), 1387-1397. doi: 10.1037/0022-3514.37.8.1387
- DeMeis, D. & Turner, R. (1978). Effects of Students' Race, Physical Attractiveness, and Dialect on Teachers' Evaluations. *Contemporary Educational Psychology*, 3(11), 77-86. doi:10.1016/0361-476X(78)90012-7
- DeShields, O.W. Jr., Kara, A., & Kaynak, E. (1996). Source Effects in Purchase Decisions: The Impact of Physical Attractiveness and Accent of Salesperson. *International Journal of Research in Marketing*, 13(1), 89-101. doi: 10.1016/0167-8116(95)00036-4

- Dion, K., Berscheid, E., & Walster, E. (1972). What is Beautiful is Good. *Journal of Personality Social Psychology*, 24(3), 285-90. doi: 10.1037/h0033731
- Dubois, M. and Pansu, P. (2004). Facial Attractiveness, Applicants' Qualifications, and Judges' Expertise About Decisions in Preselective Recruitment. *Psychological Reports*, 95, 1129-1134. doi: 10.2466/pr0.95.3f.1129-1134
- Harrison, M. & Thomas, K. (2009). The Hidden Prejudice in Selection: A Research Investigation on Skin Color Bias. *Journal of Applied Social Psychology*, 39(1), 134-168.
doi:10.1111/j.1559-1816.2008.00433.x
- Hill, E. (2000). Color Differences in the Socioeconomic Status of African American Men: Results of a Cross-sectional Study. *Social Forces*, 78(4), 1437-1460. doi: 10.1093/sf/78.4.1437
- Hunter, M. (2002). "If You're Light You're Alright" Light Skin Color as Social Capital for Women of Color. *Gender and Society*, 16(2), 175-193.
doi: 10.1177/08912430222104895
- Kahle, L.R., & Homer, P.M. (1985). Physical Attractiveness of the Celebrity Endorser: A Social Adaption Perspective. *Journal of Consumer Research*, 11(4), 954-961.
doi: 10.1086/209029
- Keh, H., Ren, R., Hill, S., & Li, X. (2013). The Beautiful, the Cheerful, and the Helpful: The Effects of Service Employee Attributes on Customer Satisfaction. *Psychology and Marketing*, (30)3, 211-226. doi: 10.1002/mar.20599
- Keith, V. & Herring, C. (1991). Skin Tone and Stratification in the Black Community. *American Journal of Sociology*, 97(3), 760-778. doi: 10.1086/229819

- Mazur, Allan & Mueller, Ulrich (1996). Facial Dominance of West Point Cadets as a Predictor of Later Military Rank. *Research in Biopolitics*, 74 (3): 823-850.
doi: 10.1093/sf/74.3.823
- Parks, F. & Kennedy, J. (2007). The Impact of Race, Physical Attractiveness, and Gender on Education Majors' and Teachers' Perceptions of Student Competence. *Journal of Black Studies*, 37(6), 936-943. doi:10.1177/0021934705285955
- Ritts, V., Patterson, M. & Tubbs, M. (1992). Expectations, Impressions, and Judgments of Physically Attractive Students: A Review. *Review of Educational Research*, 62(4), 413-426. doi: 10.3102/00346543062004413
- Senior, C., Thomson, K., Badger, J., & Butler MJ. (2007). Interviewing Strategies in the Face of Beauty: a Psychophysiological Investigation into the Job Negotiation Process. *Annals of the New York Academy of Sciences*, 1118, 142-162. doi: 10.1196/annals.1412.005
- Rule, N. & Nambady, N. (2010). Judgments of Power From College Yearbook Photos and Later Career Success. *Social Psychological and Personality Science*, 2, 154.
doi:10.1177/1948550610385473

Stimuli



Figure 1. The survey conducted had six different categories of women with three levels of phenotype and two levels of attractiveness: attractive white, attractive light-skinned black, attractive dark-skinned black, unattractive white, unattractive light-skinned black, and unattractive dark-skinned black women (Top left to bottom right, respectively). Each survey included a profile from each of the 12 women.

Mockup LinkedIn profile



Figure 2. The survey conducted had 12 different mockup LinkedIn profiles. Each survey included a profile from each of the 12 categories. Featured is a profile used for the unattractive, white candidate with a negative reference. Each LinkedIn profile included a photo, BA or BS degree with a vague major, a positive, negative, or neutral reference, and crossed out information, which participants were told was to ensure privacy of the candidate.

Attractiveness and phenotype interaction (Experiment 1)

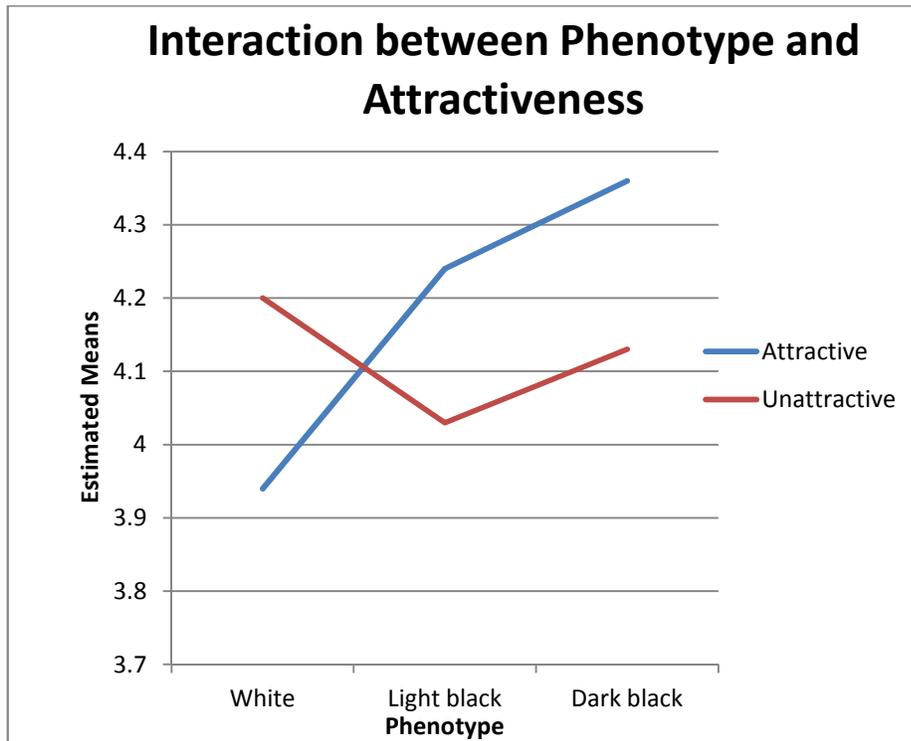


Figure 3. The data shows that attractiveness is more beneficial for darker and lighter-skinned black women. However, white women were rated higher when they were unattractive.

3-way interaction in terms of qualification level (Experiment 2)

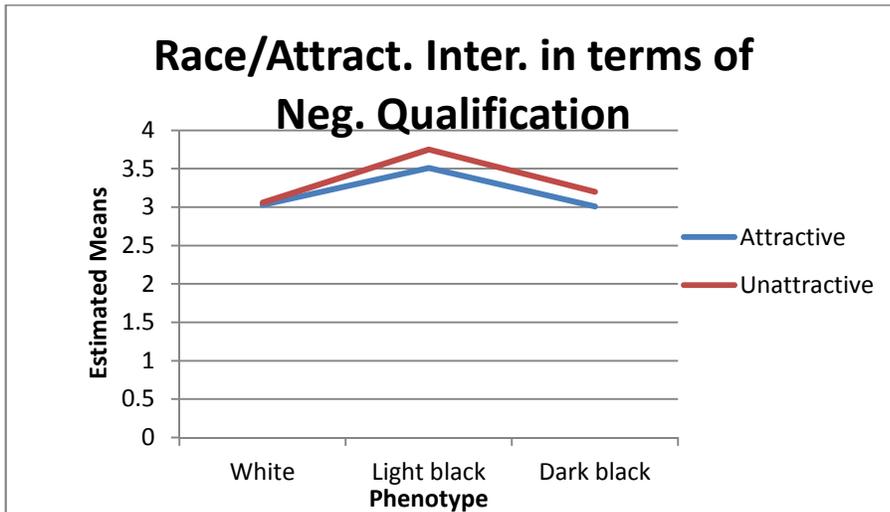
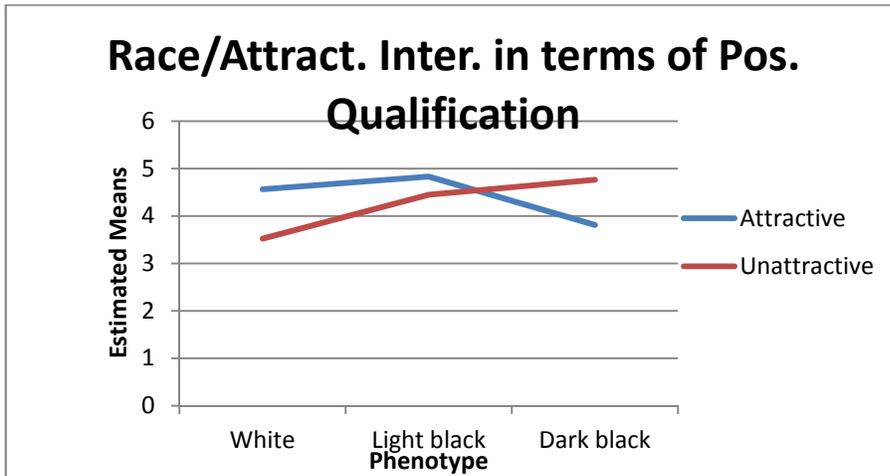


Figure 4. The data shows that attractiveness helped white women with positive profiles. On the other hand, dark-skinned black women benefited when they were unattractive with a higher profile. Attractiveness did not play a role for light-skinned black women with positive references.

When having a negative profile, attractiveness did not play a significant role for white women. Attractiveness helped for both light and dark-skinned black women with lower profiles. Participants perceived light-skinned black women the highest regardless of attractiveness when they were pair with a negative reference.

Comparison of Experiment 1 Interaction and Pretest Ratings

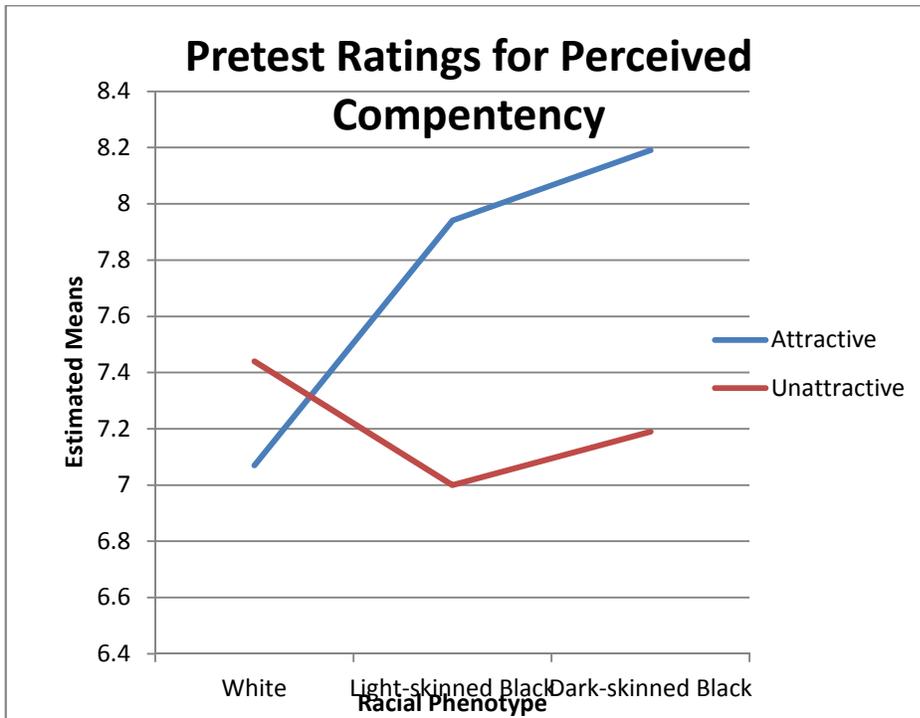
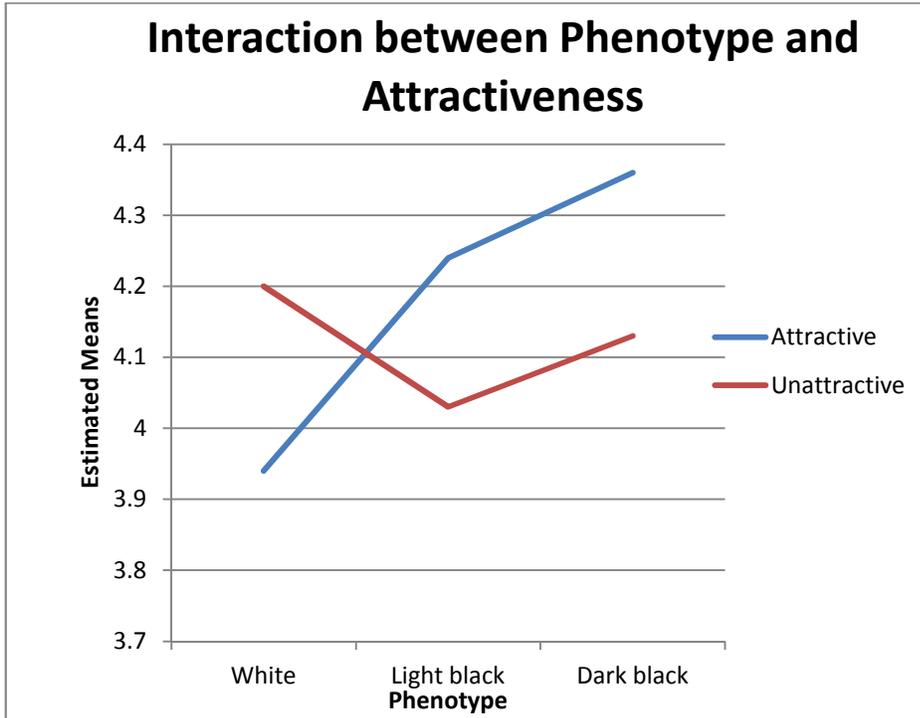


Figure 5. The data shows the similarities between the results of Experiment 1 and the Pretest ratings for perceived competency.