Analyzing the Changing Face of America: A Political Psychology Experiment to Measure the Perception of Skin Tone in Political Candidates

Presented to the
Department of Government
in partial fulfillment of the requirements
for the degree with honors
of Bachelor of Arts

Harvard College
March 2014
# Table of Contents

I. Introduction .................................................................................................................. 3

   Abstract
   Initial Thoughts

II. Literature Review ......................................................................................................... 12

   What is Colorism?
   History of Race and Skin Tone in the US
   Skin Tone in Culture: Impact on Majority and Minority Populations
   Perspectives on Skin Tone in Modern American Culture
   Contextualization of Skin Tone in Politics
   Application of Theories

III. Methods ...................................................................................................................... 23

   Methods Overview
   Research Question and Hypotheses
   Experimental Design
   Summary of Methods

IV. Results/Discussion ....................................................................................................... 46

   Data
   Discussion

V. Conclusion .................................................................................................................... 64

   Future Implications
   Final Thoughts

VI. Appendix ..................................................................................................................... 68

   Final Survey Tool
   Bibliography
Part I. Introduction
Abstract

The changing face of America could influence the visual makeup of our governing body. While political theorists have done much research surrounding race in politics, little attention has been paid to the more nuanced role of skin tone. Building upon research that has focused on voters’ snap judgments of politicians and the importance of skin tone in the assessment of Black politicians, I ran a political psychology experiment to test these criteria on photographs of United States Congressmen. Ultimately, I found that skin tone drives the degree of positivity in the opinions voters hold towards Black, White and Hispanic Congressmen. This finding holds for all character trait judgments that were tested, as well as across respondents of all races.

Initial Thoughts

Since its inception, the United States has served as a “melting pot,” eventually yielding a great amount of diversity amongst citizens. This diversity, as a result, has a profound effect on every facet of society as it relates to race. Therefore, in this thesis, I examine the intersection of race and politics in the United States, specifically focusing on the role that skin tone plays in voters' assessments of political candidates using Congressmen for the subjects in this study. While much attention has been paid to politicians’ races, as seen recently with the election and

---

1 As a note: Throughout this thesis, it should be assumed that White and Caucasian, Black and African American and Hispanic and Latino are used interchangeably to refer to the same populations.
re-election of Barack Obama, little emphasis has been placed specifically on skin tone. As American history has presented multiple examples of preference for lighter skin tones, one might reasonably assume that a study will show the same biases in the political sphere today. Not only has this been seen in the United States in the White and African American populations, but this thesis will broaden the communities of interest to show the same sentiments in the Latino community.

**Research Question**

The premise of this thesis is the suggestion that colorism impacts judgments of politicians, which then affects the electability of the candidates based on voter preferences and resulting decisions. In contrast to a general discussion of race in politics, this research focuses on skin tone in particular. In the study, politicians of African American, Hispanic and Caucasian ethnicities are analyzed by thousands of participants of all races (as identified by categories in the U.S. Census) to examine patterns that exist in many different populations. Given the varied histories and relationships with skin tone in the African American, Hispanic and Caucasian communities, the crossover of socialized reasoning processes into the political arena will likely uncover more intricate biases related to skin color itself. Racial discrepancies are evident in the make-up of the House of Representatives, as the chamber does not accurately reflect proportions of different racial groups in the broader population, yet there is not a discussion of skin tones within the body and how that might have affected
electability. This discrepancy points to the banality of racism in common discourse versus the much more subtle and less acknowledged factor of skin tone.

If it is skin color distinctions that are driving the extent to which race matters, this finding becomes relevant, as particular phenotypes have led to stereotypes for each race. As lighter or darker skin color is distinguished to be more or less typical of someone of a certain race, it is possible for opinions to shift if a participant perceives their judgment to be of a White versus Hispanic politician, for instance. Thus, in order to test the photos used in the study, a pilot test was administered to ensure that the perceived race of the politician was consistent as the skin color changed. This finding would have widespread application because of its ability to explain the mental processes that govern instant opinion formation since the pictures are being rapidly flashed. This project is also pivotal as the United States continues on a trajectory to be increasingly diverse due to the heightened number of multiracial relationships that are creating a greater variance of skin tones. Since there is no longer a strict binary of Black and White but rather a broad range of colors in between, it is fascinating that there are still very clear instances of hierarchical ordering based on skin color in society. Noting that it is broader societal views that drive politics, there is a logical progression toward the assumption of skin color preference in elections.

**Main Argument**

My hypothesis is that this seemingly minute detail may have grander implications than have been noted in prior research. I predict that the lighter
versions of Congressional candidates will be preferred across all races of survey participants, as this preference has been seen in society. Secondarily, I predict that participants rating candidates of their own race will be more negative in their initial reactions as compared to candidates of other races because of the harsher implicit judgments that the in-group feels comfortable making. This matters because politics boils down to implicit and explicit judgments. In today’s age, in which the number of people who now identify themselves as multiracial creates a much wider spectrum of colors, there will be potential shifts in the appearance of those elected to Congress, which could have broader societal impact.

The project itself begins with the selection of pictures of male Congressional representatives and the altering skin tones of the politicians in the photos to both lighter and darker shades. Each picture is altered to essentially the same degree, with the result being faces that are recognizably lighter and darker than the original. A pilot test is then run amongst Harvard students and Amazon’s Mechanical Turk users in order to get an initial read on how the sample views particular characteristics of the pictures themselves in order to make certain that the pictures are otherwise comparable and are not cuing something other than skin tone. Moving to the true experiment, using these photos and randomly assigning different versions to be used within a single survey (i.e. a lighter version of one representative, a darker version of another and the actual coloring of a third, etc.), I then flash these pictures briefly and get immediate responses on perceived attributes of the politicians in the photos. By varying only the skin tone in the official congressional photographs of six Representatives, I am able to isolate skin
tone as the variable of interest. The survey ran on Mechanical Turk, a website with hundreds of thousands of subscribers who were used to mimic a voting population.

Not only does this provide a snapshot of inherent biases that are held within society, but it also demonstrates this expression in political representation. As these politicians are Congressional decision-makers and a sample of voting-age Americans was taken, there is much that can be extrapolated from the findings to help inform actual elections. In gathering the survey information, I also collect demographic data to examine the age of those taking the survey. This is used to shed light on the future of representation in the United States and the value that colorism may hold in future elections, since younger voters may likely choose to participate in elections for decades to come.

In examining how people view skin tone in determining their opinions of politicians, contextual truths about the future of politics are revealed. It is easy for one to acknowledge the shift in the “face of America” since the racial dichotomy that used to represent the U.S. has since transformed into a much larger scale. This research will hopefully add to a guided debate about race in politics and the true role that skin tones play in perceptions of others.

**Preview of Results**

At a high level, the results supported the hypotheses about skin tone being a factor in the analysis of politicians and that the trend was one where the lightened photographs were preferred in nearly every instance. These trends held
across all of the politicians and characteristics, meaning that the research was able to expand on prior studies to show this correlation regardless of the race of the voter or the politician. Unlike my original prediction, survey participants rated photos with politicians of the same race more favorably, meaning that there is an argument for voting along racial lines as well. Colorism has become embedded in American society and has carried over into the political realm, which signals that the history of skin tone in the United States continues to have a broad impact; not only has it affected sociological factors (assessments of beauty, career success, etc.), but it has also influenced psychological processes in opinion formation and electability. These trends are also likely to extend into the future, as tests using the responses of those under 40 showed that they maintain a bias towards politicians with lighter skin.

The strategy of election campaigns could be affected, as the results of this political psychology experiment can be extrapolated to actual elections. Since there is a link between skin tone and initial judgments, there may be methods to best position candidates through their publicity materials. This could be seen as areas with greater populations of White vs. Asian vs. African American, etc. voters may see campaigns more or less focused on showing the image of the politicians on materials based on how assumptions will be formed about the candidate. For instance, as the study shows that Asian Americans respond to African American candidates with a consistent preference of lighter skin, a lighter-skinned Black candidate may want to have their image more visible in an area with largely Asian demographics whereas a darker-skinned Black candidate
may have a campaign more focused on non-visual campaign elements as opposed to showing their image to the same degree. While, in theory, elections should be determined by the platforms of the candidates, much research has suggested that appearance is also a significant factor in determining who will win an election. Thus, from a strategic perspective, attention must be paid to the positioning of the image of the candidate.

**Structure of the Thesis**

The thesis has begun by giving a general overview and the specific question I am looking to address through my research as well as a first glimpse of what is found and the resulting inferences that can be drawn. In Chapter 2, the Literature Review, I draw upon history, sociology, psychology and political science to better understand the role that skin tone currently holds in society. After examining how other studies have found that skin tone affects certain life outcomes, I postulate how these realities may play out in the world of politics and how this study extends prior findings. In Chapter 3, Methods, I describe the experimental design, including a review of online survey methods and the results of a pilot test performed before running the final survey. I also detail the hypotheses and variables of the experiment, as well as the recruitment process for subjects. Chapter 4, Data/Results, gives the summary statistics from the final experiment in addition to reviewing the findings from the calculations in R. Finally, Chapter 5, Conclusion, draws the final inferences from the data gathered in the experiment and discusses possible future implications of the findings.
Part II. Literature Review
What is Colorism?

It is undeniable that race has played a large role in shaping American society. From the time foreigners came to what is now U.S. soil, there has been a distinct separation and hierarchy formed based on appearance. With the original division of light-skinned Europeans and dark-skinned African slaves and American Indians, the power was placed within the camp of those with white skin very early on. Nuanced within race, however, there is a story that has been minimally examined and not had much of an earnest discourse. Especially as mixing occurred within the aforementioned groups and between other immigrants, skin tone has become one attribute of many that has driven stratification within races. As it relates to this project, to study colorism is to look at skin color specifically, apart from a broader discussion of race, to see the effects that the tone itself has on the perception of a person.

In boiling down the discussion from race to skin tone, there are many factors to draw upon to prove that colorism matters. In Skin Deep: How Race and Complexion Matter in the “Color-Blind” Era, a number of authors looked at the differences in treatment of individuals within the same racial group but with a variety of skin tones. One author’s research showed that “African Americans with lighter skin complexions were more likely to have higher social and economic statuses than their darker skinned counterparts” (Herring et al. 1). This fact suggests that there is a greater demand to study the intraracial facets of this discussion, as there are many determinants of lifestyle; however, the key to this
argument lies in the increased opportunities that those with lighter skin receive in life. As Dr. Cedric Herring described one view of African American culture in his book, “lighter skinned Blacks were extended more social and economic amenities than the larger group…[which] translated over successive generations into the Black elite being of fairer complexion” (Herring et al 9). In the past, opinion polls and surveys have been widely used to facilitate data collection regarding attitudes towards skin tone within politics. A 1991 study conducted by Richard Seltzer and Robert Smith that analyzed colorism within the African American community found positive effects associated with having lighter skin within American society. They found that “darker skin respondents [were] consistently less civil libertarian than their lighter-skin counterparts,” suggesting a harsher and more pessimistic view of society that could easily be correlated to “lighter-skin Blacks exhibiting higher educational and occupational attainments” (Seltzer and Smith 284). Continuing with the use of surveys to elicit information about African American attitudes towards skin color, Verna Keith and Dr. Cedric Herring looked back a few decades to find that skin stratification had played a role in life outcomes in the 1970s and 1980s. Essentially, they found that the skin tone of an African American individual was a stronger predictor of life outcomes in regard to employment, salary and other successes than any other “background characteristics,” further pointing to a need to accurately measure the strength of this effect in politics given its general meaning in life. My study aims to further show these claims on a causal level by implementing the experimental methods of political psychology in order to better elucidate the story of skin tone.
History of Race and Skin Tone in the US

From a historical perspective, there is much to be said about the initial generation of a hierarchy based on skin tone. In *Living Color: The Biological and Social Meaning of Skin Color*, author Nina Jablonski sought to better understand the scientific ties between the progression of differentiation between skin tones and why these differences have mattered in society. She found that tensions caused as a result of skin tone have lasted for centuries, even stemming back to the time of the 18th century philosopher Immanuel Kant (Jablonski 134). Europeans and those with lighter skin tones were linked with positive qualities while Africans and those with darker skin were deemed inferior. This thought process then continued with the implementation of slavery in the United States as nearly all Africans brought to America were owned and worked solely for others. Thus realizing this negative connotation with dark skin, African Americans over time turned to skin bleaching to be accepted by society (Jablonski 173). In an examination of American history with a focus on the role of skin tone, *Race, Gender and the Politics of Skin Tone* by Margaret Hunter traces the permeation of skin tone in the development of culture within the African American and Hispanic populations. She found that racialization came early in the African American community, as European colonists came in contact with African slaves and created mixed children, leading to immediate preference for lighter skin because of the clear observation of European blood (Hunter 20). Similarly, within the Hispanic population, the interactions between Europeans and the darker
indigenous people immediately provoked a status that categorized different skin
tones in the mixing of people due to the same ideas of superiority of European
culture (Hunter 21). *The Color Complex: The Politics of Skin Color in a New
Millennium* by Kathy Russell-Cole, Midge Wilson and Ronald Hall, which also
takes anecdotes from the past to contextualize skin color in America today, also
confirmed this reality from times of slavery. Within the United States, lighter skin
tones initially came about because of “White masters[s] having sex with [their]
own female slaves,” which led to the preference for lighter tones because of the
relationship to European roots (Russell-Cole et al 16). With the continuation of
mixing and the generation of a wider range of skin tones, there was a clearer
distinction between those who had more or less European blood in them and were
thus more or less preferred.

**Skin Tone in Culture: Impact on Majority and Minority
Populations**

It is crucial to note how skin color has also affected various racial groups.
In looking at the vast amount of immigration that has occurred since the inception
of America, the introduction of new groups of people and the resulting mixtures
of skin tones has widened the spectrum from a dichotomy of Black and White
(Hunter et al 173). This has also meant that both African Americans and Hispanic
Americans have faced a great amount of racial discrimination for being minority
groups in the United States. Hunter studied this reality with UCLA professors
Walter Allen and Edward Telles, concluding that skin tone was a trait of significance, and “influences social relations in contemporary American society” (Hunter et al 180). To prove this reality across racial groups and geographies, Christina Gómez studied Puerto Rican and Dominican men and women who live in the Northeast United States. While this is an area stereotyped to be more racially tolerant, there was still a discrepancy in wages amongst men with darker or lighter skin tones, extending the results of prior studies of Latino populations in America (Gómez 94). Not only is this reality common throughout the United States, but it has also been noted internationally. Within the global Hispanic community, similar attitudes towards colorism were tested in a 2002 study run by Eric Uhlmann, Nilanjana Dasgupta, Angelica Elgueta, Anthony Greenwald and Jane Swanson, as Chileans and American Hispanics both rated people of different skin tones, consistently preferring lighter skinned people (Uhlmann et al 198). Stemming from the looks of Europeans once again, the preference for lighter skin tones is a global phenomenon in determining one’s place in society.

Within the African American community, the discussion of the role skin tone plays in determining social circumstances (such as jobs or marriage) has been going on for decades. In a 1973 study of “Black Urbanites,” Ozzie Edwards found that the color scale has been likened to that of a “caste system,” in which people tend to find themselves caught with limited mobility because of the attention paid to the external factor of skin tone alone (Edwards 473). This again confirms the need for further investigation by the academic community as colorism clearly permeates many facets of culture and society. In a longitudinal
study conducted by Mark Hill that followed men from childhood and throughout adulthood, it was commonly shown that “color bias (colorism) rather than family background may be responsible for the bulk of color differences in the socioeconomic status of African American men” (Hill 1437). The scope of this bias spans time and space, again suggesting the relevance of the topic to better understanding the functioning of the forces at work in American culture and society, specifically throughout the political sphere since it has not been the focus of prior research.

**Perspectives on Skin Tone in Modern American Culture**

There are many fields that have looked at the power skin tone holds in society, including history, sociology and psychology. After looking at some of the historical ties, this section seeks to better grasp how skin color has been understood in social science. Within the realm of sociology, Margaret Hunter has done work to better explain colorism, or “skin color stratification, …[which] privileges light-skinned people of color over dark” throughout the world (Hunter 237). Not only are these racial concepts held generally throughout the American population, but they also penetrate society through outlets like the media. As a result, this has promoted a culture of people taking steps to achieve lighter shades of skin to create better outcomes in terms of social mobility.

Placing the role of skin tone in a broader discussion of social psychology, research surrounding the implicit and explicit biases that shape opinions has been
the focus of certain studies. In a 2002 study conducted at Tufts University, Keith Maddox and Stephanie Gray found that skin color affected how people were organizing information and creating stereotypes for people of different skin tones within the same racial category (Maddox and Gray 250). This finding suggests that the implicit components of judgment take a very active role in how the brain interprets difference. It also proposes that better understanding this system of schemas may help to decode why certain patterned societal outcomes prevail for those of lighter and darker skin tones. In Nayda Terkildsen’s surveys of white respondents rating African American candidates of different complexions, there seemed to be different cognitive processes used to analyze light-skinned versus dark-skinned candidates as well (Terkildsen 1032). This is crucial to interpreting political psychology, which uses interdisciplinary methods to understand these realities within the political sphere.

Contextualization of Skin Tone in Politics

After realizing the effects that skin tone plays in broader culture and society, it is interesting to now place this discussion within the realm of political science. Much of this thesis was inspired by Vesla Weaver’s attempt to better grasp colorism towards African American politicians in her study “The Electoral Consequences of Skin Color: The ‘Hidden’ Side of Race in Politics.” When surveying white respondents viewing African American photos, Weaver found that there was significance in the implicit biases that are associated with skin
color (Weaver 159). While skin color was more discernible in some tests than others, it consistently drove how people formed stereotypes based on how near or far the skin tone of the candidate they saw matched what they had pictured as a ‘Black man’ in their head (Weaver 159). Within the attempt to draw ties between skin tone and the political sphere, there has been debate as to whether skin color can predict views on policy. In their piece examining Blacks and Latinos, Jennifer Hochschild, Traci Burch and Vesla Weaver found that these groups “have moved increasingly toward in-group identification over the past century,” suggesting that skin tone impacts perception of race, which then drives political outcomes (Hochschild et al. 1). In better understanding the nuances of skin color in broader society as compared to politics, Hochschild and Weaver also find there to be unique complexes that arise within African Americans. They suggest that while there is stratification amongst blacks in society, race is the more deterministic characteristic in politics (Hochschild and Weaver 643). This finding proposes that until race is brought to an even playing field, skin tone does not take on a meaningful role as a factor within race in terms of voter opinions. However, these conclusions were drawn from examination of national surveys, as opposed to running an experiment in which skin tone can have a more direct effect. In this particular experiment, I hope to show that there is a strong deterministic factor that skin color holds amongst reviews of politicians. This study will feature candidates and respondents of different races, which will help to shed further light on cognitive processes and outcomes of skin tone across racial groups.
In understanding the ways that people tend to rate political candidates, there are many attributes that are necessary to keep in mind other than appearance. In taking a closer look at how voters tend to form judgments about politicians, Patrick Pierce studied the extent to which it is the implicit biases stemming from surface-level features, as opposed to the more explicit policy thoughts, that drive decisions. He found that “the extent to which voters engage in rationalization involving candidate personal traits and issue positions is related to political sophistication” (Pierce 21). Thus, assuming that most voters are forming opinions with limited information on candidates, the more weight appearance holds. This is even more evident in the age of media where image is key and there is a lack of data people receive about the true policy positions of candidates (especially in lower-level elections). In the extremely polarized political system that reigns within U.S. politics, the interaction between skin tone preference and party affiliation cannot be denied as this reality persists by means of in-group preferences. In a 2009 study that looked at faces of biracial candidates, Eugene Caruso, Nicole Mead and Emily Balcetis found that respondents evaluating pictures of politicians of the same party seemed to prefer lightened pictures of candidates and assumed darkened pictures represented candidates of the other party (Caruso et al. 20168). This supports the idea that skin tone remains as a major component of political opinions, and that the assumption of lighter being better draws people to want to include those of preferred skin tones in their in-group.
Application of Theories

Taking these theories and history into account, this project has been shaped by my view of colorism and its impact on the study of political science. Operating under the assumption that “lighter is better,” this preference could manifest itself in determining who has a better chance of being elected in a particular race. With the American population becoming increasingly diverse there will be potential shifts in the appearance of those both running for and elected to Congress (e.g. lighter skinned people who may appear as White but identify as African American or Hispanic), which could have broader societal impact. This methodology suggests that skin tone is a driver behind the strength of racial biases, which has a number of implications as the spectrum of colors will continue to broaden in the coming years.
Part III. Methods
Methods Overview

General Introduction

This project draws its results from an original political psychology experiment that I conducted. To test the theory that immediate reactions impact opinions, a survey was created to capture the snap judgments of voting-age Americans on pictures of Congressmen based on their appearance. Official photographs of Caucasian, African American and Hispanic members of the House of Representatives were modified, creating three pictures per politician: the original one, one with a lighter skin tone and an adaptation with darkened skin. This study design allows for comparison both within each subject and across politicians of different races. After gathering information from other projects to create the best tool and distribution methods, I placed this Qualtrics study within the vast participant pool of Amazon’s Mechanical Turk to gather original data for subsequent analysis. The recruitment of participants aimed to find a sample to replicate a cross-section of the American voting population with a diversity of ages, races and locations, albeit somewhat younger and more liberal given the online survey framework. In gathering demographic data from survey respondents in addition to getting the initial reactions to the pictures (by obtaining their views on different characteristics that they believe the politician would embody based on their photo), causal relationships were determined from the statistical tests and regressions. This will shed light on the association between race and skin tone and the implicit biases that subconsciously drive voter decision-making.
Methods Literature Review

A number of studies were precursors to determining the specific design for this particular political psychology experiment. These prior studies were useful not only in identifying photographic research as a strong proxy for voting tendencies, but also in question design and assessing the use of the Mechanical Turk pool for experimentation. One of the first photographic studies to analyze implicit judgments in political judgments came from Princeton University in 2007. Charles Ballew and Alexander Todorov completed a study that examined people’s interpretations of pictures and instant judgment of traits based off of a one-second flash of a picture. In their results, they found significant correlations between voting patterns and strong initial reactions to the pictures (Ballew and Todorov 2007). The implicit biases that people held came into fruition, as faces that were identified as being competent were linked to winners of elections at a very high rate. Given that the exposure to the pictures was only one second, the direct relationship between the distinction of competence and the actual election results suggests that electability stems from more than policy positions alone (Ballew and Todorov). Todorov has also done a series of studies investigating this same line of thinking by looking at other characteristics, which were used toward question creation for this survey. In addition a study of faces conducted by Matthew Atkinson, Ryan Enos and Seth Hill in 2009 found that, while not the deciding factor in elections, there were ties between close elections and the
quality of candidate photos used (Atkinson et al 2009). These studies affected this thesis as they showed that politicians’ appearances do matter in elections and in the formation of opinions as well as the significance of implicit biases.

Expanding upon this study with a specific racial example, Shanto Iyengar, Solomon Messing, Jeremy Bailenson and Kyu Hahn used Barack Obama as an example to show how, when people have limited information about a political candidate (as is the case in many actual elections), they use skin color as a proxy to make decisions about politicians. In the case of Obama, implicit associations led people to have more negative views of darkened pictures earlier in the campaign, alluding to the presence of these colorism issues not only on a theoretical level but also in practice (Iyengar et al 2010). There is an insinuation that colorism becomes the de facto alternate measure of a candidate when there is little known about other aspects of the platform, which is often the case in elections that are less stressed nationally, as in the House of Representatives. This research also concluded that people with more information about the politician take skin tone into account to a lesser extent yet continue to maintain their racial biases (Iyengar et al 2010). This hints that colorism may be a component of the magnitude of the effect of racial biases.

The choice of Mechanical Turk as the method of distribution also came as political science has shifted toward the use of online mechanisms to access study pools. An online platform run by Amazon, MTurk boasts a pool of hundreds of thousands of participants who fulfill a variety of requests for payment. Ranging from social science surveys to simple Google searches and Adobe Photoshop
work, this is a way to recruit participants at a fraction of the cost and time of a lab experiment. In 2010, Adam Berinsky, Gregory Huber and Gabriel Lenz did a full analysis of MTurk and found a number of advantages and limiting factors: since the platform is still somewhat new, the pool of participants have yet to be overexposed to political science surveys and results have mimicked laboratory studies after recruiting participants more inexpensively and quickly; on the other hand, this is still a convenience sample and users tend to be younger and somewhat liberal (Berinsky et al 2012). For the purposes of this research, looking at the younger generations of voters that have grown up in a more diverse America is a desirable sample pool because this will show any potential biases that have been passed down as well as hint towards preferences that may be seen in future elections. Michael Buhrmester, Tracy Kwang and Samuel Dosling also analyzed MTurk in 2011, and found that “MTurk can be used to obtain high-quality data inexpensively and rapidly” especially as subjects in the pool comprise a much broader cross-section of people than comparable websites and groups at academic institutions (Buhrmester et al 2011). After finding this research and looking into techniques to recruit reliable participants in a short amount of time, Amazon’s Mechanical Turk was the much-preferred arena in which to conduct this experiment.
Research Question and Hypotheses

Hypotheses

My research examines the extent to which skin tone plays a role in voters' assessments of political candidates. As American history has presented multiple examples of preference for lighter skin tones, my hypothesis is that a study will show the same biases in the political sphere today. The influence of colorism in America could affect public perception of candidates for political office as interracial mixing has visibly created a much broader range of colors, and I am looking to see this effect amongst voters when they examine faces of politicians (House Representatives for the purposes of this project). This is especially important in understanding implicit judgments because skin tone could affect the magnitude of the consequential significance of race in elections. The research will also shed light on the formula for a successful political race in examining how appearance may present another factor of assessment of a candidate in addition to their particular policy platform.

Hypothesis #1: Skin tone shapes perceptions of politicians.

There are a few overarching hypotheses that guide this work and differentiate it from other studies in the field. The first and most overarching proposition of this research is the suggestion that skin tone matters in voters’ perceptions of politicians and political candidates. While political scientists and psychologists have based many experiments and projects on better understanding
the role of race in perceptions and resulting analyses of others, very few have looked at skin color as a factor separate from race. While race is an implicit bias that many people hold, I believe that skin tone may be a means by which the strength of this bias is determined. If people associate lighter or darker tones of skin with a particular stereotype or set of heuristics, that would change how the candidate is judged based on image alone. As a candidate’s physical image is a major component of elections, this would be important.

**Hypothesis #2: Lighter is better.**

Second, I assume that the lighter versions of the photos will have higher proportions of people identify that they would vote for that candidate. This hypothesis is measured by the proportion of “yes” answers that are selected to the question “Would you vote for this candidate?” This pattern of preference for lightened photos should also occur across the other characteristics, as one would assume that these would factor into a person’s decision whether or not to vote for a candidate. I expect to see a statistically significant higher percentage of people who would unknowingly prefer the lighter version of the politician as compared to the normal pictures and the darkened photos. While this is a between-subjects approach, the randomization and large sample that see each photo still presents an opportunity to yield significant results. As it has already been shown empirically that lighter people tend to have more successful careers with higher salaries, the potential crossover to politics is clear. Since politicians, as representatives of their constituents, are consistently judged on their appearance and personas in the
media, the logic implies that lighter-skinned individuals would thus be preferred. Also, given the preference for lighter-skinned people across a number of different racial groups, this effect should likely hold true with all of the candidates.

**Variables**

As this project focuses on the electability of a politician based on appearance, the dependent variable most vital to the research is the question of whether the survey participant would vote for the person in the photo they just viewed. In addition, participants are also asked to judge the photos based on the perceived character traits of the politician, looking specifically at likeability, competence, intelligence and trustworthiness. The characteristics listed for each picture are analyzed both for their influence on the likelihood of voting for an individual (dependent variable) and for their ability to be predicted by the aforementioned demographic parameters. Respondents then self-reported across various demographics including age, gender, race, political party, skin tone, etc. These factors, when paired with the treatments of randomly seeing a lightened or darkened photo or the control of the original picture, became ways to determine what characteristics of the viewing audience seem to drive judgments. Once these various qualities are coded using a statistical program, they are paired to form two-sample t-tests and regressions that then inform the implications resulting from the data. Essentially, random assignment to a dark, light or normal picture will be tested to find the drivers in determining the views about the different questions (i.e. the characteristics), which were the dependent variables. The
demographic variables collected from survey participants are heterogeneous
treatment effects that may also be interacted with the treatment conditions when
forming these causal predictions to find the most meaningful traits of voters that
drive their opinion formation.

**Experimental Design**

**Use of Survey Instrument**

In the realm of political psychology, surveys are commonly utilized as
efficient tools for gathering information to create original data sets. In order to
best simulate true voting practices with a large number of people within budget
constraints, the method chosen for this was an online survey run on Amazon’s
Mechanical Turk. The survey is informative because of its ability to provide
insights into voter thinking through the respondents’ perceptions of politicians
whilst also detailing aspects of their personal background that shape these
judgments. Using the online platform aids in the ease of recruiting a diverse pool
of participants that is much more difficult to recreate in a laboratory setting.

With the data that was accumulated, a number of statistical tests were
performed that resulted in inferences that may have widespread national political
implications. Since this is an experiment, the survey instrument also provided a
means to separate the control and experimental groups and guarantee
randomization throughout the test. Running preliminary tests to ensure an
accurate reading of respondent judgments and applicability to the American
election voting population is also important for both internal and external validity. There
are also limitations, however, because demographic information cannot be
assigned and cannot show causality and because Mechanical Turk tends to cater
to an overall younger, more liberal pool. Taking everything into account, the
survey method poses the best framework for this particular experiment to obtain
accurate and noteworthy results with this specialized data set.

Survey Design

In the 45-question survey, respondents are shown a photograph of a
politician for two seconds and then asked a series of questions, reminiscent of the
2006 Willis and Todorov study. The wording of the survey questions is based on
the aforementioned research, measuring competency, likeability, trustworthiness,
intelligence and electability (see table below). Each respondent viewed a total of
six total photos, each of a different politician: all respondents saw the same six
Congressmen, but photos were randomized such that the respondent saw either
the original, lightened or darkened version of each congressman. While an
individual could see only one version of each politician’s photo throughout their
particular survey, in aggregate, approximately the same numbers of people
evaluated all three versions of each politician’s photo. Questions two through five
(see figure below) were randomly ordered, with Question 1 always appearing
first.
Figure 1. Questions Asked for Each Photo

<table>
<thead>
<tr>
<th>Question 1</th>
<th>Would you vote for this candidate?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 2</td>
<td>How likable did this candidate appear to be?</td>
</tr>
<tr>
<td>Question 3</td>
<td>How competent did this candidate appear to be?</td>
</tr>
<tr>
<td>Question 4</td>
<td>How intelligent did this candidate appear to be?</td>
</tr>
<tr>
<td>Question 5</td>
<td>How trustworthy did this candidate appear to be?</td>
</tr>
</tbody>
</table>

Each question was rated on a 4-point scale, ranging from the politician in the picture not embodying the characteristic to fully portraying that trait. After gathering responses from each participant, this allows for a calculation of the mean response to use in regressions. After completing the analysis of the pictures, those taking the survey were then asked a series of demographic questions from age and location in the U.S. to political views and their own skin tone. While remaining anonymous, there were a number of data points gathered that could be used as potential predictors in determining causal relationships. There were also spaces for participants to speculate on what the survey was measuring as well as provide their own comments upon completion. In total, it was designed as about a 5 minute 30 second survey that could be placed online to pay $0.20 per test, in line with other projects of the type on the website.

The treatments in this experiment came in the form of the version of the photo that they were randomly assigned to view and judge. In this double-blind format where neither the participant nor I knew which treatment would be assigned, the averages of each characteristic can be compared across the different
skin tones. While demographic information cannot be assigned, it clearly affects opinion formation, and can be run in regressions as it interacts with other variables. The figure below demonstrates the experimental design for a given candidate, in which a respondent would see one form of the photo and would answer the randomly ordered questions about the 5 characteristics.

**Figure 2. Explanation of Treatments**

<table>
<thead>
<tr>
<th>Candidate X</th>
<th>Candidate X</th>
<th>Candidate X</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Light</td>
<td>• Normal</td>
<td>• Dark</td>
</tr>
<tr>
<td>• Competent</td>
<td>• Competent</td>
<td>• Competent</td>
</tr>
<tr>
<td>• Likeable</td>
<td>• Likeable</td>
<td>• Likeable</td>
</tr>
<tr>
<td>• Intelligent</td>
<td>• Intelligent</td>
<td>• Intelligent</td>
</tr>
<tr>
<td>• Trustworthy</td>
<td>• Trustworthy</td>
<td>• Trustworthy</td>
</tr>
<tr>
<td>• Would you vote</td>
<td>• Would you vote</td>
<td>• Would you vote</td>
</tr>
</tbody>
</table>

**Choice of Politicians**

The choice of politicians to use in this research was based on a few major factors. First, the photos used for this study are all of Congressmen that are currently in office. The advantage of this feature is that the photos all already have a very similar layout with a basic background, a politician in formal attire and a flag placed just behind the individual. Congressmen from the House of Representatives were used for this study in particular because of the greater
diversity seen in this voting body as compared to the other branches of government. In addition, I considered the level of general anonymity that representatives hold due to the large populace of the group in comparison to Senators. These members were also not ones who have high name recognition so as to prevent respondents from being able to easily identify the politician and produce conflated judgments based on preconceived perceptions of the individual. When selecting photos, the party affiliation of the politician was not considered.

Secondly, this study only uses male members of Congress in order to isolate the variable of skin tone by eliminating the complexity that comes with an analysis of gender relations. Since women are still in the vast minority within Congress, holding only about 18% of the seats, they tend to be more recognizable and thus more difficult to use in this experimental design. Finally, I selected photos for inclusion based on stereotypical phenotypes and potential for believable manipulation. When thinking of facial features, I looked for consistency with general sentiments (i.e. stereotypical phenotypes) so as to deter survey respondents from having clouded opinions due to any confusion of the race of the respondent. When the skin color was darkened and lightened, the ideal was to make sure the perceived race of the politician remained intact. After completing the vetting process for politicians to use in the study, the Congressmen that emerged were the following:

David Scott (GA-D)--African American

Emanuel Cleaver (MO-D)--African American

Juan Vargas (CA-D)--Hispanic
Ben Lujan (NM-D)--Hispanic
Doug LaMalfa (CA-R)--White
Ted Yoho (FL-R)--White

Photo Manipulation

The photos were manipulated using an editing tool called Gimp, which is an open-source software similar to Adobe Photoshop. In this program, the coloring of pictures is modified by adding layers to the different portions of the image and then retouching certain sections of the original. For the purposes of this study, there was a brown color chosen and saturated to ~90-100% when placed over the face of the politician. Conversely, a shade of pink was used for the White and Hispanic pictures and a very light brown used with the African American photos to lighten them using the same process. The same colors could be utilized with the many faces because the layer solely tinted the face and could provide the same lightening or darkening effect regardless of the original color of the face. The goal was to create three versions of the photograph: the original, a copy that was 2-3 shades more “pale,” and a copy that was 2-3 shades more “tan.” These levels created a distinct difference in appearance, but one that would not be detected when looking at an individual picture. In order to determine the saturation of photos, each candidate had a clear distinction in skin color that appeared to be to the same degree across all six candidates.
**Figure 3.** Example of manipulated pictures (Emanuel Cleaver)

<table>
<thead>
<tr>
<th>Light</th>
<th>Original</th>
<th>Dark</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Light Image" /></td>
<td><img src="image" alt="Original Image" /></td>
<td><img src="image" alt="Dark Image" /></td>
</tr>
</tbody>
</table>

**Pilot Testing**

Before deploying the final survey tool, a pilot test was run on 146 participants from Mechanical Turk (Mturk) and 85 participants from Harvard University in order to collect preliminary data on the photos of the politicians. There was a $0.20 payment given for completing the task in Mturk and a raffle for a $10 gift card amongst students in order to solicit a large number of responses. In the test, one of the three versions of each politician (i.e. original, lightened or darkened) was shown with two questions accompanying the photo. The pictures did not flash so that any irregularities with the picture would be noted. This was done for all six candidates in question to see if any adjustments would need to be made for the various photos. At the end of the pilot test, respondents provided
basic demographic information so as to get a better sense of the sample that was being collected.

**Figure 4. Questions Asked for Each Photo in the Pilot**

<table>
<thead>
<tr>
<th>Question 1</th>
<th>Does anything about this picture seem weird or abnormal? Does it seem normal and fine?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 2</td>
<td>What race do you think this person is? Please be as specific as possible.</td>
</tr>
</tbody>
</table>

Both of the questions required respondents to analyze the pictures thoroughly in order to point out any potential flaws in the coloring and to better contextualize the material they were viewing. Question 1 encourages the person taking the survey to focus on minute aspects of the photo in addition to the full frame in order to best assess the proposition of potential abnormality. As participants may feel obliged to fully answer the question in order to ensure payment for their work, there is a tendency to be very detailed in pinpointing any faults of the picture; note that the survey instrument used in the experiment features pictures that flash for only a few seconds, and are thus subjected to less scrutiny. Question 2, which asks about the race of the candidate, is designed to detect a potential interaction between skin tone and race. As lighter or darker skin tones could produce different readings of race (especially between the White and Hispanic pictures), this may yield different answers due to any biases that are associated with a different race that would affect judgments due to the perceptions that accompany racial stereotypes. With the concern that the results of the survey
could be driven by racial biases as opposed to ones related to skin tone, the pilot gains importance as a means of validation for the final results. Since respondents were overwhelmingly consistent in their assessments of perceived race, we can be assured that the final survey is truly testing for differences in the skin tone of the politician as opposed to racial groups. Ultimately, the pilot tested the validity of the photos in order to assure that they did not detect noticeable imperfections, and finalized the six candidates and 18 treatment conditions seen in the final survey.

This pilot testing also served as a means to becoming acclimated to running a survey of this kind on Mechanical Turk. There were certain keywords that seemed to help the survey (e.g. “quick” or “short”), and the results confirmed that a price tag of $0.20 was reasonable. The responses also came in somewhat rapidly, with nearly 100 responses gathered in a matter of a few days. Finally, when looking at the demographic information, it gave insight on the population that could be expected for the final survey. The respondent pool showed a range of people not only racially, but also in terms of age, location and gender. With regard to building a power calculation for sample size of the final project, it was also useful to extend estimates to ensure a large enough number of African American and Hispanic people were included to derive significant probabilities in the regressions and hypothesis testing (based on the proportions of respondents from the pilot test).

In addition, the results previewed a potential interaction between skin tone and racial categorization. The Hispanic candidates, Ben Lujan and Juan Vargas, were often marked as White/Caucasian, especially in their lightened form. If
people hold different stereotypes towards Hispanic populations versus White populations, this iteration could be problematic because racial judgments could drive respondent opinions as opposed to those based on skin tone. After screening for this and other comments about the elements of the pictures themselves, I deemed the project ready to run because no patterns were found in the comments about the photos and any discrepancies between White and Hispanic politicians that were identified in the pilot could be adjusted through editing.
**Figure 5. Pilot Test Results**

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Photo Version</th>
<th>% Who Guessed Race Correctly</th>
</tr>
</thead>
<tbody>
<tr>
<td>LaMalfa</td>
<td>Light</td>
<td>100%</td>
</tr>
<tr>
<td>LaMalfa</td>
<td>Normal</td>
<td>100%</td>
</tr>
<tr>
<td>LaMalfa</td>
<td>Dark</td>
<td>95%</td>
</tr>
<tr>
<td>Yoho</td>
<td>Light</td>
<td>100%</td>
</tr>
<tr>
<td>Yoho</td>
<td>Normal</td>
<td>100%</td>
</tr>
<tr>
<td>Yoho</td>
<td>Dark</td>
<td>100%</td>
</tr>
<tr>
<td>Cleaver</td>
<td>Light</td>
<td>96%</td>
</tr>
<tr>
<td>Cleaver</td>
<td>Normal</td>
<td>100%</td>
</tr>
<tr>
<td>Cleaver</td>
<td>Dark</td>
<td>97%</td>
</tr>
<tr>
<td>Scott</td>
<td>Light</td>
<td>100%</td>
</tr>
<tr>
<td>Scott</td>
<td>Normal</td>
<td>97%</td>
</tr>
<tr>
<td>Scott</td>
<td>Dark</td>
<td>99%</td>
</tr>
<tr>
<td>Vargas</td>
<td>Light</td>
<td>48%</td>
</tr>
<tr>
<td>Vargas</td>
<td>Normal</td>
<td>56%</td>
</tr>
<tr>
<td>Vargas</td>
<td>Dark</td>
<td>59%</td>
</tr>
<tr>
<td>Lujan</td>
<td>Light</td>
<td>28%</td>
</tr>
<tr>
<td>Lujan</td>
<td>Normal</td>
<td>38%</td>
</tr>
<tr>
<td>Lujan</td>
<td>Dark</td>
<td>69%</td>
</tr>
</tbody>
</table>
Figure 6. Excerpt of the Pilot Test (Doug LaMalfa, lightened version)

Does anything about this picture seem weird or abnormal? Does it seem normal and fine?

What race do you think this person is? Please be as specific as possible.

- White/Caucasian (1)
- Black/African American (2)
- Hispanic/Latino (3)
- Asian/Pacific Islander (4)
- American Indian (5)
- Other (6) ____________________
Summary of Methods

Methods of Data Analysis

Upon collection of the data in the Qualtrics survey platform, the recorded responses were then downloaded to a comma separated values format to be imported into the statistical software program $R$. In this software, the variables were then named to clearly delineate the different demographic aspects (age, gender, location, etc.) as well as ratings of each characteristic (trustworthy, competent, etc.). Once the labeling took place in $R$, there were a number of tests run on the data to derive causal inferences from the data. Both t-tests and regressions were performed on the data to draw conclusions from the results.

The t-test examined the difference of means in the proportions of people who said they would vote for the politician (i.e. responded “yes” to the question “Would you vote for this candidate?”) in the various versions of the photographs. The t-test was run between each modified photo and the original as well as pairing the dark and light picture to one another. The original version served as the control in terms of monitoring the proportions of those who would vote for the politician. An ANOVA regression was also performed in order to see any variance in patterns across photos, which showed that there were differences in the ratings across the participant demographics.

Multiple regression was the other method that was used to test the differences to see the effects of skin tone. The skin tone of the photograph, used as the primary independent variable as the determinant for voting, was paired with
other demographic information to predict the voting behaviors of respondents.
These mixed effects regressions were run with potential interactions between characteristics such as location and skin tone of the respondent. In addition, different demographic characteristics were tested in order to most accurately predict the factors that account for readings of competence, likability, trustworthiness and intelligence. The likelihood of voting for a politician was also an effect measured by the aforementioned characteristics, to see which carries the most weight in determining inclination towards or against certain political figures.

By running interaction terms as well within the regressions to see how race and age interacted with the color treatments of the photos, a relationship was shown to exist in both categories as to how they tend to judge politicians (i.e. judgments in younger versus older voters and voters of different races).

With the variety in the demographic information collected in the survey, the most respondent- and politician-specific important factors that drive judgments were discovered throughout statistical tests. In addition to finding the most important information about the respondents in terms of what causes them to form certain opinions, the character traits that were most crucial in the analysis of politicians were also availed through interactions in the tests. As this is a randomized experiment, skin tone was determined to cause differences in ratings of photos with statistical significance across each of the characteristics. The comparison of the results between the lightened and darkened photographs, as well as looking across the politicians of different races, also aided in yielding results across multiple dimensions in order to best understand how ratings are
determined for politicians of different races. In this way, there was increased clarity in the photo assessment process and the implicit and explicit biases that drive them, especially as a result of different demographics.
Sample Portion of the Survey Tool (see Appendix A for the full survey)

Figure 7. Excerpt of the Final Survey (Ben Lujan, darkened version)

Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Part IV. Results and Discussion
Data

Coding in Excel

Once the data was downloaded from Qualtrics into Excel, a scrubbing process took place to remove responses deemed inappropriate for inclusion. First, previews of the survey that served as tests that the tool was functioning correctly were thrown out. Next, the five participants who had not agreed to the informed consent were removed. One check for validity within the survey came as respondents were asked a question about a picture that flashed in the same manner as those of the candidates--if the answer was incorrect, that test was thrown out.² Finally, since the IP addresses and MTurk IDs of each participant were recorded, only the oldest attempt (as determined by the time stamp) was kept in the case of duplicate responses. After scrubbing the original data, the refined set ended up having 2,422 responses (out of 2,638 total originally).

Explanation of Variables

- Vote: Coded in the form “Vote_candidate_skin tone,” this variable measured the answer to the yes/no question “Would you vote for this candidate?” (e.g. “Vote_lamalfa_light”).

² The test used to check for validity in this survey was one created for a project used in a recent prototype for a study by Assistant Professor Ryan D. Enos. In the task, a picture that showed two bars (one red column and one blue column) flashed for 2 seconds, the same format as the pictures of the politicians. The respondent was then asked how many colored bars there were in the image. Thus, the task served two purposes--to see if the respondent was truly paying attention to the assignment as well as to prepare them for how the actual experiment would function (see Appendix A for a picture of the chart that was shown).
• **Likeable**: Coded in the form “Likeable_candidate_skin tone,” this variable measured the answer to the question “How likeable did this candidate appear to be?” (e.g. “Likeable_lamalfa_light”) and was measured on a 4-point scale.

• **Competent**: Coded in the form “Competent_candidate_skin tone,” this variable measured the answer to the question “How competent did this candidate appear to be?” (e.g. “Competent_lamalfa_light”) and was measured on a 4-point scale.

• **Intelligent**: Coded in the form “Intelligent_candidate_skin tone,” this variable measured the answer to the question “How intelligent did this candidate appear to be?” (e.g. “Intelligent_lamalfa_light”) and was measured on a 4-point scale.

• **Trustworthy**: Coded in the form “TrustworthyCandidate_skin tone,” this variable measured the answer to the question “How trustworthy did this candidate appear to be?” (e.g. “Trustworthy_lamalfa_light”) and was measured on a 4-point scale.

• **Gender_respondent**: Measured the self-reported gender of the participant (either male or female).

• **Age_respondent**: Measured the self-reported age of the participant (<19, 20-30, 31-40, 41-50, 51-60, 61+).

• **U.S. Citizen_respondent**: Measured the self-reported citizenship of the participant, who needed to be in the United States to have access to take the survey (either a U.S. citizen or not a U.S. citizen).

• **State_respondent**: Measured the self-reported state of residence of the participant (each individual state was a response, so it could be broken into regions during the data analysis phase).

• **Hispanic_respondent**: Measured whether or not the participant self-identified as Hispanic/Latino.
• **Race_respondent**: Measured the self-reported race of the participant (White/Caucasian, Black/African American, Asian/Pacific Islander, American Indian or other).

• **Party_respondent**: Recorded the self-reported political party of the participant (Democrat, Republican, Independent or other).

• **Strengthviews_respondent**: Measured the extent to which the participant identified as being liberal or conservative on a 5-point scale (ranging from very liberal to very conservative).

• **Skintone_respondent**: Measured the self-identified skin tone of the participant (respondents selected a color based on a wide palette, similar to a make-up sample).

• **Recognize_respondent**: Recorded whether or not a participant recognized one of the politicians, since this could taint their initial judgment if they were familiar with the policies, etc. of a particular representative.

**Findings**

After the coding was completed in Excel, I used the statistical package R to analyze the data. Each variable was coded as a factor or interval variable, with the intent of running descriptive statistics, difference of means t-tests and multi-variable regressions. The statistical tests were run for every candidate, treatment and question asked on the survey, which yielded hundreds of tests. As such, I will highlight certain tests in order to summarize the findings.
General Ratings

To establish a baseline for the ratings, I ran a test for a difference in mean ratings of the original photographs for each candidate across the four major characteristics of competence, likeability, trustworthiness and intelligence. Likelihood of voting for a particular politician is measured on a different scale from the others, thus it was not included in this summary. As each characteristic was measured on a scale with four choices (from not embodying the trait, somewhat not, somewhat and fully embodying the trait), these were then coded to become the 4-point scale used for calculations. In deciding which version of the photos to use in finding a baseline, I felt that the pictures without alteration would be best because the resulting graph then depicts reactions to the official photos released by the House of Representatives. The average ratings across the four traits were then averaged for each candidate and are presented in Figure 8 below.

Figure 8. Means of Main Characteristics for Each Candidate
As Figure 8 above shows, there was minimal variation between the overall judgments of the candidates. This is preferable because an analysis can be performed across the candidates without major outliers. While there was not a test that could lead to conclusive reasons as to why there were differences in ratings (as opposed to simply liking one politician or another), the difference in these means could relate to the age of the representatives, as the older politicians had the highest averages. The averages were all in the upper 2/lower 3 range, meaning that there was generally a neutral stance on how much or how little a politician fulfilled a certain characteristic at first glance.
<table>
<thead>
<tr>
<th></th>
<th>Competent</th>
<th>Trustworthy</th>
<th>Likeable</th>
<th>Intelligent</th>
<th>Electable*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lamalfa</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light version</td>
<td>2.818182</td>
<td>2.576588</td>
<td>2.70486</td>
<td>2.798257</td>
<td>1.566625</td>
</tr>
<tr>
<td>Normal version</td>
<td>2.803461</td>
<td>2.619283</td>
<td>2.76391</td>
<td>2.78492</td>
<td>1.530284</td>
</tr>
<tr>
<td>Dark version</td>
<td>2.785185</td>
<td>2.581481</td>
<td>2.71605</td>
<td>2.739506</td>
<td>1.606173</td>
</tr>
<tr>
<td><strong>Yoho</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light version</td>
<td>2.930864</td>
<td>2.819753</td>
<td>3.00864</td>
<td>2.903704</td>
<td>1.445679</td>
</tr>
<tr>
<td>Normal version</td>
<td>2.897628</td>
<td>2.776529</td>
<td>2.98502</td>
<td>2.878901</td>
<td>1.47191</td>
</tr>
<tr>
<td>Dark version</td>
<td>2.885327</td>
<td>2.695438</td>
<td>2.90012</td>
<td>2.848335</td>
<td>1.511714</td>
</tr>
<tr>
<td><strong>Cleaver</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light version</td>
<td>3.166047</td>
<td>2.959108</td>
<td>2.87732</td>
<td>3.226766</td>
<td>1.375465</td>
</tr>
<tr>
<td>Normal version</td>
<td>3.131127</td>
<td>2.954657</td>
<td>2.92892</td>
<td>3.234069</td>
<td>1.377451</td>
</tr>
<tr>
<td>Dark version</td>
<td>3.100125</td>
<td>2.937422</td>
<td>2.84856</td>
<td>3.172716</td>
<td>1.394243</td>
</tr>
<tr>
<td><strong>Scott</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light version</td>
<td>2.932416</td>
<td>2.822278</td>
<td>3.08761</td>
<td>2.988736</td>
<td>1.440551</td>
</tr>
<tr>
<td>Normal version</td>
<td>2.937346</td>
<td>2.821867</td>
<td>3.03317</td>
<td>2.976658</td>
<td>1.422604</td>
</tr>
<tr>
<td>Dark version</td>
<td>2.873918</td>
<td>2.739184</td>
<td>2.95056</td>
<td>2.887515</td>
<td>1.487021</td>
</tr>
<tr>
<td><strong>Vargas</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light version</td>
<td>2.805419</td>
<td>2.745074</td>
<td>2.87685</td>
<td>2.820197</td>
<td>1.495074</td>
</tr>
<tr>
<td>Normal version</td>
<td>2.882939</td>
<td>2.763387</td>
<td>2.94396</td>
<td>2.864259</td>
<td>1.485679</td>
</tr>
<tr>
<td>Dark version</td>
<td>2.805452</td>
<td>2.672862</td>
<td>2.87113</td>
<td>2.826518</td>
<td>1.526642</td>
</tr>
<tr>
<td><strong>Lujan</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light version</td>
<td>2.726601</td>
<td>2.642857</td>
<td>2.78325</td>
<td>2.742611</td>
<td>1.589901</td>
</tr>
<tr>
<td>Normal version</td>
<td>2.74042</td>
<td>2.660074</td>
<td>2.86403</td>
<td>2.719407</td>
<td>1.562423</td>
</tr>
<tr>
<td>Dark version</td>
<td>2.705368</td>
<td>2.612984</td>
<td>2.78402</td>
<td>2.670412</td>
<td>1.58427</td>
</tr>
</tbody>
</table>

*For this scale, 1 corresponded to “Yes” and 2 corresponded to “No” when answering the question “Would you vote for this candidate?”.
After viewing Figure 9, it is clear that there is a clear trend in the data. Across all characteristics, with each candidate, the lighter and original versions of the photos are consistently preferred. This suggests that there is a definite difference in the snap judgments and predicts the direction of this bias as well. While not every test had as low of a p-value, a number of t-tests that were run also showed statistical significance in this difference, meaning that the treatment did matter in how respondents rated politicians. See Figure 10 for the detailed output of each two-sample t-test.

Discussion

Relationship Between Findings and Hypotheses

Hypothesis #1: Skin tone shapes perceptions of politicians

In order to test this hypothesis, I ran t-tests to study the difference of means in the scores between the light and dark photographs. Under the null hypothesis of zero difference in the scores of the light and dark photographs and the alternative hypothesis that the true difference was not equal to zero, I studied the ratings for every candidate and question asked. (Since the point of interest was in seeing the difference between the light and dark photographs, comparisons were not made with the original pictures.) Figure 10 includes the tests that were run, showing that for Caucasian, African American and Hispanic politicians, there
is statistical significance in many tests to show that there is a difference in their
ratings across the characteristics.

These tests are not only meaningful because of their low p-values, but also
because they all feature different attributes and candidates and all show a pattern
in the difference in the means where the lighter photo is rated more highly. Thus,
these tests not only show that there is a difference, but they also point to the
direction of this distinction. With such a large value for degrees of freedom,
suggesting a large sample size with which the calculations were made, these
statistically significant results prove reliable as well. This demonstrates that
colorism is a factor that should be taken into account when analyzing a politician
because it clearly does make a difference how light or dark a person’s skin tone is
when judging one’s character traits based on their picture.
<table>
<thead>
<tr>
<th>Figure 10. T-Tests for Difference in Means</th>
<th>Competent</th>
<th>Trustworthy</th>
<th>Likeable</th>
<th>Intelligent</th>
<th>Electable</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lamalfa</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t)-value</td>
<td>0.9263</td>
<td>-0.1229</td>
<td>-0.2811</td>
<td>1.6584</td>
<td>-1.6128</td>
</tr>
<tr>
<td>( df )</td>
<td>1608.929</td>
<td>1608.996</td>
<td>1605.227</td>
<td>1610.989</td>
<td>1610.163</td>
</tr>
<tr>
<td>( p)-value</td>
<td>0.3544</td>
<td>0.9022</td>
<td>0.7787</td>
<td>0.09743 .</td>
<td>0.107</td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>(-0.04, 0.10)</td>
<td>(-0.08, 0.07)</td>
<td>(-0.09, 0.07)</td>
<td>(-0.01, 0.13)</td>
<td>(-0.09, 0.01)</td>
</tr>
<tr>
<td><strong>Yoho</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t)-value</td>
<td>1.2723</td>
<td>3.2543</td>
<td>2.7311</td>
<td>1.5769</td>
<td>-2.6653</td>
</tr>
<tr>
<td>( df )</td>
<td>1618.592</td>
<td>1618.878</td>
<td>1618.976</td>
<td>1618.993</td>
<td>1618.968</td>
</tr>
<tr>
<td>( p)-value</td>
<td>0.2034</td>
<td>0.00116 **</td>
<td>0.006381 **</td>
<td>0.115</td>
<td>0.007769 **</td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>(-0.02, 0.12)</td>
<td>(0.05, 0.20)</td>
<td>(0.03, 0.19)</td>
<td>(-0.01, 0.12)</td>
<td>(-0.11, -0.02)</td>
</tr>
<tr>
<td><strong>Cleaver</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t)-value</td>
<td>1.866</td>
<td>0.5718</td>
<td>0.7339</td>
<td>1.5842</td>
<td>-0.773</td>
</tr>
<tr>
<td>( df )</td>
<td>1603.627</td>
<td>1603.984</td>
<td>1602.586</td>
<td>1603.999</td>
<td>1603.414</td>
</tr>
<tr>
<td>( p)-value</td>
<td>0.06223 .</td>
<td>0.5676</td>
<td>0.4631</td>
<td>0.1133</td>
<td>0.4397</td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>(-0.00, 0.14)</td>
<td>(-0.05, 0.10)</td>
<td>(-0.05, 0.11)</td>
<td>(-0.01, 0.12)</td>
<td>(-0.07, 0.03)</td>
</tr>
<tr>
<td><strong>Scott</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t)-value</td>
<td>1.609</td>
<td>2.1</td>
<td>3.436</td>
<td>2.8175</td>
<td>-1.8692</td>
</tr>
<tr>
<td>( df )</td>
<td>1605.998</td>
<td>1605.509</td>
<td>1602.357</td>
<td>1604.635</td>
<td>1605.948</td>
</tr>
<tr>
<td>( p)-value</td>
<td>0.1078</td>
<td>0.03589 * 0.0006054 ***</td>
<td>0.004899 **</td>
<td>0.06177 .</td>
<td></td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>(-0.01, 0.13)</td>
<td>(0.06, 0.16)</td>
<td>(0.06, 0.22)</td>
<td>(0.03, 0.17)</td>
<td>(-0.10, 0.00)</td>
</tr>
<tr>
<td><strong>Vargas</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t)-value</td>
<td>0</td>
<td>1.8775</td>
<td>0.1419</td>
<td>-0.17</td>
<td>-1.2703</td>
</tr>
<tr>
<td>( df )</td>
<td>1608.695</td>
<td>1612.28</td>
<td>1609.454</td>
<td>1589.839</td>
<td>1616.963</td>
</tr>
<tr>
<td>( p)-value</td>
<td>0.9993</td>
<td>0.06063 .</td>
<td>0.8871</td>
<td>0.8613</td>
<td>0.2041</td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>(-0.07, 0.07)</td>
<td>(0.00, 0.15)</td>
<td>(-0.07, 0.08)</td>
<td>(-0.08, 0.06)</td>
<td>(-0.08, 0.02)</td>
</tr>
<tr>
<td><strong>Lujan</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t)-value</td>
<td>0.5935</td>
<td>0.7842</td>
<td>-0.0193</td>
<td>2.0373</td>
<td>0.2296</td>
</tr>
<tr>
<td>( df )</td>
<td>1610.921</td>
<td>1610.209</td>
<td>1610.205</td>
<td>1610.558</td>
<td>1610.604</td>
</tr>
<tr>
<td>( p)-value</td>
<td>0.553</td>
<td>0.433</td>
<td>0.9846</td>
<td>0.04179 *</td>
<td>0.8185</td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>(-0.05, 0.09)</td>
<td>(-0.04, 0.10)</td>
<td>(-0.08, 0.08)</td>
<td>(0.00, 0.14)</td>
<td>(-0.04, 0.05)</td>
</tr>
</tbody>
</table>

Significance codes: \( * \) \( p \leq 0.05 \); \( ** \) \( 0.01 < p \leq 0.005 \); \( *** \) \( p \leq 0.001 \); \( * * * \) \( p \leq 0.0001 \). 95\% confidence intervals are rounded to the nearest hundredth; Electable is measured on a scale where 1 = would vote for the candidate and 2 = would not vote for the candidate.
Hypothesis #2: Lighter is better

To reiterate what was seen in Figure 9, the lighter photographs consistently had better ratings in almost every test. This finding is consistent with research that has been done across other facets of life (such as salary or professional titles), yet extends the research into politics as well. Linking the ideas of snap judgments, implicit biases and historical preference, the pattern implies that the hierarchy that favored those with lighter skin historically has become so ingrained in American society that it has become a bias that is held implicitly and is revealed in instantaneous judgments. Not only does this pattern hold, but it also then affects voter mentalities, which determines the makeup of our governing body.

Another component of this research that separates it from studies done in the past is in the sample population. Unlike previous surveys, which featured White respondents rating African American politicians, this experiment looks at both politicians and respondents of multiple racial populations. Simulating what is seen in actual elections, respondents of all races were making judgments about candidates of various races. Taking this fact into account, the results garner a stronger meaning because the heuristic of lighter being better has similarly evolved with different racial groups. Even though participants of many races were giving their opinions, they tended to follow the same pattern, which was seen in the diverse group of politicians.
### Figure 11. Mixed Effects Regressions

<table>
<thead>
<tr>
<th></th>
<th>Competent</th>
<th>Trustworthy</th>
<th>Likeable</th>
<th>Intelligent</th>
<th>Electable</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>3.18***</td>
<td>2.98***</td>
<td>2.93***</td>
<td>3.25***</td>
<td>-0.73***</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>American Indian</td>
<td>-0.17</td>
<td>-0.15</td>
<td>-0.02</td>
<td>-0.12</td>
<td>0.74</td>
</tr>
<tr>
<td>respondents</td>
<td>(0.13)</td>
<td>(0.13)</td>
<td>(0.13)</td>
<td>(0.13)</td>
<td>(0.45)</td>
</tr>
<tr>
<td>Asian/Pacific Islander respondents</td>
<td>0.01</td>
<td>0.02</td>
<td>-0.02</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>Black/African American respondents</td>
<td>0.15***</td>
<td>0.10*</td>
<td>0.09*</td>
<td>0.21***</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>0.06</td>
<td>0.06</td>
<td>0.04</td>
<td>0.10**</td>
<td>-0.20</td>
</tr>
<tr>
<td>respondents</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>Other race</td>
<td>-0.15*</td>
<td>-0.12</td>
<td>-0.19**</td>
<td>-0.11</td>
<td>0.50*</td>
</tr>
<tr>
<td>respondents</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.23)</td>
</tr>
<tr>
<td>LaMalfa</td>
<td>-0.33***</td>
<td>-0.36***</td>
<td>-0.16***</td>
<td>-0.44***</td>
<td>0.96***</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Lujan</td>
<td>-0.41***</td>
<td>-0.31***</td>
<td>-0.07***</td>
<td>-0.50***</td>
<td>1.03***</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Scott</td>
<td>-0.22***</td>
<td>-0.16***</td>
<td>0.14***</td>
<td>-0.26***</td>
<td>0.35***</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Vargas</td>
<td>-0.30***</td>
<td>-0.23***</td>
<td>0.01</td>
<td>-0.37***</td>
<td>0.63***</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Yoho</td>
<td>-0.23***</td>
<td>-0.19***</td>
<td>0.08***</td>
<td>-0.33***</td>
<td>0.49***</td>
</tr>
<tr>
<td></td>
<td>(-0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Independent</td>
<td>0.01</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.30***</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>Other political</td>
<td>-0.15***</td>
<td>-0.22***</td>
<td>-0.13**</td>
<td>-0.15***</td>
<td>1.04***</td>
</tr>
<tr>
<td>party</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.14)</td>
</tr>
<tr>
<td>Republican</td>
<td>0.10***</td>
<td>0.12***</td>
<td>0.11***</td>
<td>0.09***</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>Light photo</td>
<td>0.03*</td>
<td>0.06***</td>
<td>0.05**</td>
<td>0.05***</td>
<td>-0.17***</td>
</tr>
<tr>
<td>version</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.01)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Original photo</td>
<td>0.05***</td>
<td>0.07***</td>
<td>0.08***</td>
<td>0.05***</td>
<td>-0.24***</td>
</tr>
<tr>
<td>version</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.01)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Respondent age</td>
<td>0.00</td>
<td>0.08</td>
<td>0.05</td>
<td>0.00</td>
<td>-0.05</td>
</tr>
<tr>
<td>(19 or younger)</td>
<td>(0.04)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Respondent age</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-0.04</td>
<td>-0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>(20-30)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>Respondent age</td>
<td>0.04</td>
<td>0.05</td>
<td>0.05</td>
<td>0.08*</td>
<td>-0.10</td>
</tr>
<tr>
<td>(41-50)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>Respondent age</td>
<td>0.03</td>
<td>0.07</td>
<td>0.07</td>
<td>0.04</td>
<td>-0.29*</td>
</tr>
<tr>
<td>(51-60)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Respondent age</td>
<td>0.00</td>
<td>0.06</td>
<td>0.03</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>(61+)</td>
<td>(0.07)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.07)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>Male respondents</td>
<td>-0.14***</td>
<td>-0.14***</td>
<td>-0.13***</td>
<td>-0.16***</td>
<td>0.18***</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.06)</td>
</tr>
</tbody>
</table>

| AIC                   | 30647.08  | 32853.98    | 33722.57 | 29768.64    | 18690.31  |
| BIC                   | 30829.06  | 33035.97    | 33904.55 | 29950.62    | 18864.75  |
| Log Likelihood        | -15299.54 | -16402.99   | -16837.28| -14860.32   | -9322.16  |
| Num. obs.             | 14532     | 14532       | 14532    | 14532       | 14532     |
| Deviance              |           |             |          |             | 18644.31  |
| Num. groups: TurkID   |           |             |          |             | 2422      |
| Variance: TurkID      |           |             |          |             | 1.27      |
| Variance: Residual    |           |             |          |             | 1.00      |

***p<0.001, **p<0.01, *p<0.05
Other Findings

In testing my secondary hypothesis of harsher in-group judgments, using a dummy variable to represent times when participants were evaluating candidates of the same race, people showed higher ratings for candidates of the same race across the characteristics tested. This was especially apparent in the linear mixed model, where the strength of the interaction between the race of the candidate and the respondent caused the model to show an error because they were so highly correlated. Whereas I believed that people would be more apt to be harsher within their race because they would feel free to voice stronger opinions with those whom they view as similar, the opposite was true. This suggests that people believe more in those who they presume to be in the same racial category, which makes a case for descriptive representation holding very real meaning in politics as another factor that helps drive electability of candidates. While this could be chalked up to inferred partisanship, including politicians of both parties suggests that in-group preference is the true factor of interest. It is also important when thinking of a voting base, because candidates can still generally count on having the support of their racial group in elections.

Whilst assessing race of respondents as a factor, I found a few major trends in the data. First, when looking across races, African Americans gave photos higher ratings overall, followed by Hispanic/Latinos and finally Whites. As Figure 11 displays, it was statistically significant that African Americans had higher more positive judgments than Whites across the board, with Hispanics also generally voting above the average of White respondents. Thus, it is the White
majority that produces harsher initial judgments as compared to the minority groups. Secondly, Whites also responded to the color treatment of the photos to a significant amount across tests measuring each characteristic tested. After running regressions, it was the Caucasian respondents who had stronger, significant responses to the lighter versions of the photos across the characteristics. Seen in Figure 12, this culminates in an assumption that it is Whites who are holding onto the colorist story most firmly in politics today. Finally, as predicted, the story of lighter being better has become so entrenched in all races such that there is a preference for pictures with lightened skin, regardless of the race of the respondent. Figure 11, based off of a linear mixed effects model, shows statistical significance that the lighter and original versions of the photos were preferred consistently. Thus, skin tone is a definite variable of interest as this pattern exists throughout the voting population.
In addition to proving the hypotheses set forth by this thesis, further regressions shed light on the impact of age, political party and gender on rating tendencies. When the data was subset to only show voters between the ages of 18 and 40, t-tests showed similar responses to the tests within the full data set, again with a preference for lightened photos (see Figure 13). While many would expect this result for an older population, it is interesting that the pattern holds true for those between the ages of 18-40. This signals that a partiality towards lighter skin has permeated the younger voting crowd, and will thus continue in future
elections. This effect was seen across character traits as well (in some cases to a
significant extent), meaning that the consistent results are meaningful.

Figure 13. T-Tests for Difference in Means with Respondents Under 40

<table>
<thead>
<tr>
<th></th>
<th>Competent</th>
<th>Trustworthy</th>
<th>Likeable</th>
<th>Intelligent</th>
<th>Electable</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lamalfa</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-value</td>
<td>0.5727</td>
<td>-0.3348</td>
<td>-0.3622</td>
<td>1.3041</td>
<td>-1.6135</td>
</tr>
<tr>
<td>df</td>
<td>1379.915</td>
<td>1379.728</td>
<td>1372.521</td>
<td>1382.84</td>
<td>1381.827</td>
</tr>
<tr>
<td>p-value</td>
<td>0.567</td>
<td>0.7378</td>
<td>0.7173</td>
<td>0.1924</td>
<td>0.1069</td>
</tr>
<tr>
<td>95%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>confidence</td>
<td>(-0.05, 0.10)</td>
<td>(-0.10, 0.07)</td>
<td>(-0.10, 0.07)</td>
<td>(-0.03, 0.13)</td>
<td>(-0.09, 0.01)</td>
</tr>
<tr>
<td><strong>Yoho</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-value</td>
<td>0.9212</td>
<td>3.252</td>
<td>2.6684</td>
<td>1.2077</td>
<td>-2.6301</td>
</tr>
<tr>
<td>df</td>
<td>1382.972</td>
<td>1382.621</td>
<td>1383.92</td>
<td>1383.68</td>
<td>1383.517</td>
</tr>
<tr>
<td>p-value</td>
<td>0.3571</td>
<td>0.001174**</td>
<td>0.007709**</td>
<td>0.2274</td>
<td>0.01**</td>
</tr>
<tr>
<td>95%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>confidence</td>
<td>(-0.94, 0.11)</td>
<td>(0.05, 0.22)</td>
<td>(0.03, 0.20)</td>
<td>(-0.03, 0.12)</td>
<td>(-0.12, -0.02)</td>
</tr>
<tr>
<td><strong>Cleaver</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-value</td>
<td>2.3338</td>
<td>0.9785</td>
<td>1.1681</td>
<td>1.777</td>
<td>-1.1339</td>
</tr>
<tr>
<td>df</td>
<td>1383.469</td>
<td>1384.863</td>
<td>1383.801</td>
<td>1384.121</td>
<td>1382.924</td>
</tr>
<tr>
<td>p-value</td>
<td>0.01975*</td>
<td>0.328</td>
<td>0.243</td>
<td>0.07579</td>
<td>0.257</td>
</tr>
<tr>
<td>95%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>confidence</td>
<td>(0.01, 0.16)</td>
<td>(0.04, 0.12)</td>
<td>(-0.03, 0.13)</td>
<td>(-0.01, 0.14)</td>
<td>(-0.08, 0.02)</td>
</tr>
<tr>
<td><strong>Scott</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-value</td>
<td>1.0163</td>
<td>1.7345</td>
<td>3.0338</td>
<td>2.2935</td>
<td>-1.2125</td>
</tr>
<tr>
<td>df</td>
<td>1367.999</td>
<td>1367.585</td>
<td>1364.526</td>
<td>1366.761</td>
<td>1367.761</td>
</tr>
<tr>
<td>p-value</td>
<td>0.3096</td>
<td>0.08305**</td>
<td>0.00261**</td>
<td>0.02197*</td>
<td>0.2255</td>
</tr>
<tr>
<td>95%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>confidence</td>
<td>(-0.04, 0.12)</td>
<td>(-0.01, 0.16)</td>
<td>(0.05, 0.22)</td>
<td>(0.01, 0.17)</td>
<td>(-0.09, 0.02)</td>
</tr>
<tr>
<td><strong>Vargas</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-value</td>
<td>-0.163</td>
<td>1.4559</td>
<td>0.2637</td>
<td>-0.4559</td>
<td>-0.8825</td>
</tr>
<tr>
<td>df</td>
<td>1370.43</td>
<td>1370.828</td>
<td>1368.374</td>
<td>1354.693</td>
<td>1372.797</td>
</tr>
<tr>
<td>p-value</td>
<td>0.8705</td>
<td>0.1456</td>
<td>0.7921</td>
<td>0.6486</td>
<td>0.3777</td>
</tr>
<tr>
<td>95%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>confidence</td>
<td>(-0.09, 0.07)</td>
<td>(-0.02, 0.14)</td>
<td>(-0.07, 0.10)</td>
<td>(-0.10, 0.06)</td>
<td>(-0.08, 0.03)</td>
</tr>
<tr>
<td><strong>Lujan</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-value</td>
<td>1.0144</td>
<td>0.8478</td>
<td>0.001</td>
<td>2.3448</td>
<td>-0.2407</td>
</tr>
<tr>
<td>df</td>
<td>1384.428</td>
<td>1385.893</td>
<td>1385.708</td>
<td>1384.13</td>
<td>1385.986</td>
</tr>
<tr>
<td>p-value</td>
<td>0.3106</td>
<td>0.3967</td>
<td>0.9992</td>
<td>0.01918*</td>
<td>0.8098</td>
</tr>
<tr>
<td>95%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>confidence</td>
<td>(-0.04, 0.12)</td>
<td>(-0.05, 0.12)</td>
<td>(-0.08, 0.08)</td>
<td>(0.01, 0.17)</td>
<td>(-0.06, 0.05)</td>
</tr>
</tbody>
</table>

Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 1 ‘.’ 1: 95% confidence intervals are rounded to the nearest hundredth;
Electable is measured on a scale where 1 = would vote for the candidate and 2 = would not vote for the candidate.
When assessing the results as they pertained to different political parties, the results gathered showed an interesting trend. Republicans tended to rate the photos higher than Democrats on the whole, many times to a statistically significant extent. There was a much smaller sample of Republicans to Democrats (14% versus 43% of survey respondents), however, which could skew the findings. Looking at the differences between male and female respondents in Figure 11, it was statistically significant that men were more harsh in their views towards the character traits yet were more likely to say that they would vote for the candidate overall.

**Summary of Data**

In all, these findings provide evidence that supports both hypotheses--skin tone does matter in forming opinions of politicians and lighter skin tone is preferred. Using the aforementioned tests and the descriptive statistics from the data set, the clear pattern that emerges tells a story of a minute difference in one attribute driving a difference in the snap judgments of voters. Today, lighter skin holds the same value that it maintained throughout both time and numerous other facets of society, including permeating the political sphere. As a study, this is important as political theorists begin to more fully construct a story around skin tone because of its widespread applicability. In comparison to past studies, this research takes into account politicians and respondents of various races, extending the results past White respondents judging African American politicians. This has many implications for politicians looking to run for office as well as strategists.
who are looking to market their candidate, especially in a time where appearances carry a lot of weight in decision-making.

**Limitations**

While this study was able to make statistically significant claims, there are always shortcomings in research. For one, the survey ideally would have been more equally stratified across age and party. Working within the time and budget constraints, Mechanical Turk posed the best method to gather a data set large enough to analyze, yet this community tends to be younger and more liberal on average. If older, more conservative people responded differently to the treatment effects, that would not be depicted; however, there were people who self-identified as both conservative and older who were incorporated in these results. Without being able to monitor participants in a laboratory environment, there is no way to regulate factors such as computer brightness and quality, which would affect how the pictures would be judged. Given the tests within the survey, the randomization of photo treatments for respondents and the number of respondents, this does not pose a major area of concern. Additionally, I only examined candidates that were White, Black and Hispanic men in order to keep the number of conditions at a manageable number, yet there are also female members of Congress and Asian Americans are another major minority group. This made sense for the scope of this project, yet could be expanded upon in the future.
Finally, regarding external validity, there is a point of contention because party bases are strong and it is difficult to predict how much information voters have about candidates. This is also an online experiment, yet it is being applied to actual practices seen in the real world. A response to this would be that in low-information elections, this will definitely be a driver if appearance is one of the few factors of value in the decision-making process. Also, given the visual marketing campaigns habitual to elections, these findings are sure to remain important. Lastly, the online component very closely mirrors a laboratory experiment but on a grander scale. Given the sample size, randomization and pilot test results, there is a large case for these findings to contribute to the academic discussion of political strategies and colorism and race in the political sphere.
Part V. Conclusion
Future Implications

The findings yield multiple implications for both the general voting population and politicians/political strategists. With the trends seen in other studies and extended in this experiment, there is much to be said about new perspectives towards political decision-making and tactics.

Supposing that candidate appearance is paramount in elections, knowing how to properly position a politician is key to winning. Using the results of this survey, it may impact the choice of a picture for publicity materials. Campaign managers may look to edit pictures to adjust to a light skin tone or choose pictures with different amounts of lighting, as well as decide how to tailor the advertising component of their campaign based on the skin tone of the candidate. When working with a minority aspirant, the same mentality holds true--while they will generally have more support from those who identify as being of the same race, a lighter minority contender would be easier to market to a White population. Conversely, candidates’ opponents could darken pictures to make the politician less electable, as seen in negative campaign advertising. Thinking strategically, the results of this study are very helpful for campaign managers and political consultants in measuring the electability of an individual, especially on a national level.

As the population of multiracial people in the United States continues to increase, this will likely impact both the pool of voters and those who choose to run for political office. Based on the repetition of the trends in the results, lighter
skin should be expected to be preferred, yet more races of people will be linked to lighter shades as mixing continues. This presents an interesting conundrum, as a politician like Marc Morial, current President of the National Urban League, identifies himself as an African American yet is on the lighter end of the spectrum of skin tone for all Blacks in America. With this dual boost, one would suspect that it might be similar candidates who make up more of the governing body in the future. This also makes a case for descriptive representation, or the idea that people tend to vote for those who resemble them because of supposed similarities in political ideas based on appearance.

With the large sample of younger voters and the statistically significant results demonstrating their affirmation of the stereotype that lighter is better, it is probable that analogous judgments will be made for many years to come. This means two things: this train of thought is being passed down from generation to generation and this data will be pertinent for campaign design. If these ideas continue to exist in younger groups—even when racial relations are much better than they were historically—this suggests that there should be a greater discussion of skin color. Thus, there is still much research to be done regarding colorism, especially within politics, in order to expand the knowledge surrounding the topic.

**Final Thoughts**

While many political theorists have pondered the role of race in the political sphere, very little has been said about skin tone. In designing this project, I hoped to find a connection between the opinions voters hold today and the
history of color hierarchy that has plagued the United States since its formation. Although nuanced within a discussion of race, skin tone itself holds more power than one may think. In society, it is thought to be responsible for determining beauty within minority populations and has affected the opportunities people have received to be promoted or earn higher salaries. In politics, the results lead one to question how skin tone has affected the outcomes of past elections and how political strategists will use this information in the future. There is a clear preference for lighter skin, with the instant implicit opinions participants held coming to the forefront in their explicit evaluations of politicians. In the future, as the amount of diversity in America increases, I wonder whether the light/dark skin profile will drive how race is perceived and votes are cast. As this experiment has shown, all racial groups prefer photographs of politicians with lighter skin, and this will continue to be a point of interest in the future as younger voters hold true to the pattern. Clearly, skin tone impacts our assessment of others in the political sphere and is key to understanding politics in today’s era. The Civil Rights Act of 1964, one of the most crucial pieces of legislation in the 20th century, prevents “discrimination on account of race, color, religion, sex or national origin.” While many have questioned the presence of “color,” this thesis reminds us that skin tone holds true meaning in American politics.
Part VI. Appendix
A. Final Survey Tool
**Thesis Survey--Final Experiment**

Informed Consent: Please consider this information carefully before deciding whether to participate in this research.

Purpose of the research: To examine how people analyze the characteristics of political figures.

What you will do in this research: If you decide to participate, you will be asked to make some judgments about politicians.

Time required: Participation will take less than 10 minutes to complete.

Risks: There are no anticipated risks associated with participating in this study. The effects of participating should be comparable to those you would experience from viewing a computer monitor for less than 10 minutes and using a mouse or keyboard.

Benefits: There are no direct benefits from participating in this study.

Compensation: If you complete the study, you will receive $0.20 for your participation. Confidentiality: Your participation in this study will remain confidential.

Participation and withdrawal: Your participation in this study was completely voluntary, and you may withdraw at any time without penalty. You will receive payment based on completing the study. You may withdraw by informing the experimenter that you no longer wish to participate (no questions will be asked).

To Contact the Researcher: If you have questions about this research, please contact Chelsea Celistan, Harvard University, Department of Government, 1737 Cambridge Street, Cambridge, MA 02138; 617-495-2148; ccelistan@college.harvard.edu

Whom to contact about your rights in this research, for questions, concerns, suggestions, or complaints that are not being addressed by the researcher, or research-related harm: Harvard University Committee on the Use of Human Subjects in Research, 1414 Mass. Ave., Room 234, Cambridge, MA 02138. Phone: 617-495-5459. E-mail: cuhs@fas.harvard.edu

- I hereby give my consent to participate in this study. (1)
- I do not consent to participating in this study. (2)

You will be asked to evaluate the person seen in a picture. There are no right or wrong answers. Each picture will appear for about 2 seconds and then you will be asked to judge the characteristics of the person you saw in the photo. On the next page, you will see one example using a figure in order to see how long the pictures will flash on the screen. This test will be followed by a series of pictures of politicians. After the tasks, you will be asked a few short questions about yourself. The entire survey should take no more than a few minutes.
How many colored bars were there on the chart that you just saw?

- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)

Remember, for the next series of questions, the photos will only appear for about 2 seconds. After the photos appear, you will then be asked to click the answer to indicate how much you agree or disagree with the characteristic of the politician in the photo. Ready to proceed?

- Yes (1)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
Would you vote for this candidate?
- Yes (1)
- No (2)

How likable did this candidate appear to be?
- Not likable (1)
- Somewhat unlikable (2)
- Somewhat likable (3)
- Very likable (4)

How competent did this candidate appear to be?
- Not competent (1)
- Somewhat incompetent (2)
- Somewhat competent (3)
- Very competent (4)

How intelligent did this candidate appear to be?
- Not intelligent (1)
- Somewhat unintelligent (2)
- Somewhat intelligent (3)
- Very intelligent (4)

How trustworthy did this candidate appear to be?
- Not trustworthy (1)
- Somewhat untrustworthy (2)
- Somewhat trustworthy (3)
- Very trustworthy (4)
What is your gender?
- Male (1)
- Female (2)

How old are you?
- 19 years or younger (1)
- 20-30 (2)
- 31-40 (3)
- 41-50 (4)
- 51-60 (5)
- 61+ (6)

Are you a U.S. citizen?
- Yes (1)
- No (2)

What state are you from?
- Alabama (1)
- Alaska (2)
- Arizona (3)
- Arkansas (4)
- California (5)
- Colorado (6)
- Connecticut (7)
- Delaware (8)
- Florida (9)
- Georgia (10)
- Hawaii (11)
- Idaho (12)
- Illinois (13)
- Indiana (14)
- Iowa (15)
- Kansas (16)
- Kentucky (17)
- Louisiana (18)
- Maine (19)
- Maryland (20)
- Massachusetts (21)
- Michigan (22)
- Minnesota (23)
- Mississippi (24)
- Missouri (25)
- Montana (26)
- Nebraska (27)
- Nevada (28)
- New Hampshire (29)
☐ New Jersey (30)
☐ New Mexico (31)
☐ New York (32)
☐ North Carolina (33)
☐ North Dakota (34)
☐ Ohio (35)
☐ Oklahoma (36)
☐ Oregon (37)
☐ Pennsylvania (38)
☐ Rhode Island (39)
☐ South Carolina (40)
☐ South Dakota (41)
☐ Tennessee (42)
☐ Texas (43)
☐ Utah (44)
☐ Vermont (45)
☐ Virginia (46)
☐ Washington (47)
☐ West Virginia (48)
☐ Wisconsin (49)
☐ Wyoming (50)
☐ Other (51)

Are you of Hispanic/Latino origin?
☐ Yes (1)
☐ No (2)

What race/ethnicity are you?
☐ White/Caucasian (1)
☐ Black/African American (2)
☐ Asian/Pacific Islander (3)
☐ American Indian (4)
☐ Other (5) ____________________

Which party do you associate yourself with?
☐ Democrat (1)
☐ Republican (2)
☐ Independent (3)
☐ Other (4) ____________________
## Strength of Views

<table>
<thead>
<tr>
<th>How would you consider your political views? (1)</th>
<th>Very Liberal (1)</th>
<th>Somewhat Liberal (2)</th>
<th>Moderate (3)</th>
<th>Somewhat Conservative (4)</th>
<th>Very Conservative (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

Do you vote in major elections (e.g. Presidential, Congressional, Mayoral, etc.)
- Yes (1)
- No (2)
Based on the graphic seen above, which color most closely matches your skin tone?

- Light 1 (1)
- Light 2 (2)
- Light 3 (3)
- Light 4 (4)
- Medium 1 (5)
- Medium 2 (6)
- Medium 3 (7)
- Medium 4 (8)
- Med-Dark 1 (9)
- Med-Dark 2 (10)
- Med-Dark 3 (11)
- Med-Dark 4 (12)
- Dark 1 (13)
- Dark 2 (14)
- Dark 3 (15)
- Dark 4 (16)

Did you recognize any pictures of the candidates you were rating?

- Yes (1)
- No (2)

What do you think this test intended to measure?

Do you have any comments about the study?
Overview of Survey Demographics

In total, there were 231 respondents who took the pilot test, comprised of 146 responses from Mechanical Turk and 85 Harvard students. As a result, each photo had nearly 80 qualitative ratings. All participants saw the exact same survey, yet the populations were separated into two distinct groups to be analyzed separately. I decided to separate the groups in this manner because I wanted to get a sense of the demographics of respondents from MTurk in order to have a better sense of what to expect when the final survey experiment was run; moreover, I wanted to include the perspectives of Harvard students because I suspected this group would be much more critical of the pictures. Noting the strong ties between social media and photographs and the constant exposure that college students have to photos that have been edited, I thought that any flaws in the picture would be obvious to this group.

The population for the MTurk group was varied, as there was a wide range in age, location and party identification. While most people who took the survey were under the age of 30 and somewhat liberal, the pool was diverse in that respondents fell into every category option within each demographic listed. As expected, this was quite different from the Harvard population, from which students were recruited through one upper-class House e-mail list and one extracurricular student group e-mail list.

In the final experiment, there were 2,631 total responses, meaning that nearly 900 people saw each of the 18 total photographs (6 politicians with 3
treatments each). All of the participants were recruited through postings on Mechanical Turk over about a 14-day period. Using the demographics from the pilot test to predict the variation primarily in race, a power calculation for sample size was constructed with the program G*Power 3. Looking at various calculations for multiple regressions and t-tests within the software, it seemed that having a sample size of about 200 would be ample to attain statistical significance with the chosen computation. Since the goal was to show meaningful results for each racial group (White, African American, Hispanic and Asian), I then used the results from the pilot test to decide how to set the parameters on the MTurk posting. As I could expect about 9% of respondents to be Hispanic and 9% to be African American based on the pilot test, I decided to aim to receive about 2,200 total responses.

Summary Data for the Final Survey

- Male: 59% (1,551 participants)
- Female: 41% (1,080 participants)
- <19 years old: 5% (119 participants)
- 20-30 years old: 59% (1,558 participants)
- 31-40 years old: 22% (580 participants)
- 41-50 years old: 7% (193 participants)
- 51-60 years old: 5% (137 participants)
- 61+ years old: 2% (44 participants)
• Every state was represented, with the most responses coming from California, Florida, Pennsylvania and Texas respectively

• White/Caucasian: 73% (1,926 participants)

• Hispanic/Latino: 8% (200 participants)

• Black/African American: 6% (156 participants)

• Asian/Pacific Islander: 10% (254 participants)

• Democrat: 43% (1,128 participants)

• Republicans: 14% (379 participants)

• Independent: 38% (992 participants)
B. Bibliography


   http://www.hks.harvard.edu/inequality/Seminar/Papers/Hochschild05.pdf.


