The Pennsylvania State University

The Graduate School

Intercollege Graduate Degree Program in Genetics

USING GENOMIC ANCESTRY AND DEMOGRAPHIC VARIABLES TO STUDY PERCEPTION IN HUMAN FACES

A Thesis in

Genetics

by

Megan Alicia Rogers

Submitted in Partial Fulfillment
of the of the requirements
for the degree of

Master of Science

May 2012

The thesis of Megan Alicia Rogers was reviewed and approved* by the following:

Mark D. Shriver
Professor of Anthropology
Thesis Advisor

David Vandenbergh
Professor of Biobehavioral Health
Committee Chair

Stephen Schaeffer
Professor of Biology

Nina Jablonski
Professor of Anthropology

David Puts
Professor of Anthropology

Ross Hardison
Professor of Biochemistry and Molecular Biology

Robert Paulson
Professor of Veterinary and Biomedical Sciences
Head of the Intercollege Graduate Genetics Program

*Signatures are on file in the Graduate School

Abstract

There is substantial evidence supporting the contention that both interindividual and intergroup variability affect our perceptions of human faces. Indeed, there is substantial variability in how persons perceive faces. Like other primates, humans are highly visual and rely on sight much more than they rely on other senses. Humans show a vast capacity for communicating via facial expressions and an incredible facility for distinguishing and remembering faces. It is also clear that our facial perceptive abilities are environmentally modified and need to be explored because we can then understand the roles of facial phenotypes and perception in in-group biases like colorism or racism. Also, by applying this research, we can promote an acceptance and appreciation of phenotypic diversity.

I investigated specifically how people rate genomic ancestry of populations admixed with European and West African ancestry, and the relationship between a person's demographic background (e.g. education level, socioeconomic status, geographical location, racial/ethnic breakdown of family and friends, etc.), and his or her ability to estimate genomic ancestry in human faces. Additionally, I investigated whether sex differences were present among how the stimulus faces were rated, as well. I first collected demographic survey information from each participant. Then, I presented two sets of 42 photos and asked them to estimate the proportion of African ancestry in each face for one set, and categorize each face as "Black" or "White" in the other set. There were 100 participants collected in Georgia and Pennsylvania. My analyses determined that, raters are better than randomly simulated ancestry estimates (p=0.0001). Also, female stimulus faces were rated significantly (p=1.6e-11) better than male stimulus faces. To test whether demographic factors determined one's ability to rate genomic ancestry, I ran a multiple regression, and individual t-tests or ANOVAs. The multiple regression via factor analysis and principle components resulted in none of the demographic factors being significant. For the t- tests and ANOVAs, I applied the Bonferroni correction and only one factor was significant: the participants' self described ability to recognize Black faces (p=0.0006). Three factors did not exceed the significance threshold of 0.00263. These trends, geographical location before the age of ten (p=0.041), participants' current neighborhood (p=0.031), and the proportion of European Americans in their high school (p=0.011), and possible explanations for them will be discussed.

Keywords: Genomic Ancestry, perception, faces

Acronyms: AIMs- Ancestry Informative Markers; WA- West African; ORE- Other Race Effect; EDS- Euclidean Distance Score

Table of Contents

List of Tables

List of Figures

Acknowledgements

First and foremost, I'd like to thank God for putting all things in place for me. Without Him, none of this would be possible. Second, I'd like to thank my parents and family members for their never-ending support throughout my graduate career and my entire life. I am forever grateful to have them. Thank you to Dr. Richard Ordway and the Genetics program, for seeing something in me and giving me the opportunity to study at The Pennsylvania State University. To Dr. Guy Barbato and Robab Katani, thank you for laying my graduate school foundation and teaching me invaluable tools to succeed.

To Dr. Mark D. Shriver, I appreciate your knowledge and expertise regarding genetics. I am thankful to have worked with you to gain essential techniques and the critical thinking needed to excel at any career I choose. I'd also like to thank you for providing me with the ancestry data and the 3D faces for my project. To my lab-mates Jennifer Wagner, Laurel Pearson, Ellen Quillen, and Denise Liberton, thank you for your support, advice and patience. Thank you to Yan Huaru, Stephanie Rossman, Betty Blain, Faye Manning and Kim Miller for making my time at Penn State a stress free breeze, administratively. Finally, thank you to my committee chair Dr. David Vandenbergh for your countless time helping me see the light at the end of the tunnel, and to the remainder of my committee members for your support and mentoring, thank you Dr. Nina Jablonski, Dr. David Puts, Dr. Stephen Schaeffer, and Dr. Ross Hardison.

CHAPTER 1: INTRODUCTION

This thesis provides some insight into the factors that may be involved in how humans perceive genomic ancestry in faces. Understanding the relationship between genotypes and phenotypes, such as facial components and perception abilities is useful for the field of genetics, as well as various other interdisciplinary areas. From studies involving genomics and perception, we might learn more about how humans have evolved in terms of perceptual abilities, as well as how humans behave socially. This study introduces some claims regarding the underlining determinants of various social structures that exist like discrimination. Additionally, this information is useful when considering the evolution of human interactions. How people rate genomic ancestry may show some correlations with the history of mate choice. Perhaps an innate ability to see faces genomically has acted upon humans since our beginning. For example, we know that humans' mate choice is dependent on factors such as, symmetry, health cues, and masculinity and femininity in faces. I posit that there may be a relationship between the characteristics that are acting on mate choice, and the cognitive ability to rate genomic ancestry in faces. This is a new interdisciplinary approach that combines cultural anthropology, sociobiology, genetics, and cognitive science to contribute to understanding our evolution collectively, as a species. Although this work is simply a screenshot of this time, it's useful to compare genomic estimates among populations both in the United States, as well as around the world. Evolution has taught us that we are constantly changing. For instance, the United States has a growing multi-cultural population. This population shift will cause people to see faces of many different genomic ancestry combinations. Thus, a study like this could be replicated at future time periods (e.g. 1 year from now, 5 years from now, 10 years from now etc.) to investigate evolutionary characteristics of human populations' perception abilities.

Significance of Faces

Faces are used in many different ways and are very important in everyday interactions. We use hundreds of muscles in our face to express emotion. The face is key in recognition of family, friends, significant others, celebrities on television, athletes, politicians etc. There is a wide spectrum of facial perceptive abilities. For example, the condition prosopagnosia or "face blindness" is characterized by a limited ability to recognize faces. Some affected individuals

cannot recognize even the faces of close friends and family members [Ramon, 2009]. Although clinically significant, prosopagnosia is rare; it has been suggested that as many as 2% of people suffer from some mild form of prosopagnosia [Kennerknecht, 2006]. Further, the opposite end of the spectrum must be considered. “Super-recognizers” are as far above average at face recognition as developmental prosopagnosics are below average. Russell and collaborators (2009) showed that persons who claimed to have significantly better than ordinary face recognition ability were far above average regarding two separate tests of facial recognition-one involving recognizing faces of people before they were famous, and another involving learning unfamiliar faces. Indeed, there is substantial variability in how persons along the “normal” spectrum perceive faces, which is the subject of this investigation.

In some cultures, being able to recognize facial features may be more advantageous than in other cultures. In some populations, specific characteristics are understood by society as desirable and thus selected more. For instance, a significant difference in preference for symmetry has been shown when assessing attractiveness between two groups: Europeans from the UK and the Hadza people from Tanzania [Little, 2007]. Both the Europeans from the UK and the Hadza people in Tanzania were shown five pairs of photos (Europeans observed European faces and the Hadza observed Hadza faces) and asked which face was more attractive. The Hadza people showed a stronger preference for symmetry and it has been suggested that the reason for this is due to their higher mortality rates. Symmetry is correlated with stronger immune systems and this allows for a higher survival rate. Overall, this study showed that a relationship exists between face recognition and cultural decisions, thus contributing to sexual selection and evolution.

It is clear that the Little (2007) study shows that, facial perceptive abilities are culturally modified. An additional factor to consider in face recognition is “the other-race effect” (ORE), which is the diminished ability of persons to recognize faces of members of “races” compared to people within their own “race” [Malpass, 1969]. There is a significant body of work supporting the contention that the ORE is largely environmentally mediated, stemming from limited exposure to “other-race” faces at critically important developmental periods in childhood [De Heering, 2009].

One of the outcomes of the ORE is that it may mediate social treatment, such as discrimination in health care and in the criminal justice system [Steffensmeier, 2001]. This study is important because, perhaps there is a path involving demographic factors, face recognition, and genomic ancestry that results in discrimination. Faces are a major component of non-verbal communication. Variability exists in both faces and in perception abilities and thus, due to this constant interaction over time, facial feature and perceptive traits may have co-evolved. It is important to study both factors collectively because they can be useful in genetics, criminology, health care, and psychology.

The Other-Race Effect (ORE)

In 1969, Roy Malpass (University of Illinois) and Jerome Kravitz (Howard University) began investigating the phrase, "…they all look alike". They showed African American and European American students at both universities a series of photos that included, in random order, 40 Black and 40 White students' faces. After the first set of photos was shown, the students were exposed to a second set of photos that included both previously seen photos and some new photos. The images were shown in succession (seconds apart) and the students were then asked to indicate whether they had seen the face in the first set or not. This study provided the first published evidence that people are better at recognizing "in-group" faces compared to "out-group" faces [Malpass & Kravitz, 1969]. This paper also suggested that the results were consistent with the "hypothesis of differential experiences." That is, experience with individuals of another "race" is related to increased recognition ability of the other "race." This research was key to developing the field of face perception and the phenomenon known as, the "other-race effect" (ORE), also known as "own-race bias" and "cross-race effect." In this thesis, I will use the term "other-race effect" or ORE, which is the observation that people are generally more accurate at recognizing faces from the population group they belong to (e.g. European American, African American, etc.) compared with faces from other "races" or population groups.

According to the ORE literature, experience and exposure to other racial/ethnic groups are two key factors that mediate the degree to which a person is able to recognize and remember faces of one's own race compared to faces of other races. There are two hypotheses on which the

ORE is based: social-cognitive and perceptual expertise [Sporer, 2001]. The social cognitive model focuses on the different ways people process information as a function of categorizing others as in-group or out-group members; the perceptual expertise model is based on the idea that the ability to extract information from an environment improves with experience [Hehman, 2009]. Meissner (2001) did a meta-analysis of all of the own-race bias studies and summarized that this effect is present in African Americans, European Americans, and populations all around the world. Studies have been done in France and South Africa, as well [Sangrigoli et al 2004; Horry et al 2010]. Sangrigoli conducted a study in England using three groups: Korean adults who were adopted young by European families, Europeans who resided in France since birth, and Koreans who were born and raised in Korea and had lived in Europe for less than 11 years. The participants saw one face, and then the next screen displayed two faces and they had to choose the face that matched the original target. The researchers found that the adopted Koreans recognized European faces better than they did Korean faces. These findings support the contention that this phenomenon is not only seen in the United States, but all over the world, and that experience and exposure are key. Horry and colleagues (2010) sampled both White and Black South Africans at the University of Cape Town and presented them with a series of photos, each face with a different background location. After seeing all of the photos, the participants had to say whether the face was old or new and if it were old, they answered whether they "remembered, knew, or guessed (RKG)" their answer. They found that all participants "remembered" more faces of their own-ethnicity than faces of other ethnicities, but that the effect was larger in the White participants, meaning they were worse recognizing faces of other ethnicities than the Black participants.

While the ORE exists in many different cultures, it is important to consider whether this effect can be modified, as well. Children who were adopted into different "race" homes show that the effect can be reversed [Sangrigoli, 2005], but only if that exposure is long enough and occurs during the neurological development of the face recognition system [de Heering, 2009]. Face processing is mature at four years of age [de Heering 2007]. In adults, Lebrecht and colleagues (2009) showed that the ORE can be diminished with training and exposure to more faces. Participants were given memory, categorization, and individualization tasks with training

phases (pre-testing), testing phases, and post-training phases. The researchers measured both the ORE and the participants' implicit racial bias before and after and showed that both were improved after having the training phases. Implicit biases are discriminatory biases based on implicit attitudes or implicit stereotypes and can produce behavior that diverges from a person's endorsed beliefs or principles [Greenwald, 2006]. Greenwald also speculated that perhaps improving ones ORE will then improve their implicit bias. This speculation also supports this thesis in that understanding all aspects of face perception can elucidate implicit biases and even discrimination that results from such biases. Specifically, in this thesis, I will explore the perceptual expertise model. I propose that experience and exposure are two key factors that contribute to the effect of the ORE and will also contribute to one's ability to rate genomic ancestry in human faces.

Categorical Perception

While experience and exposure are two key components involved in the ability of a person to recognize faces of his or her own "race", there is another area of study in perception that provides a theoretical framework for the ORE. Categorical perception has been studied in both language and faces and occurs when discriminations between stimuli within one category are more difficult than discriminations between stimuli that span the boundary between categories [Levin & Beale 2001]. Unlike the ORE, this process has more to do with the cognitive perceptual system; the ORE is more culturally involved. Calder and colleagues (1996) examined categorical perception of morphed facial expressions. They showed participants three series of faces that were morphed along an emotion continuum that went from happiness → sadness, sadness → anger, and anger → fear and they found that the faces that were in the middle of the continuum were identified as either of the two emotions but the faces that were on either side of the categorical boundaries were identified as the correct emotion. Levin & Angelone (2001) did a similar study (using the same study design) but they used the continuum Black → Black, Black → White, and White→ White. They showed that categorical perception on between-"race" continua (Black → White) was stronger than within-"race" continua (Black → Black or White → White). I hypothesize that there will be some categorical perception at play in both tasks of rating genomic ancestry and categorizing faces as Black or White.

Faces and Genomic Ancestry

Klimentidis and Shriver (2009) showed that people are better than simulated estimates when rating genomic ancestry components in faces admixed with Native American and European ancestry. For their study, and the one presented here to be possible, the relationship between ancestry and faces had to be examined. Over the past few decades, the field of genetics has grown and now lies at the forefront of science. With the advancement of DNA technology, scientists are able to investigate how genotypes and phenotypes are connected.

Population structure can be defined as variability in the distribution of alleles across a population or alternatively, a correlation among alleles at unlinked loci and can result from a number of different causes, for example, non-random mating because of some reproductive barrier. Continuous gene flow may also cause population structure because populations moving from one region to another and reproducing will allow for new genetic combinations to be introduced, which result in allelic associations present from both parental population groups. The more this process occurs, over time, the more population structure present in the population. Admixture stratification is population structure in a recently mixed population and varies in extent among various population groups for a number of reasons. Traits that vary between parental populations are correlated in admixed populations as the result of admixture stratification [Pfaff, 2001]. Consider two traits that vary on average between Europeans and West Africans (e.g. nose width and skin color). Even though the genes that determine population differences in nose width and skin color are not the same, and perhaps not even on the same chromosomes, admixture stratification results in trait value correlations in admixed populations.

In face perception, population stratification is important to consider when analyzing how genomic ancestry is detected due to relationships among associated traits and ancestry. I speculate that observing facial features and skin pigmentation can yield accurate estimation of genomic ancestry. Only if a population is structured will associations between unlinked loci be detected. Population structure remains in a population through assortative mating and gene flow. Because both gene flow and assortative mating have occurred in African Americans and

Brazilians, there are significant associations between unlinked markers, and these associations potentially result in correlations between ancestry and facial traits [Parra, 2004].

European and West African Genomic Ancestry: Admixture

Charles Darwin, in his book "Descent of Man, and Selection in Relation to Sex" (1871), was the first to postulate that humans have a single common ancestor. Through much advancement in multiple scientific disciplines, this claim has been investigated using genetics of mitochondrial DNA [Cann, Stoneking & Wilson, 1987], as well as, anthropological study of fossils [Lieberman, 2002; White, 2003]. Modern Humans originated in Africa between 90-130 kya [Jobling 2004], and through generations of the forces of evolution, have formed many populations groups. The term "population" is used by anthropologists to replace the word "race". Joseph Graves (2004) defines "race" as a group of animals or plants that 1. Can have its own distinct genetic lineage, meaning that it evolved in enough isolation that it never reproduced with individuals outside its borders or, 2. The genetic distance between one population and another has to be significantly greater than the genetic variability that exists within the populations themselves. The word "race" has many meanings associated with it. It is a socially constructed term that was originally used to describe the various groups of people according to differences in physical appearance e.g. skin color, hair, facial features etc. Although the word "race" was originally used to describe categories of persons that looked different from one another, scientists believed that these categories also correlated with biological differences among groups, as well. Because "race" is a combination of both cultural and biological features of metapopulations, it is not an accurate term when describing groups.

"The definition of a 'population' is not simple, but reflects a combination of geographical proximity, a common language and shared ethnicity, culture, and religion" [Jobling 2004]. Although there are many populations that exist today, I will focus on samples of persons mixed with West African and European ancestry.

When an individual or population has ancestry from more than one parental group, that individual or population is considered "admixed". The term "admixture" is defined as the formation of a hybrid population through the mixing of two ancestral populations [Jobling

2004]. West Africans were first brought to America in 1619 and it is estimated that intermixture among Europeans and Africans in North America began around 1675 [Glass 1953]. This estimate suggests approximately 14 generations of admixture to the present time, assuming approximately 25 years per generation. African Americans are admixed showing mostly West African and European, but also some Native American ancestry [Parra, 1998].

Proportions of genetic admixture are measured by comparing allele frequencies among the parental populations, and the admixed populations. First explored by Neel [1973], markers or alleles that showed population differences were called “private”. Later that same year Reed [1973] used “ideal”. As time progressed, the terminology changed: “unique alleles” [Chakraborty 1991] or, “population specific alleles” (PSAs) [Shriver 1997]. The most widely accepted term used today is Ancestry Informative Markers (AIMs) [Pfaff, 2001]. AIMs are used to determine degree of admixture in populations with combinations of two or more parental population ancestries. Unpublished data in the Shriver lab collected by Denise Liberton and Kerri Matthes has shown strong correlations between facial features and ancestry and thus make it feasible to study perception using genomics and structured populations, such as African Americans.

In addition to African Americans, Brazilians are also “admixed”. Brazil is a large country and includes many subpopulations. Brazilians have a complex history of mixture involving Europeans represented by the Portuguese, enslaved West Africans, and Native Americans [Parra, 2003]. Assortative mating and continuous gene flow are two processes that occur in admixed populations and result in allele associations, allowing for an investigation like mine to proceed. Because African Americans and Brazilians are recently admixed populations, associations among ancestry and some traits are still present.

Racial Passing

An important phenomenon to consider in this study is “passing”. There are many different types of passing including sex (men passing for women or vice versa), sexual orientation (homosexual passing for heterosexual or vice versa), or “race” (Black passing for White, or Mexican passing for Indian) (Doniger, 2004). “Passing” is not recognized by the researchers who

investigate this phenomenon. It is studied through personal testimonies of those people who have passed in their lives and chose to share their stories. Generally, an individual is "passing" when he or she is successful at pretending to be something or someone else. I reviewed racial passing in this section because this phenomenon could contribute to how people perceive genomic ancestry in human faces. Although this phenomenon has existed for centuries, unlike the ORE, racial passing has not been studied as extensively. Racial passing is defined as being capable, based off of physical appearance, of not being identified as belonging to the socially stigmatized "race" [Kennedy, 2001]. This act occurs in various populations ranging from Whites passing for Native American (advantageous for specific government aid) to Asians passing for White. In the United States, passing mostly refers to a Black or multiracial person passing for White. From 1861-1950, an estimated 600,000 persons were passing [Stuckert, 1958]. Although the civil rights movement helped put an end to institutionalized racism, it is speculated that passing still occurs and often results in better treatment. For example, my friend is a Black person who passes for White. She has had numerous lunches with her European American friends and receives better service (food comes faster, server more hospitable etc.) than when she has lunch with her African American friends.

For reasons that haven't been proven scientifically, African Americans, "Negroes", or "Coloreds" have most times been able to distinguish when an individual is "passing" [Crary, 2003]. Having low levels of West African ancestry many times goes undetectable by Whites. Ralph Ellison in his 1946 essay stated, "although the sociologists tell us that thousands of light skinned Negroes become White each year undetected, most Negroes can spot a paper thin 'White Negro' every time" [Ellison, 1953]. Perhaps, this ability results from the specific concept: "takes one to know one", a perspective that postulates only those who are a part of that culture (being Black) can know when a person is Black, but may be "passing" for White [Robinson, 1994]. Beyond the scope of this thesis lies an opportunity for someone to use my experimental approach to investigate racial passing by using genomic ancestry estimations. Comparing genomic estimates between different population groups would be one way to generate data that supports or refutes claims that Blacks are better than Whites at detecting persons passing for White.

"Estimating Genetic Ancestry Proportions from Faces"

Recently, Drs. Yann Klimentidis and Mark Shriver asked subjects to look at facial photographs and estimate the genomic ancestry of the faces. They showed that observers' answers depended on the self-defined "race" and ethnicity of the observer [Klimentidis, 2009]. This work was carried out in Albuquerque, New Mexico at both The University of New Mexico and at a local MVD (Motor Vehicle Division) and focused on variability of face perception across the Indigenous American and European genomic ancestry axis. The facial photos shown to observers were comprised of 14 college aged students who self-identified as Hispanic or Latino Americans with family origins in Europe or the Americas. The 14 were selected from a larger sample of 55 students enrolled at the Pennsylvania State University with family origins in Europe and the Americas. These individuals were genotyped using ancestry informative markers (AIMs) and the results of which were used to estimate genomic ancestry [Halder, 2009]. Criteria for selection were low levels of African and East Asian genomic ancestry and varied levels of Native American and European ancestry proportions. Their photos were then used as stimulus for 241 observers. Students enrolled in introductory Biology and Anthropology at the University of New Mexico were recruited [n=134], as well as patrons at a location in New Mexico [n=107]. All participants were asked to estimate the proportion of Native American admixture they thought each person had.

Klimentidis and Shriver (2009) collected demographic variables, namely, age, sex, income, ethnicity, self-estimated genomic proportions, and they tested for relationships between the demographic variables and participants' accuracy at judging genomic ancestry. The formula for Euclidean Distance was used to calculate an error score, which explains an observers' facility at rating ancestry. The Euclidean Distance Score (EDS) formula summarizes the difference between the observer estimates and the genomic estimates of ancestry. Briefly, they showed that when compared to random simulations of ancestry estimations, observers on average had lower error scores (better ancestry estimates). Additionally, there were no significant correlations found regarding the observers' facility at estimating ancestry proportions and their sex or where they lived most of their lives.

However, age was a demographic variable that showed a significant correlation with lower error scores in the observer's ratings. Although, the authors of this paper hinted at, this may have been due to the fact that the college students were presented photos of people who were closer to their age or that they were in college and perhaps being stimulated, intellectually, more than those individuals at the MVD. The mean ages were 21.6 for the University and 36.8 for the MVD. Any person who was 18 or older was eligible to participate, thus, there were no specific age selections made for this study. Klimentidis and Shriver (2009) also showed that there is variability in self-identified "race" or ethnicity and estimated ancestry proportions. Persons self-identifying as European, or having higher self-estimated proportions of European ancestry also had lower error scores, and they, on average, overestimated Native American admixture proportions of the stimulus photos. Those who self identified as Native Americans, on average, underestimated Native American admixture proportions. Both groups were more likely to exclude a stimulus face, than include it in their respective population group. This trend supports the contact/differential experience hypothesis for facial form, which posits that an increase in facial recognition among persons in one's own group is due to more contact and experience with that group [Bothwell, 1989].

Conversely, Native Americans had higher error scores at rating high European admixed faces; perhaps they were more isolated and had not had the opportunity to become familiar with these faces. The relationship between genetic and socially based identification measures is confounded with historical, cultural, social and possibly phenotypic factors [Sinha, 2006]. Klimentidis and Shriver (2009) explored the degree of concordance between physical appearance and genomic measures of "race" and ethnicity in admixed populations with Native American and European ancestry. In summary, they suggested that this relationship was far from perfect, complex, and context specific. In addition, they concluded that their findings warrant this type of research on other admixed populations. Thus, in this thesis, I will examine whether differences in the ability to recognize and assign faces of admixed individuals from populations of African Americans and Brazilians depends on demographic experiences.

CHAPTER 2: MATERIALS/METHODS

Keywords:

Euclidean Distance Score (EDS) -The squared difference between the measured genomic ancestry and the rater's estimate of ancestry for all the faces in the panel. The lower this number, the closer the rater's estimates were to the calculated genomic ancestry.
Raters, Participants, Subjects (used synonymously)- The people in the study (n=100) who completed the demographic survey, then after viewing the stimulus faces, categorized one panel, and estimated ancestry in the other panel.
Stimulus faces- The two panels of 42 faces comprised of African American, Brazilian, and European faces (spanning a West African Ancestry spectrum of 0-100%) presented as printed photographs and observed by participants.

Design

This study examined how a set of subjects rated the genomic ancestry of human faces. I hypothesized that the factors that affect this rating process are dictated by the amounts of experience and exposure one has with faces of diverse "racial"/ethnic backgrounds. The study done by Klimentidis and Shriver (2009) involved the use of a short questionnaire, and 14 2D photos of persons from admixed European and Native American ancestry, to determine the degree of concordance between observer and genetic estimates. This thesis extends Klimentidis and Shriver's work by using:

1. a much more extensive demographic survey
2. printed 3D photos (n=84: two panels of 42)
3. faces admixed with West African and European genomic ancestry

The survey has been extended to include 41 questions to gather more information. Also, different from 2 dimensional photos, we used a 3 dimensional camera to capture components of the face and printed out those photos for this study. Finally, our populations were admixed with West African and European ancestry and not Native American and European ancestry. In addition to reporting on the relationship among genomic and rater estimates, I tested whether perception of genomic ancestry, measured by Euclidean Distance Score (EDS), is mediated by

the perceiver, or the face that is being perceived. It is possible that this accuracy is mediated by both the perceiver and the face being perceived, or neither of the two. Perhaps certain background characteristics of an individual may make him or her more accurate in rating genomic ancestry. For instance, persons who are bi- or multi-racial may be better at rating faces that fall in the middle of an ancestry spectrum. Conversely, a biological or genetic predisposition may be responsible for this apparent skill, or lack thereof. For the scope of this thesis, I focused on properties of the participants. Perhaps there are facial features that make the estimating task easier or more difficult.

This thesis tests whether demographic factors could predict the Euclidean Distance Score (EDS), or accuracy of ancestry estimation. I analyzed a total of 20 factors collected from a survey, and evaluated whether these factors are related to variation in EDS scores. Demographic measures such as education, socioeconomic status (SES), location of upbringing, and racial/ethnic breakdown of friends and schools are factors that I speculate will help capture important aspects of an observer's experience with faces. It has been shown that experience and exposure to faces enhances perception abilities at the categorizing level and memorizing level (Have you seen this face before? Is this face old or new?) [Meissner, 2001]. Furthermore, we know that humans perform better than random simulations at estimating genomic ancestry with photos [Klimentidis, 2009]. Thus, I hypothesize a connection between demographic variables and genomic perception.

I used a demographic survey to determine what variables regarding experience and exposure might or might not be responsible for ancestry estimation accuracy. I presented two panels of printed 3D photos and for one panel, asked each participant to categorize the photos as "Black" or "White" and for the other panel, asked participants to estimate from 0-100 how much African ancestry they thought each face had. Similarly, I focused on demographics specific to experience and exposure to other racial/ethnic groups. Equally important was what types of trends were observed in genomic ancestry estimation. Although all of the factors I tested were not significant, the data suggest that they were close to significance. Finally, using the facial categorization data, I analyzed where along the ancestry spectrum were the averages that corresponded to a face categorization of "Black" or "White". I hypothesized that participants'

EDS will be lower than the simulated ancestry estimations. I also hypothesized that there would be a relationship between EDS and the demographic factors. Lastly, I hypothesized that, due to a full range of genomic ancestry being used in this stimulus set; there will be no trends in over and underestimating genomic ancestry. Up until now, many researchers studied the two key components of facial recognition (experience and exposure) and used only memorizing and categorizing tasks; I would like to widen this area of research by introducing a genomic rating component.

Participants

Participants in this study rated the stimulus photos. I sampled 100 participants on the campuses of Morehouse College (Atlanta, GA) *n=53*, the Pennsylvania State University (University Park, PA) *n=10*, and Spelman College (Atlanta, GA) *n=37* (see Table 1). Although only three sample locations were used, the birthplaces of the raters spanned the United States, Barbados, England, Germany, Ghana, Iraq, South Korea, Spain, and the Sudan (See Table 1). Institutional Review Board (IRB) protocols were approved at each institution and students, staff, and faculty on each campus who were over the age of 18 and had proficiency in English reading were eligible to participate. Three undergraduate students at Morehouse College (Justin Whitt, Marquette Moore, and Jeron Rowland) and I recruited participants for two days from one of the dorms on Morehouse College's campus in May 2010. In November, 2010, in the lab of Dr. Mark D. Shriver, he and I recruited raters on the campus of Penn State via email and class announcements. Another trip was made to Atlanta in March 2011 and the same three Morehouse College undergraduates and I sampled students, staff, and faculty, following a lecture I gave in a Health and Economics class, at Spelman College. During that trip we, once again, sampled at Morehouse College in the Psychology lab of Dr. Sinead Younge. We recruited mostly fliers, lectures, class announcements and emails. There was no compensation for participating in the study.

Table 1. Rater's Racial/Ethnic Self Identification

	Total	Mean Age	Black/ African/African American	Arab	Asian/ Pacific Islander	Hispanic	Multiracial	White/ European/ European American	Other
Female	**39**	**26.1**	**25**	**2**	**3**	**1**	**5**	**2**	**1**
Male	**61**	**21**	**53**	**0**	**1**	**1**	**4**	**2**	**1**

Demographic Survey

Considering the demographic factors that might inform on experience and exposure, and consulting with Dr. Mark Shriver, Dr. David Puts, and Dr. Rob Burris (Penn State Anthropology), Dr. Bryant Marks, Dr. David Wall Rice and Dr. Sinead Younge (Morehouse Psychology), I created a demographic survey. With the survey, I hoped to determine what demographic factors were responsible for accuracy in ancestry estimation. It is comprised of a total of 41 questions that capture information regarding background, upbringing, and environment: *Which of the following best describes the area in which you currently live? <Rural; Suburban; Urban>. Before the age of 10, how many friends did you have who were European American (White)? <None (0%); Almost none (less than 25%); About half (about 50%); Mostly (more than 75%); All (100%)>. What percentage of your elementary, middle, and high school was European American (White)? <None (0%); Almost none (less than 25%); About half (about 50%); Mostly (more than 75%); All (100%)>. What is the racial/ethnic breakdown of your favorite television show? <All Black; mostly Black; mostly White; all White; racially diverse; other_____>* (See Table 2).

Table 2. Sample demographic statistics presented as counts (n) and average EDS

Area you currently live?		
rural	**suburban**	**urban**
9 (3.84)	46 (3.6)	44 (3.37)

European American friends before 10?				
none (0%)	**almost none (25%)**	**about half (50%)**	**mostly (75%)**	**all (100%)**
19 (3.92)	36 (3.3)	22 (3.59)	20 (3.45)	3 (3.6)

%age of Elementary, Middle, High School European American?					
	none (0%)	**almost none (25%)**	**about half (50%)**	**mostly (75%)**	**all (100%)**
elementary	16 (3.72)	25 (3.51)	29 (3.51)	26 (3.36)	4 (3.88)
middle	13 (3.29)	25 (3.49)	32 (3.79)	27 (3.26)	3 (4.16)
high	11 (3.77)	23 (3.40)	35 (3.56)	28 (3.45)	2 (3.62)

Racial/Ethnic breakdown of favorite TV show?				
all Black	**mostly Black**	**mostly White**	**all White**	**racially diverse**
9 (3.36)	17 (3.42)	35 (3.37)	6 (3.93)	27 (3.75)

Highest level of education?			
high school diploma	**some college**	**bachelor's degree**	**grad school/degree**
32 (3.6)	37 (3.43)	13 (3.43)	15 (3.64)

Stimulus Faces

Stimulus faces used in this study were selected from a larger dataset collected under the project titled "The Genetics of Human Pigmentation, Ancestry and Facial Features" (GHPAFF), with principal investigator Dr. Mark D. Shriver. Lab members Ellen Quillen and Laurel Pearson from Penn State sampled in Brazil. In addition, colleagues in other Universities at the international sites assisted with the collection efforts, in particular Brian McEvoy, Sandra Beleza, and Rinaldo Perieria in Brasilia. Three-dimensional images (3dMD, Atlanta, GA), finger stick blood samples (Whatman, Florham park, NJ), buccal cell swabs, as well as demographic and phenotypic information in the form of surveys were collected. The buccal cells were used to obtain biogeographical ancestry using ancestry informative markers (AIMs) [Halder, 2009].

The stimulus faces for this study were selected using three criteria: an age between 18 and 35 to help control for any morphological changes with age; limited facial hair so as not to distract from facial features; and less than a combined 15 percent of measured admixture of East Asian

and Native American ancestry [Matthes Master's Thesis, 2008]. In total, I chose 84 faces from the database and divided them into two panels of 42 photos with equal amounts of ancestry distributed in each panel (see Figure 1). In both panels, there were 24 stimulus faces with 0-40% African ancestry and 18 stimulus faces with 41-100% African ancestry. The distribution was not uniform for the panels because there were not enough faces with varying amounts of measured genomic ancestry. The sample locations for the stimulus chosen from GHPAFF for this study include State College, Pennsylvania; Williamsport, Pennsylvania; Porto, Portugal; Rome, Italy; Warsaw, Poland; and Brasilia, Brazil (see Table 3).

Table 3. Stimulus Faces sample characteristics

		Total	African American	Brazilian	European
Panel 1	Female	n=26	9	16	1
	Male	n=16	3	8	5
Panel 2	Female	n=31	12	16	3
	Male	n=11	4	4	3

I used 3dMD patient software 4.1, Adobe Photoshop Cs4, and Microsoft PowerPoint 2007 for the editing, assembling, and printing of each stimulus face. Dr. Peter Claes applied a quasi-landmark based approach to orient and remap the stimulus subject faces providing me with a standardized face trimmed along the hairline in front of the ears and just below the chin (see Figure 1 below). Each face in the stimulus set was saved as a serial 2D image (five separate angles: right profile, right oblique, front, left oblique, left profile). As shown in Figure 1, I used a 50% grey background color, to alleviate the contrast that a 100% Black or 100% White color might cause to the stimulus faces in the photos, as suggested by Dr. Rob Burris. I also used Adobe Photoshop in Creative Suite 4 to edit each individual photo file to the serial composite. For each panel, the serial composites were assembled using PowerPoint (see figure 1). After editing and assembling was complete for each of the two groups of 42 photos, I randomized their order using the random function in Microsoft Excel 2007 and gave each stimulus face serial composite an order number.

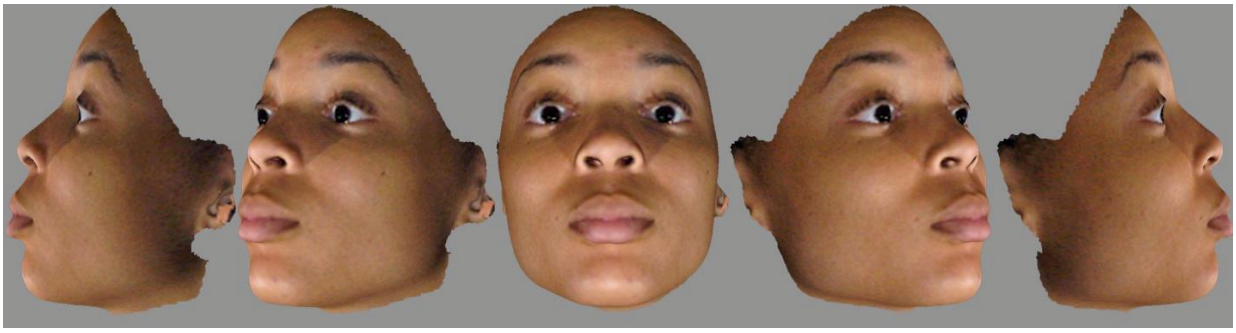


Figure 1. Serial composite of stimulus face.
Remapped to quasi-landmarks and edited in 3dMD patient software 4.1, Adobe Photoshop Cs4 and Microsoft PowerPoint 2007.

Procedure

The raters who read and agreed to the IRB Consent forms were included in this study. Following the consenting process, they were asked to complete the demographic survey and then to complete two tasks. Although the tasks were randomized per rater, one task was to categorize each stimulus face in panel 1 or panel 2 as "Black" or "White". The other task was to estimate the proportion of "African" genomic ancestry they thought each stimulus face had on a scale from 0-100 represented as percentages; each participant estimated one panel of stimulus faces and categorized the other panel of stimulus faces, so that no participant saw the same face more than once. Using the Halder et. al. (2008) methods, we specifically summarize amounts of West African ancestry by using West African allele frequencies as one of the parental populations, but to avoid having to give explanations that may affect the genomic estimates given by raters, I simplified the descriptor from "West African" to "African". There were eight different combinations of how the rater would complete the study according to task (categorize or estimate), panel (see panel 1 first or panel 2 first), and photo order (ascending or descending) to control for any order affects that the stimulus faces may cause in ancestry estimation (see Table 4). For example, the first participant would estimate panel 1 in ascending order first, and categorize panel 2 in descending order last. Table 4 shows the eight combinations I used. On the 9th rater, the order started over.

For data entry, we created a text filter in MicrosoftExcel which standardized all the answers for statistical analysis. This filter also reduced data entry time by allowing us to select the answer from the survey and move to the next, allowing for more time to evaluate for data entry errors.

There were 53 participants recruited during the first sampling, 10 for the second, and 37 for the last sampling. Although no names were recorded, we stored all data in MicrosoftExcel files on password protected computers to keep the identity of the participants secure [See Appendix A for sampling details].

Table 4. Eight combinations designated to allow randomization of rating and categorizations tasks.

Estimating and Categorizing Tasks		
Rater 1	P1, estimate, ascending	P2, categorize, descending
Rater 2	P1, estimate, descending	P2, categorize, ascending
Rater 3	P1, categorize, ascending	P2, estimate, descending
Rater 4	P1, categorize, descending	P2, estimate, ascending
Rater 5	P2, estimate, ascending	P1, categorize, descending
Rater 6	P2, estimate, descending	P1, categorize, ascending
Rater 7	P2, categorize, ascending	P1, estimate, descending
Rater 8	P2, categorize, descending	P1, estimate, ascending

Statistical Analyses

I used several statistical software packages to analyze the data, Matlab, SPSS 17, and Microsoft Excel 2007. As stated in the introduction, this work was aimed at understanding not only the rater who perceives genomic ancestry, but also the face that is being perceived. The statistical analyses reflect both the rater, and the stimulus faces. Regarding the types of analyses conducted for this study, I employed two different methods. I used both individual and group analyses to assess the contribution of each demographic factor to the accuracy of Euclidean Distance Score. That is, because most of the demographic factors examined by the demographic survey had multiple levels, I wanted to look at both the individual and group effects of these factors in determining EDS. The remainder of the thesis will be organized according to nine main areas of focus: 1. Accuracy in estimating genomic ancestry (overall distribution of EDS), 2. Accuracy in estimating genomic ancestry (trends present in over and underestimating), 3. EDS simulations 4. Aggregating panel 1 and panel 2 data, 5. Average ancestry assigned to “Black” and “White,” 6. EDS per stimulus face, 7. EDS per population group, 8. Sex differences in ratings, and 9. Demographic predictors of EDS.

Accuracy in Estimating Genomic Ancestry: EDS

I plotted the genomic ancestry estimates for each stimulus face against the average raters' estimates in order to provide an overview of the measured vs. rater ancestry proportions. To determine each rater's accuracy in estimating genomic ancestry, the Euclidean Distance formula was used to calculate the Euclidean Distance Score (EDS) in Microsoft Excel 2007. This was done using the following formula: $\frac{\sqrt{\sum_{i=1}^{n}(GE-RE)^2}}{n}$

The EDS summarizes the magnitude of difference between two measures, in this case, the genomic estimates and the rater's estimates. GE is the genomic estimate, RE is the rater's estimate, and *n* is the number of faces each rater provided an estimate for in case raters did not provide a response for every stimulus face. I also used Excel to calculate descriptive statistics to observe the overall distribution of the EDS and summarize the data set.

Accuracy in estimating genomic ancestry: Over/Underestimation and βeta values

To investigate trends in the direction in which the participants estimated genomic ancestry can be investigated using two methods: (taking the difference and beta values). Both methods were used, but only the differences were used to analyze trends in over and underestimations. I subtracted the genomic estimate (GE) from the rater's estimate (RE) for each of the 42 stimulus faces, for each participant; positive values correspond to overestimations and negative values correspond to underestimations. The other method is using the regression line's slope (β) values for each rater to evaluate the magnitude of direction of estimate per rater, per stimulus face. Slopes greater than zero indicate more overestimations than underestimations, while slopes less than zero indicate more underestimations than overestimations.

EDS Simulations

To model random participants rating genomic ancestry, I ran 100,000 simulations with the statistical software MATLAB 2009 Student Version. In summary, I calculated what would be Euclidean Distance Scores for 100,000 random participants. To do this, I wrote computer code in Matlab that involved three main steps: first, import the 42 measured genomic estimates

from one of the panels, arbitrarily (panel 2), second, shuffle the 42 genomic estimates 100,000 times to create new datasets, and third, use those shuffled datasets to calculate the EDS 100,000 times (recall EDS formula above). Following the calculation of the simulated EDS, I also output a histogram of the simulations, a *t test* comparing means, as well as descriptive statistics (See Appendix C for MATLAB script). I used the shuffling approach so not to assume that the data fell along any particular distribution.

Aggregating Panel 1 and Panel 2 Data

I used *t tests* to determine if there was a statistically significant difference between the EDS for panel 1 and panel 2. Both panels were created to have similar ranges of stimulus faces that spanned the ancestry spectrum from 0-100% (range is 2% and 96%). Figure 2 below displays a histogram of each stimulus face's measured West African (WA) genomic ancestry proportions to show that both panels have the same ancestry distributions. In order for the data to be aggregated, panel 1 and panel 2 EDS means cannot differ statistically.

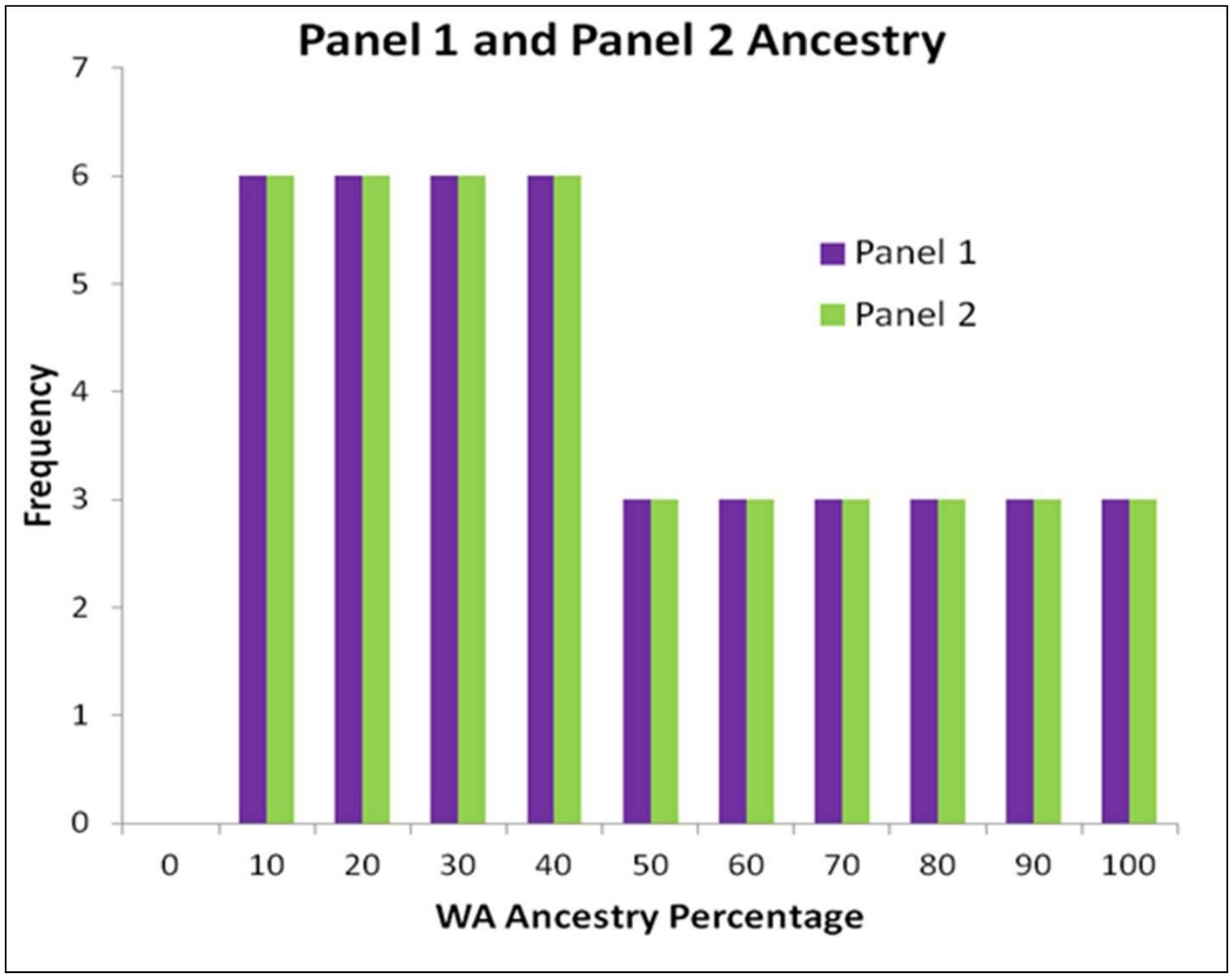


Figure 2. Panel 1 and Panel 2 Ancestry
Histogram of WA ancestry for stimulus faces in panel 1 and panel 2.

Average Ancestry Assigned to "Black" and "White"

I used logistic regression to determine the average ancestry percentile that corresponded to where, on average, participants categorized a face as "Black" or "White". After coding each response (Black-0; White-1), for the "Black" and "White" responses, the averages of the measured genomic ancestry percentages were calculated. For instance, if a rater had a frequency of 31 faces categorized as "Black" and 11 faces categorized as "White", I took the 31 genomic estimates for the subjects he labeled "Black" and averaged them together and that was established as the "Black" average genomic estimate for that participant. For the same participant, to get their average genomic estimate assigned to "White", I averaged the remaining 11 faces that were categorized as "White".

Euclidean Distance Score per Stimulus Face

To determine if the stimulus face was pivotal in rating genomic ancestry, I investigated if there were significant differences among EDS between stimulus faces. I pulled all of the genetic estimates (GE) and raters' estimates (RE) for each stimulus face and grouped them together. Thus, because each participant rated all faces in the panels, for all 84 targeted faces, there were 100 (one per rater and panel 1 and panel 2 combined) GE and RE that I used to calculate an EDS per stimulus face. Note that the formula for EDS was modified so that n= 50, for the two panels to represent the number of total observations per stimulus face, as compared to n=42 when calculating the participants' EDS. Then, I ran an *ANOVA* to compare means. For example, the results show that in panel 1, stimulus face #7 had an average EDS of 1.59 (corresponding to a very accurate ancestry rating) and stimulus face #34 had an EDS of 5.6 (corresponding to an inaccurate ancestry rating).

Euclidean Distance Scores per Population Group

ANOVA was used to determine if there was a difference in EDS among the three population groups in the stimulus set (African American, Brazilian, and European). As discussed in the introduction, these three populations (African American, Brazilian, and European) are structured differently, meaning their different population histories have resulted in different allele combinations responsible for facial features. For instance, linkage disequilibrium may be responsible for two particular traits being co-inherited in African Americans, but not in Brazilians. Although the ancestry proportions are similar, it is important to explore this relationship to further understand the contributions of both the perceiver, as well as what is being perceived.

Does Sex (male or female) of Rater and/or Stimulus Face Affect EDS?

I used both a general linear model and an *ANOVA* determine if there was a difference in EDS between men and women, men estimating men, women estimating women, men estimating women, and women estimating men. The sex of a face is very distinguishable and I wanted to explore whether this characteristic played a role in the rating process.

Demographic Predictors of EDS

From the demographic survey of 41 questions, I pulled 20 factors to determine which variables might be significantly related to ancestry accuracy (See list below). The survey was originally designed to detect differences among various different racial/ethnic groups. Because the majority of my sample was African American, the answers to some questions were the same for most of the participants and thus, I had to eliminate questions that resulted in no basis of comparison e.g. "*What 'racial'/ethnic group are you most attracted to? What 'racial'/ethnic group do you typically date?"* Although I asked questions pertaining to after the age of 10 regarding household numbers, as well as ages of people in the house, I later decided that I wanted to focus on an age range during the development of perception (ages birth-10). Thus, I chose to analyze only those data. The remaining questions that were not used in the analyses were excluded in an effort to reduce redundancy, and also because a lack of variability in the responses e.g. all raters fell into one category. Independent *t*-tests or *ANOVAs* were run for the 19 factors to examine the affects each predictor had at an independent level (See Appendix H).

1. Age
2. **Sex**
3. **People in household before ten**
4. Location before 10
5. **Current neighborhood**
6. **European friends before ten**
7. **Elementary school European-American population**
8. **Middle school European American population**
9. **High school European American population**
10. **Current European American friends**
11. **Highest education level**
12. **Mom highest education level**
13. **Dad highest education level**
14. **Parents income before ten**
15. **Self reported ability to recognize Black faces**
16. **Self reported ability to recognize White faces**
17. **Self reported ability to recognize actors on TV**
18. **Type of websites frequented**
19. **Hours of TV**
20. **Racial ethnic breakdown of favorite TV show**

To analyze the demographic factors collectively, I ran a multiple regression. In order to do this, the number of facts had to be reduced. For data reduction, I employed a factor analysis. To run the factor analysis, I needed the number of predictors to be reduced from 20 to about 12 (rule of thumb states there should be 10 observations per factor in factor analysis; for 100 participants, 10 factors is sufficient for the optimal power in the regression). However, in order to maintain as much of the original data as possible, I narrowed it down to 12 factors. Per the list above, the factors bolded in Black were individually measured and the factors bolded and underlined were merged. Factor analysis also requires interval variables. For example, age was split into two categories: raters 21 and under (n=71) and raters 22 and over (n=29). The factor "people in household before 10" was divided into two categories: less than four in household (n=60), and four and up (n=40). As explained above, the "current neighborhood" was divided into two categories: urban (n=45) and non urban (n=55). All of the factors and codes are listed in Appendix G. The factor analysis produced five principle components that were used as independent factors, along with "age" in multiple regression, and EDS as the dependent factor.

CHAPTER 3: RESULTS

Accuracy in estimating genomic ancestry: EDS

As shown in Figure 3, there is a strong relationship between the measured genomic estimates and the rater's average perceptions. This, along with the simulations, contributes to the proof that people have some skill in perceiving genomic ancestry in human faces. We can also infer the direction of over or underestimation from this plot. It appears that towards the middle of the ancestry spectrum, there is more overestimation than underestimation, based on how the points in the plot are above the trend line. I also measured a rater's accuracy of ancestry estimation by the Euclidean Distance Score (EDS): the lower this number (minimum was 2.13 and maximum was 7.75) the *more accurate* the ancestry estimation or, the better the rater was at estimating/perceiving genomic ancestry.

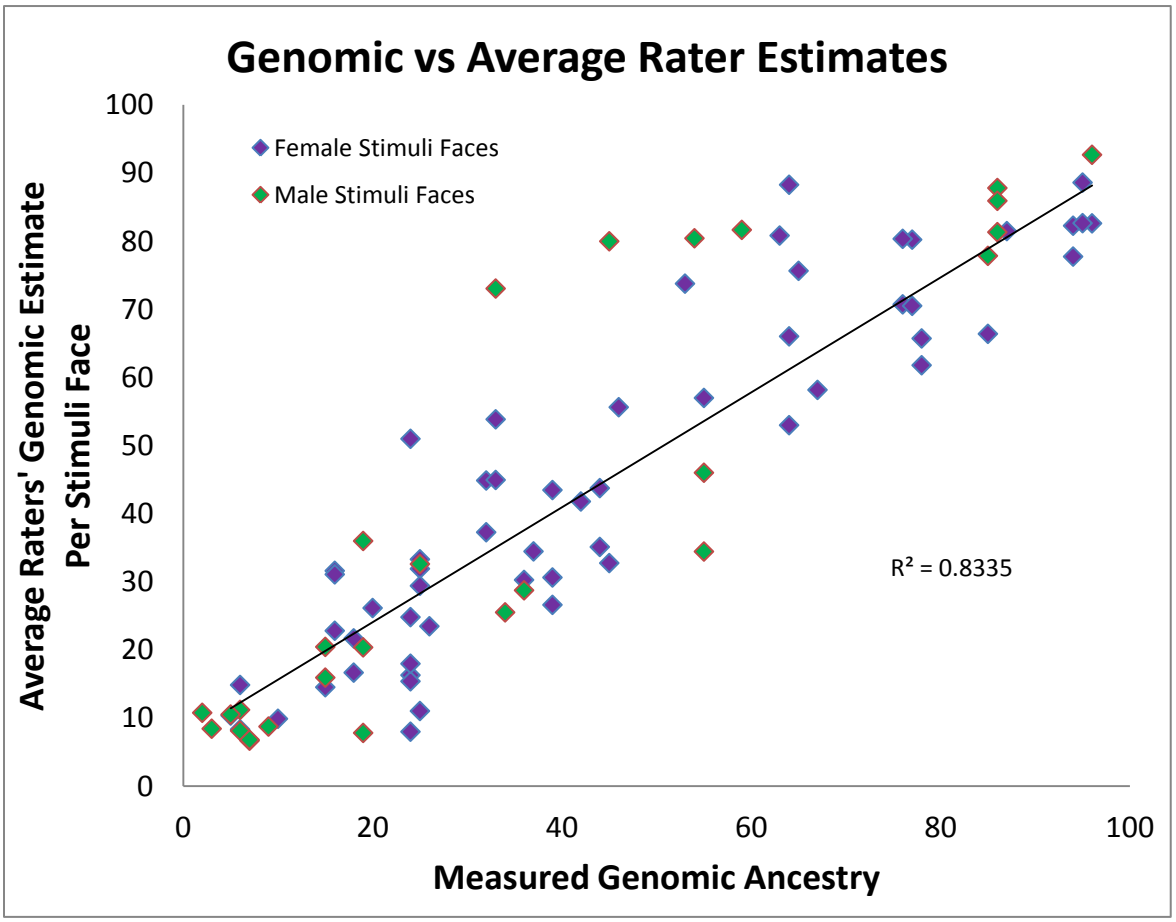


Figure 3. The perceptual estimates compared to the genomic estimates.

Table 5. Descriptive statistics (EDS) for the actual data (n=100) and the simulations for (n=100, 1,000, and 100,000)

Descriptive Statistics (EDS Distribution)				
	Actual	***100 Simulations***	***1,000 simulations***	***100,000 simulations***
Mean	3.52	6.04	6.14	6.15
Standard Deviation	0.91	0.5	0.48	0.49
Median	3.34	6.01	6.16	6.17
Range	5.62	2.95	2.92	3.89
Minimum	2.13	4.33	4.47	3.84
Maximum	7.75	7.28	7.39	7.73
Count	100	100	1,000	100,000

The null hypothesis states that there is no difference among the participants' EDS and the random simulations EDS. Raters were found to be better than chance at estimating genomic ancestry. I ran 100, 1,000, and 100,000 simulations and the mean EDS was 6.04, 6.14, and 6.15 respectively (See Figures 3 and 4 for the actual EDSs and the 100,000 simulations). I ran an ANOVA and found that the means for the actual, compared to all three simulations are significantly different (*p=0.0001)*. From this we can conclude that some skill, whether cognitive or cultural, is at play during the process of ancestry genomic estimation in human faces.

After running the two-sample *t*-test with and without the assumption of equal variances in panel 1 and panel 2, I found in both scenarios that there was no statistical difference between panels (*p=0.9) (See Appendix C for output)*. As no statistical difference in EDS between panels was found, the data from panel 1 and panel 2 were aggregated when analyzing the race threshold, as well as the information from the demographic survey (e.g. age, sex, SES, and education).

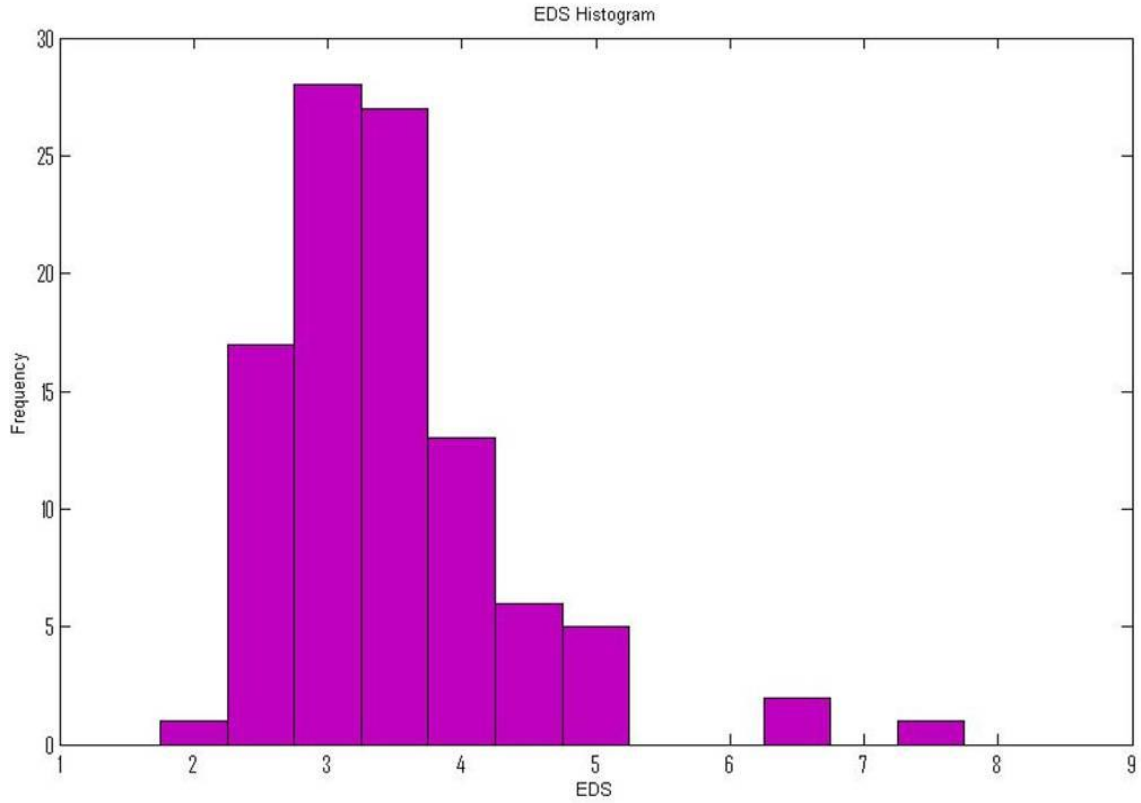


Figure 4. Actual EDS Distribution.
The distribution of Euclidean Distance Scores for all raters (n=100).

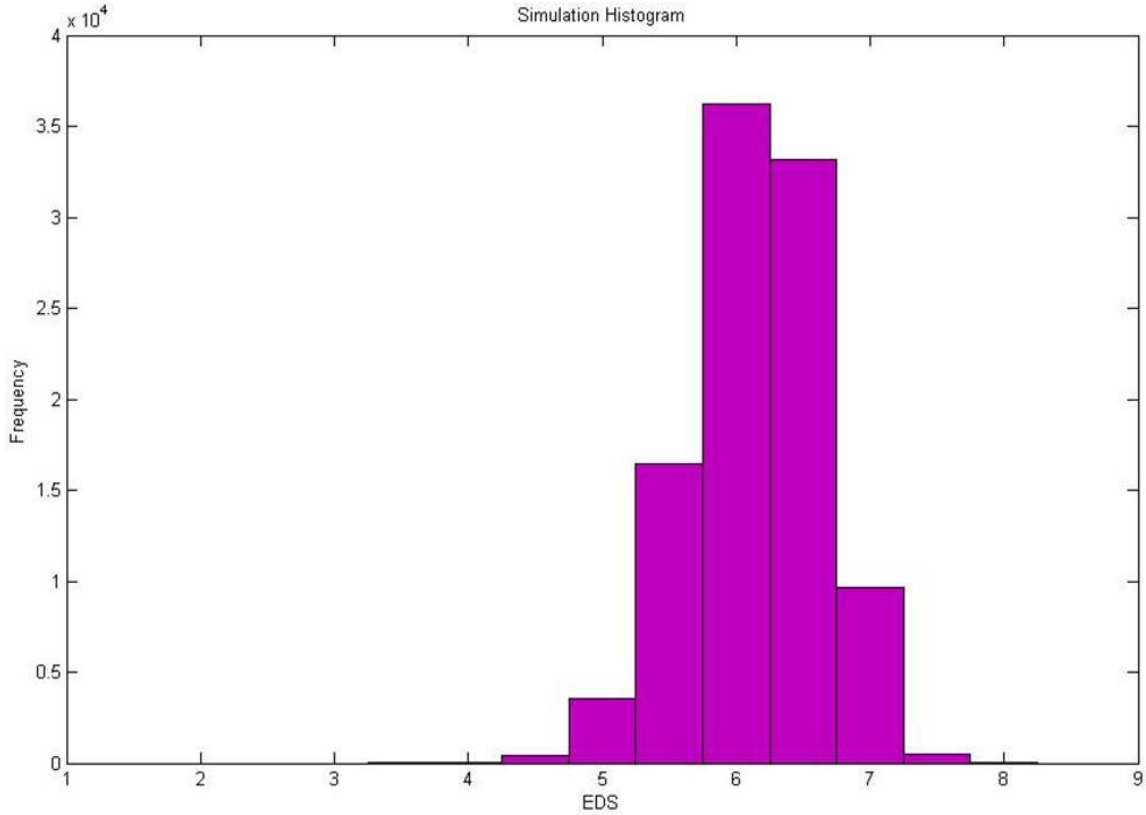


Figure 5. Simulation EDS Distribution.
The distribution of the simulations (n=100,000).

Accuracy in estimating genomic ancestry: Over/Underestimation Score

In addition to calculating the EDS to measure genomic ancestry estimation, there is an additional method to measure a rater's accuracy in ancestry estimation. I estimated the difference in the genetic estimate (GE) and rater's estimate (RE) to determine the direction of over or underestimation. Out of 4,101 observations, there were 1956 (~48%) overestimations and 2145 (~52%) underestimations. There were 275 (~46%) overestimations, and 320 (~54%) underestimations of European faces. Conversely, there were 1707 (~48%) overestimations, and 1825 (~52%) underestimations of non European faces. These numbers added together do not equal the expected 4200 observations because there were 99 perfect ratings. I ran an ANOVA comparing three categories (total, European and non European) of overestimations and underestimations and found no statistical significance (p = 0.9). Although no significant difference was found, all categories examined show that more faces were underestimated for West African Ancestry, than overestimated.

Average Ancestry Assigned to "Black" and "White"

After compiling all of the data from the categorizations that the raters completed, along with the measured genomic estimates of the stimulus faces categorized, I calculated the median West African Ancestry percentage for the "racial"/ethnic categorizations (Figure 5). The median West African ancestry percentage a face was categorized as "Black" was 53% and 25% was the median ancestry at which a face was categorized as "White". There was no single point in the spectrum where a face was categorized as "Black" or categorized as "White". The median ancestry was used to capture a more accurate rating. Using the average ancestry would have given a substantial amount of weight to extremes, for example, to those ratings that may have been a result of a participant not following instructions (e.g. rating a European face as having 100% WA ancestry, while thinking that their 100% designation signified European ancestry).

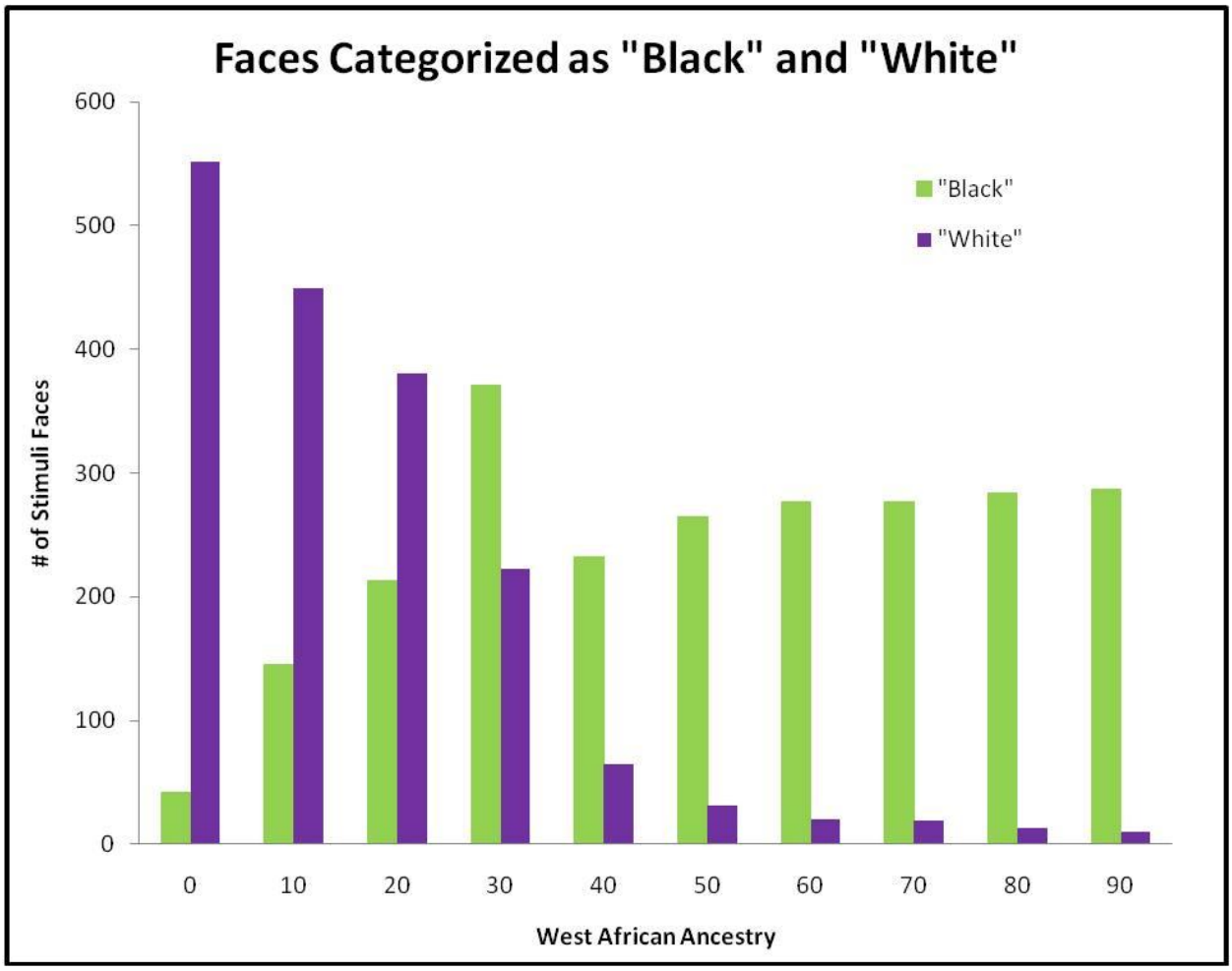


Figure 6. Histogram of faces categorized as "Black" and "White".
All of the categorizations plotted against frequency counts.

Properties of Stimulus Faces and Population Group

In this thesis, I tested hypotheses regarding properties of not only the rater, but also the face that is being rated. I examined whether the stimulus faces in panel 1 and panel 2 had significant differences in EDSs. The regression was used to test how well ancestry predicted variability in EDS. That is, considering faces of varying ancestry levels, how much does the ancestry estimation accuracy change. As I hypothesized, for both panel 1 and panel 2, a curvilinear relationship existed between genomically determined West African ancestry and EDS, but, it was only significant for panel 1 (*p=0.02)* (See Figure 6). This relationship was expected because it is easier at both ends of the ancestry spectrum for participants to rate a face that has higher or lower ancestry proportions. Moreover, a non significant result might allude to the proposition that more than skin color is being perceived. Some of the variation in EDS, as shown

in Figure 7, Panel 2, is explained by ancestry (skin color), but this non significant result supports the contention that the rater is using other features to estimate genomic ancestry.

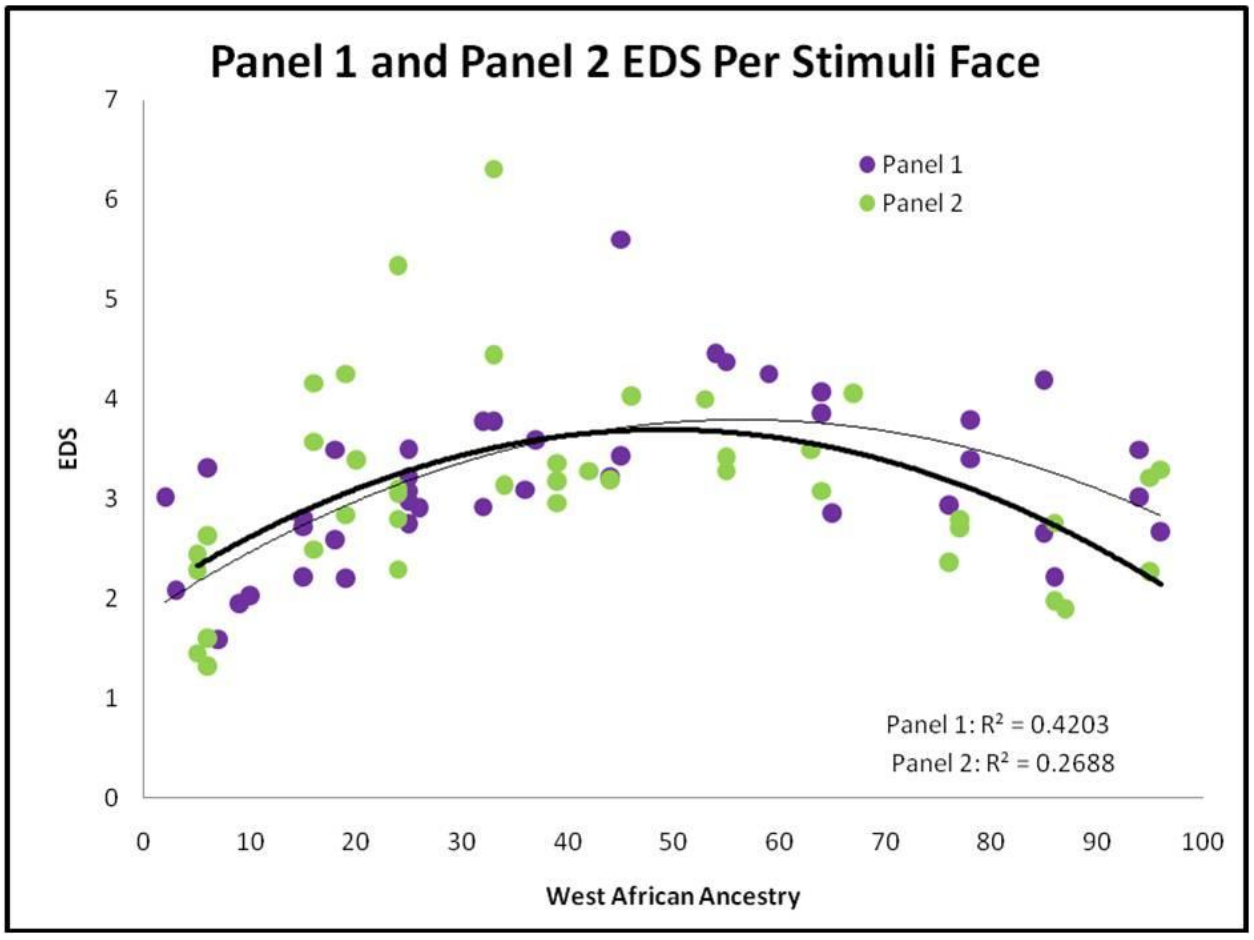


Figure 7. Panel 1 and Panel 2 EDS Per Stimuli Face
West African ancestry is plotted against the EDS for each stimulus face in Panel 1 and Panel 2.

In addition to exploring properties of the individual stimulus face, I grouped the stimulus faces into population groups as well. The average genomic ancestry for each population group was 75.6% West African for the African American group; 30.7% West African for the Brazilian group; and 5.9% West African for the European group. A box plot below shows the average EDS per population group of all 84 stimulus faces (Figure 8). As explained above, each rater saw two sets of photos, which both had stimulus faces from all three population groups: European, African-American, and Brazilian. I ran an ANOVA to compare means of the Euclidean Distance Scores in order to determine whether the "racial"/ethnic background of the stimulus face affected the genomic perception (See Appendix F). The mean EDS was African American (3.13), Brazilian (3.40) and for the European (2.14). Among the three population groups, there was no

statistically significant difference (p=0.052). However, because the p value was so close to significance, I ran individual t-tests, assuming unequal variances, between each population group and did find significant differences between the African-American and European stimulus faces (p=0.03), as well as between the Brazilians and European stimulus faces (p=0.04) (See Appendix E). Faces of predominantly European ancestry were easier to estimate and had a lower EDS as shown in Figure 7. These results indicate that "racial"/ethnic background of the face that is being perceived plays a role in the outcome of genomic ancestry estimation.

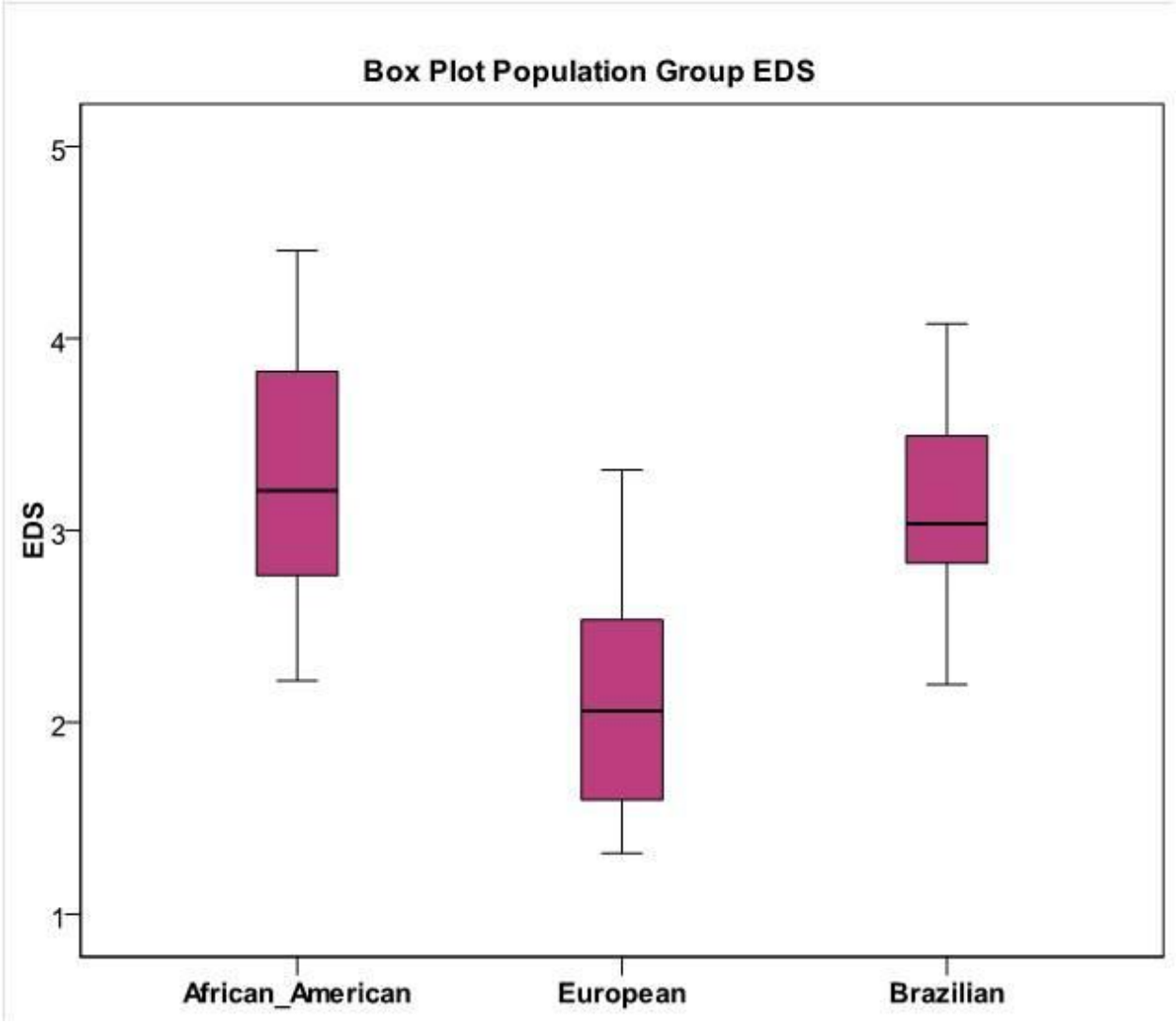


Figure 8. Box Plot Population Group EDS
Box Plots showing the range of EDS per population group.

The Role of Sex in Perceiving Genomic Ancestry

A Euclidean Distance Score was calculated, per 100 raters, for every sex relationship in the data. That is, I calculated two additional EDS for each rater; both an EDS score for their ratings of

female faces, and an EDS for their ratings of male stimulus faces. I ran a general linear model for sex differences (p=0.001) and found that the female stimulus faces had significantly lower EDS than the male stimulus faces. I also tested for interaction between sex and EDS to ensure that the difference in number of male and female stimulus faces did not result in false positives (p=0.402). Figure 9 displays a bar graph showing the variation in EDS among sex rating groups. There was a significant difference between how the female faces were rated, compared to how the male faces were rated (See Appendix I).

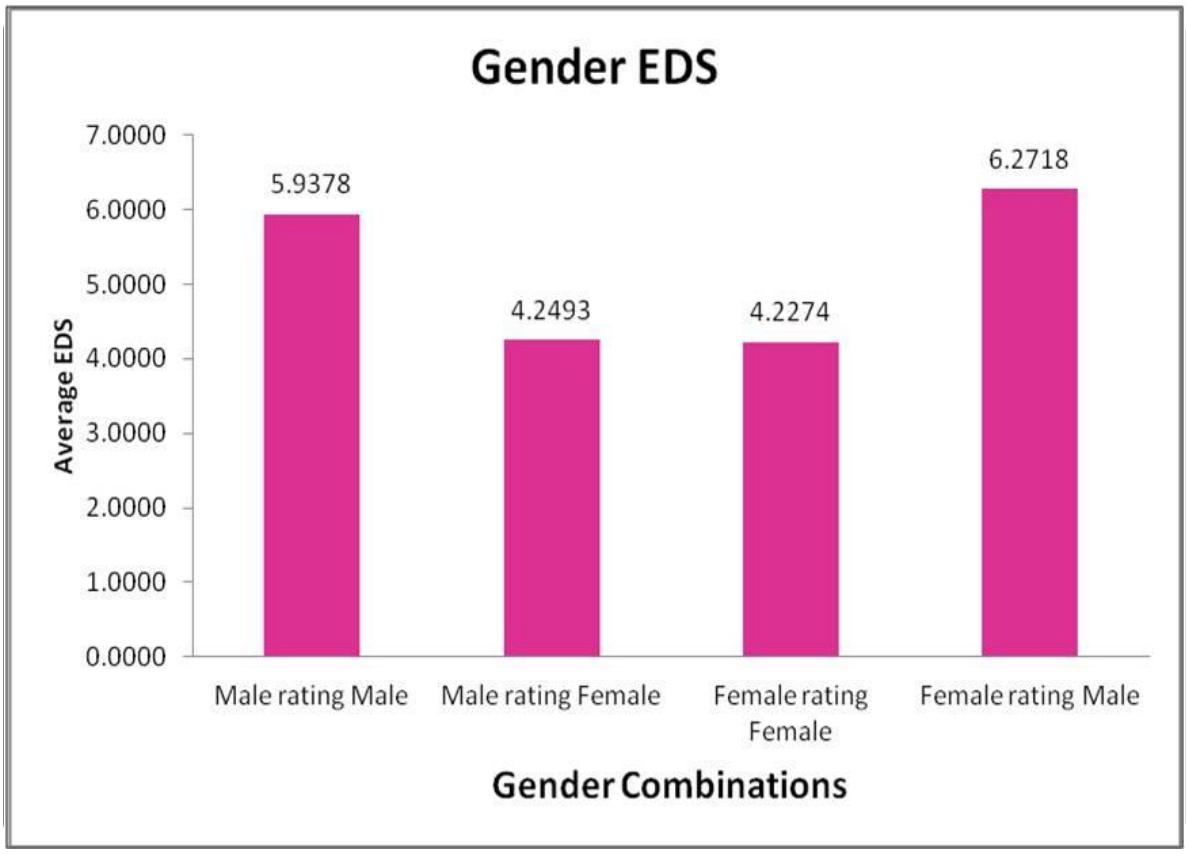


Figure 9. Bar Graph of Sex EDS.
Average EDS all raters rating males and females: male rating male, male rating female, female rating female, and female rating male. The numbers above each bar represent the average EDS per sex rating combination.

Demographic Predictors of EDS

Out of 19 factors, excluding sex, that I tested, three variables were strong trends and, according to the Boneferroni correction, one variable exceeded the significance threshold of 0.00263. The three trends were: “location before 10” (p=0.041); “current neighborhood” (p= 0.031); and “high school European American population” (p=0.011). The one significant factor found was “self reported ability to recognize Black faces” (p=0.0006) (See Appendix F for all 19 predictors

analyzed). Although significant, I found an inverse relationship with this factor. Raters who self reported as being “very good” had an average EDS of 3.69 and those raters who reported as being “average” at recognizing Black faces had an average EDS of 3.15.

The “location before 10” survey question was analyzed in terms of geographical location across the United States. The groups were: Northeast, Midwest, South, and West. The “South” groups’ means were significantly different from the other locations. The “current neighborhood” factor was divided into three categories: urban, rural, and suburban. The “urban” groups’ means were significantly different from the other two groups. Regarding “high school European American population” (p=0.011), the categories were divided up into 0%, 25%, 50%, and 75%. The 50% high school European American high school population category was significant compared to the other groups. Finally, there was one significant inverse relationship (p=0.0006) with raters who responded that their ability to recognize Black faces was “average” (mean EDS= 3.15), as compared with “very good” (mean EDS= 3.69), (See Figure 11).

Although only one factor was found to be significant, I found trends towards significance in almost all the other categories. For instance, raters who responded with having a 50% European population during Elementary school had a trend to having a lower EDS (p=0.052) (See Figure 10). The mean for raters attending high schools with 50% European American populations showed a trend toward (lower) EDS (p=0.11), as compared to students attending high schools with 0%, 25%, 75%, and 100% European American populations. Raters who spent more time on news and blog sites had a lower EDS (p=0.056) than raters who spent more time doing social networking (See Figure 12). There was a trend toward significance in means among internet type frequency.

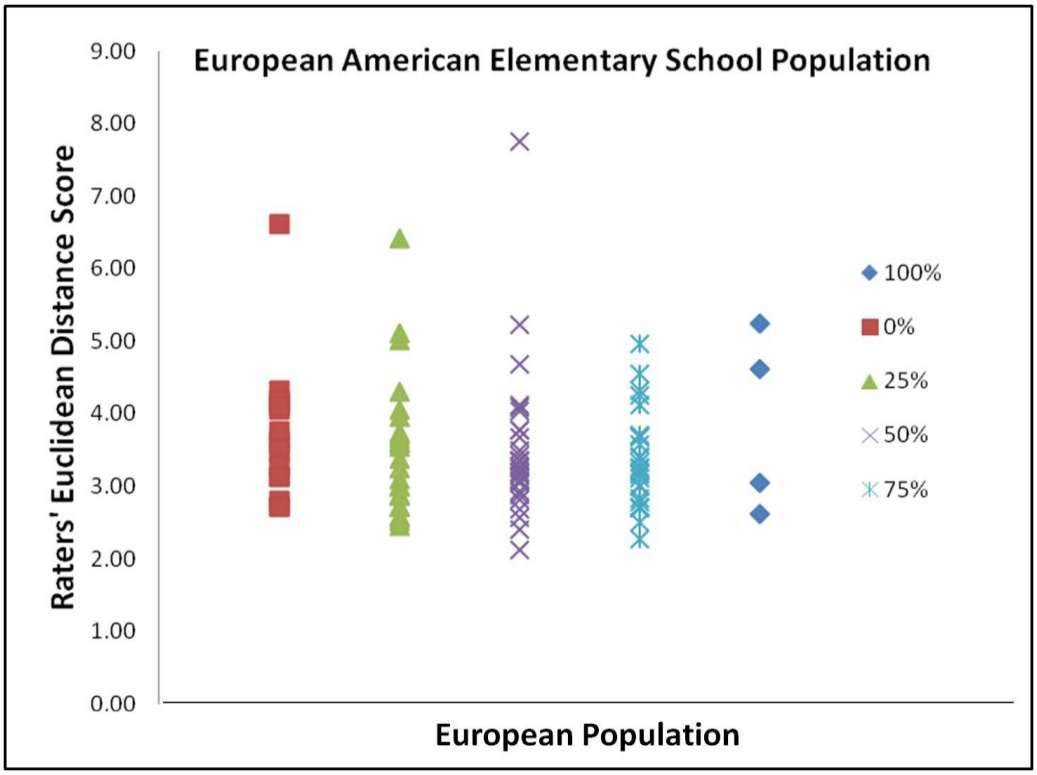


Figure 10. Raters' response to Elementary School population.
The raters with the 50% European Elementary school population have the lowest EDSs.

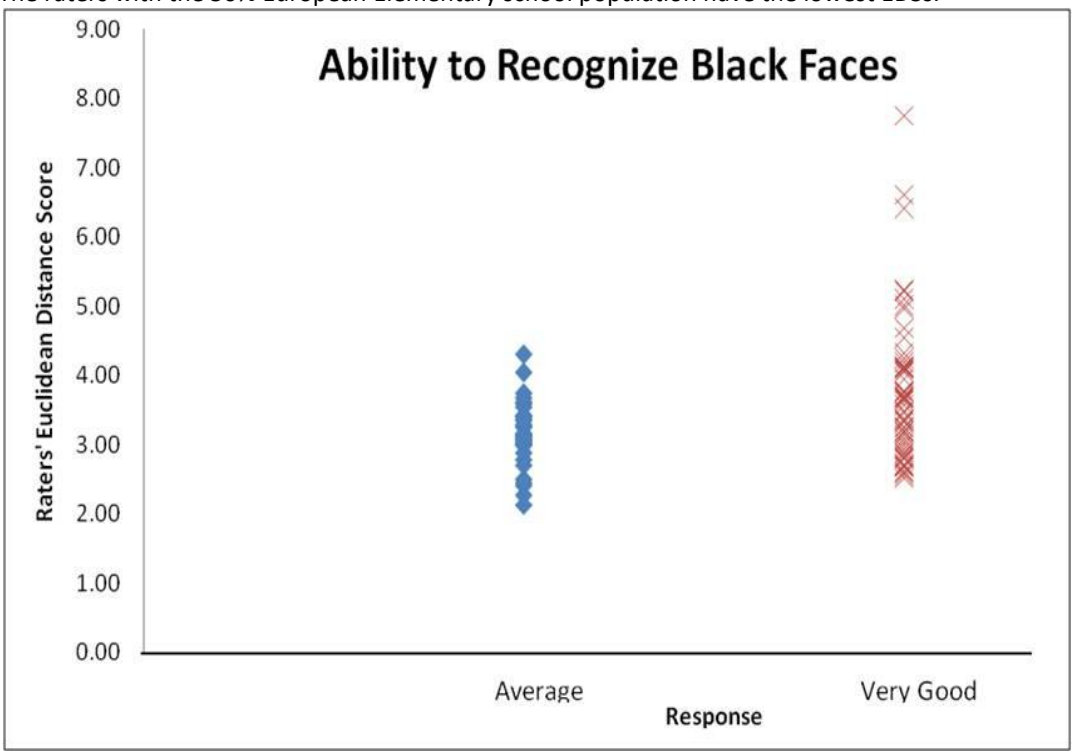


Figure 11. Raters' response to ability to recognize Black faces.
The raters who responded "average" had a lower EDS. Rationale explained in Discussion.

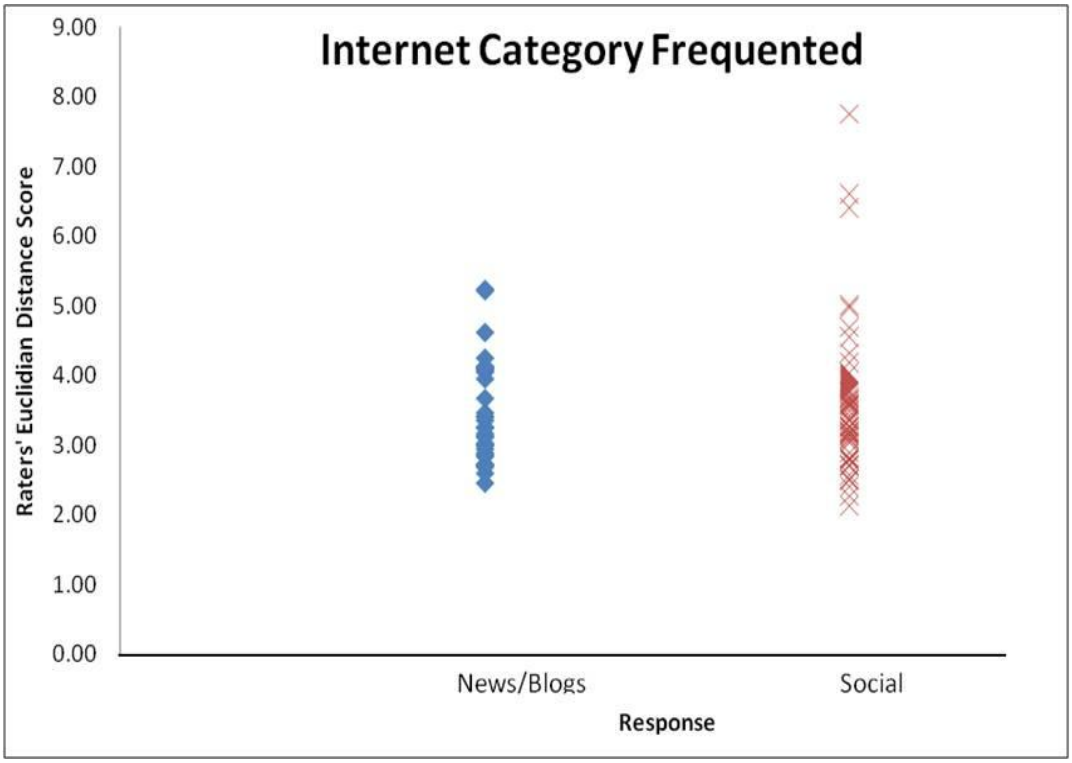


Figure 12. Raters' response to internet type frequented.
The means EDS are 3.31 for news and blogs, and 3.63 for social networking.

In addition to analyzing each predictor at an independent level, I investigated whether all the variables considered together had any influence on the outcome variable Euclidean Distance Score; I applied the Bonferroni correction to ensure that multiple factor testing did not result in false positives and the cutoff for significance given 19 factors was 0.00263. I ran a multiple regression after data reduction via factor analysis. This reduced the number of factors from 20 (including sex) to 12. “Geographic region” was eliminated because we didn’t have enough participants to fully analyze zip code effects, and it was not useful to make this variable dichotomous. I reduced the three categories of “Elementary, Middle, and High School European American population” to “average European friends during Elementary, Middle and High School combined” by averaging the dichotomous coded variables for each category. For example, a participant who had “0%” (0) for Elementary, “not 0%” (1) for Middle, and “not 0%” (1) for High, was coded as (0.67). The same method of combining was done for “mom and dad education”, “self-identified ability to recognize Black faces, White faces, and TV actors“; and, “hours of TV” and “’racial’/ethnic breakdown of favorite TV show” was multiplied for interaction effects (See Appendix G for data reduction list). The factor analysis produced five

principle components that were used as independent variables, along with “age” in multiple regression, and EDS as the dependent factor. The multiple regression resulted in none of the factors being significant in predicting EDS (Table 6).

Table 6. Multiple Regression for PCs of demographic predictors

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	EDS	3.275	.225		14.529	.000
	Age	.010	.009	.114	1.055	.294
	PC1	-.059	.085	-.074	-.699	.487
	PC 2	-.010	.081	-.013	-.129	.898
	PC 3	.117	.081	.146	1.441	.153
	PC 4	.006	.081	.008	.078	.938
	PC 5	-.063	.082	-.079	-.769	.444

CHAPTER 4: DISCUSSION & CONCLUSION

Accuracy in Estimating Genomic Ancestry

Consistent with Klimentidis and Shriver's (2009) work, I found that raters' ability to estimate genomic ancestry in faces was better than if they were to estimate randomly. Klimentidis & Shriver (2009) computed 241 simulations (same as sample *n*) and used a permutation approach. In my study, the larger number of simulations, as well as the shuffling approach both support the hypothesis that raters are better than a random chance at rating genomic ancestry in human faces (p=0.0001). There is some support that most people, with exception of those who may experience face perception disorders e.g. prosopagnosia (although not measured here) rate faces more accurately than estimates that were generated randomly. Three populations (European, African American and Brazilian) have been studied and the trends found in this study should be investigated further and studied in additional populations, including those admixed with East Asian ancestry. Additionally, perception can be studied using genomic ancestry and studying both the perceiver and the face being perceived. Future studies should involve ancestry estimates for both the perceiver and stimulus face to examine possible relationships. Different from my study, it would be equally interesting and important to compare a participants' self genomic estimates to their measured genomic estimates. I suspect a relationship to be present because we saw in my results that humans show some ability in rating faces that they had never seen before. Thus, having a lifetime of experience and exposure to their own face, participants might do better with these ratings. Second, I suspect that there will be a relationship between a participants' self genomic estimate and his or her stimulus estimates. In addition to Klimentidis and Shriver (2009), Chiao and colleagues (2006) observed a relationship between bi-racial participants' and their ability to categorically perceive faces. Participants who were asked to write about their White parent were better at perceiving the White faces and bi-racial participants who were asked to write about their Black parent identified Black stimulus faces fastest. Although these studies were at the "racial"/ethnic level, I believe this relationship may be at the genomic level also.

Perhaps the more that we discover regarding the genes involved with facial features, the more advanced the field of perception can be. Kerri Matthes showed by measuring facial features that there are differences that correspond to more or less West African Ancestry e.g. longer

nose width, or longer lip height. More studies, in order to bridge the cognitive gap between perception and genomics, need to be carried out that focus specifically on the identification of genes responsible for facial features. Based off of the work that M. Shriver has done on facial feature genetics, our many conversations, and the literature on craniofacial disorders, my speculation is that out of approximately 20,000, there are a relatively small number (10-20) of genes responsible for all the facial diversity that is present. Revealing this information will be another major step in alleviating racial and facial discrimination. Educating the world on the genomic similarities that we all share as humans can break down barriers that are due to phenotypic differences. In addition, it is valuable to look at the causal relationship among genes responsible for facial features and perception to understand more fully how humans have evolved. For example, an interesting follow up question to this study is which facial features each participant focused on and whether or not those facial features have evolved more or less rapidly than others to form the wide range of facial features that we see today. Studying these multiple fields cohesively can effectively contribute to the knowledge needed to solve this complex genetic puzzle and cultural issue.

Average Ancestry Assigned to "Black" and "White"

Overall, I showed that there was variation in both the categorical ratings, as well as the genomic estimations. A face was classified as "Black" at a median of ~53% West African Ancestry (over 40 points away from completely West African 100%), while the median for faces categorized as "White" was ~75% European ancestry (only 25 points away from 100% European). As expected, the ancestry percentage at which a face was categorized as "Black" was higher than that of classifying a face as "White". Although higher than I hypothesized 25% was considerably low on the spectrum of West African ancestry. Individuals who have any African admixture are typically labeled "Black" due to the colloquial use of the "one-drop rule". Historically, the rule, which was made a law in some states, posited that any individual with one drop of African blood was considered Black. Looking at the genomic and phenotypic relationship, my results show faces that appear to be "Black" may have genetic markers that support having European ancestors. The median genomic ancestry for classifying a face as "White" (~25% European Ancestry) is lower than that of categorizing a face as "Black" (53%

West African Ancestry). Thus, a stimulus face could fall within a larger range of West African ancestry to be classified as “Black” and a smaller range to be classified as “White”.

There are many other factors that may explain this trend, a combination of both biological and cultural. It is possible that the facial features genes that make a person appear “Black” are stronger (perhaps even dominant), therefore less West African ancestry needs to be present to be detected. In addition, skin pigmentation possibly plays a bigger role than facial features. Figure 4 showed a histogram of the facial categorizations and the relationship observed was consistent for both categorizations. That is, as West African ancestry increased, the number of faces categorized as “White” decreased, and the number of faces categorized as “Black” increased. I also compared the facial categorizations to the melanin indices and saw a similar pattern (See Appendix I). The higher the melanin index (the darker the skin), the more facial categorizations of “Black” were found. The same is true when we see lower melanin indices corresponding to more faces categorized as “White”. Perhaps skin pigmentation was the strongest determinant in the categorization task. Moving forward, studies that involve this protocol, along with stimulus faces that are black and white (false colored) can help elucidate the weight that skin color plays in genomic perception. The United States is incredibly racially conscious, but this aspect of perception needs to be studied more both here and other countries. For instance, Brazilians have a completely different color caste system. A study conducted in Brazil would have had to include more color categories for the faces, and it would be interesting to analyze how the Americans in the dataset would be categorized, compared to the Brazilian stimulus photos. I would propose that because Brazil has such a complex ethnicity nomenclature, they would be better at both rating and categorizing due to a constant interaction, identification, and classification of many ethnicities.

It is interesting to quantify how people dichotomously categorize “race”. We know that humans have a vast capacity for making judgments that then trigger a categorical response. This relationship is known as categorical perception [Beale & Keil 1995]. Although not directly measured in this thesis, I think some of the relationships I observed were due to categorical perception in terms of both categorizing faces and estimating genomic ancestry. I showed that there was variation of perception with both the categorizing task, as well as the genomic rating

task. At an individual level, stimulus faces that had West African ancestry proportions of less than 10% were categorized as “Black” in some cases. Similarly, stimulus faces that had a measured genomic ancestry of less than 10% West African, were rated as having high proportions >80% of West African ancestry. Consistent with Klimentidis and Shriver’s paper (2009), this thesis shows that perception is complex at both the social and biological level.

Euclidean Distance Score per Stimulus Face

Overall, the stimulus faces at the far ends of the ancestry spectrum had the lower Euclidean Distance Scores (See Figure 7). This trend was expected because assumingly, faces that are less admixed would be easier to rate on a percentage scale, close to 0 or 100. However, the EDS for the faces at either end of the ancestry spectrum were not as low (nearly a perfect relationship) as I hypothesized.. I showed a plot of the measured ancestry proportions vs. the Euclidean Distance Scores. We can infer that the closer a stimulus face is to a categorical boundary (in terms of ancestry), the more accurate the genomic ancestry estimates for that face, and the opposite is true for faces that lie in the middle of a categorical spectrum. Because the mind naturally and instinctively puts visual stimulus into categories (Gauthier 2001), perhaps choosing between 0 and 100 was more difficult than choosing between smaller ancestry bins, or choosing between two categories. Ancestry has no elasticity at the ends of the spectrum i.e. no one has negative ancestry proportions. If a face looks to be “White”, and the measured genomic ancestry is 0% West African, an average rating of 10% yields a difference of ten. On the other hand, if a stimulus face has ancestry of 50%, the rater could choose either 40% or 60%, thus making the probability of the estimate being wrong, twice as high. It is more difficult to choose an accurate rating because the potential difference between genomic ancestry and perceived ancestry is greater in the middle of the spectrum.

In this study, each of the 84 stimulus faces had different Euclidean Distance Scores. The lowest EDS for a face was 1.59 with a measured WA genomic ancestry of 7% and the highest EDS was 5.6 with a measured WA genomic ancestry of 45%. Future studies should compare facial landmark values (measured distances between particular locations on the face e.g. nose width, lip height, etc.) among these faces, and other measures that could explain this significant difference. For example, the face with the highest EDS and 45% WA ancestry perhaps may fall

in the middle of the facial feature landmark distances, or this face may have certain facial features that fall along the spectrum of what might be expected regarding ancestry. That is, the nose may have a landmark distance that correspond to more European ancestry (narrower) and lip height landmark distances that correspond to more African ancestry (wider). For future studies, researchers should also examine the effects both with and without skin pigmentation.

Euclidean Distance Scores per Population Group

The sample in this study was comprised of majority African-American raters (*n*=78). Based on the definition of the other-race effect and because the majority of my sample self identified as African American, those stimulus faces in the dataset should have had the lowest or most accurate Euclidean Distance Score. This was not the case regarding the three populations in the stimulus set. However, when I compared the average EDS in the Brazilian (3.18) stimulus faces to the African-American (3.29) stimulus faces, although not significant, there was a trend (*p*=0.62) towards a lower EDS on average for the African-American stimulus faces. More analyses should be carried out to specifically address this question. That is, if there was a bigger sample of both African-American and Brazilian faces (different population groups, but similar ancestry proportions), more substantial conclusions could be drawn regarding the depth of the other race effect and how "racial" perceptions impact ratings. In addition, a sample like mine could be divided into groups based on ancestry (0-10%; 11-20%; 21-30%...91-100%) and EDSs compared among raters to determine whether faces with more or less WA ancestry are rated better or worse. I would speculate that, as shown with the European ratings, faces with both lower and higher proportions of WA ancestry would be rated more accurately than faces with mid range genomic ancestry proportions.

Based on an analogy with the ORE, I speculated that African Americans would be better at rating African-American stimulus faces. However, there was more variation than I hypothesized in the EDS among population groups. As shown in the Box Plot (Figure 7 above), all three population groups overlapped. This may have been the case because of outlier participants who did not fully comprehend the instructions and were estimating ancestry incorrectly. For example, for a face that they perceived as having 0% African ancestry, they answered 100% because they interpreted the face as having 100% European ancestry and vise versa. This is

speculation, but according to the data for a few participants, these trends in backwards ratings were present. For future studies, comprehension may be improved by going through an example, a face that is not included in the study, following the consenting process. I instructed the participants to rate how much African ancestry they thought each face had on a scale from 0-100%. It is possible that when participants' heard "percent", some automatically assigned that quantity to a stimulus face that they felt was at either end of the ancestry spectrum. Another possibility is that the participant was not truly evaluating each face and responded categorically to most of the genomic ratings. For example, rater #44 had responses in multiples of 10 for all 42 stimulus faces including 16 100% genomic estimates. More specific directions can be given indicating that all of the faces will be rated using just one ancestry axis e.g. European, Native American, East Asian, or West African may clarify seemingly confusing instructions.

Does Sex (male or female) of Rater and Stimulus Face Affect EDS?

There was a strong significant relationship regarding how the stimulus faces were rated according to sex (p=0.001). The female stimulus faces were rated more accurately than the male stimulus faces. The literature documents many areas of gender differences in many everyday tasks like performing mathematical problem solving in high school [Hyde, Fennema, & Lamon 1990], risk taking [Byrnes, Miller, & Schafer 1999] etc.; however, no study has compared genomic ratings between sexes. The research does support the contention that there are differences in perception of sex. Nesse, Silverman, and Bortz (1990) showed two faces (a child and an adult), side by side, to participants and asked them to indicate whether they thought the two people in the photo were related. Consistent with my results, they showed that photos with a mother (female) were significantly easier for participants to judge than those that included a father (male). The researchers alluded to the fact that this relationship might have been observed because children's faces are more similar to women, then they are to men. The researchers also found an own sex bias i.e. participants were better able to judge faces that had the same sex as they did. One reason I speculate why female faces may be easier to perceive is that women have evolved to be more distinguishable, exuding effeminate features for sexual selection reasons. However, it is important to note that, future analyses should include ancestry

in the model to test for gender differences. It is possible that my significant results may have resulted from my dataset of having twice as many females than males. Thus, female faces being rated more accurately may be due to having more female faces in the dataset closer to the ends of the ancestry spectrum and thus having lower EDSs associated with them.

Demographic Predictors of EDS

After analyzing 20 variables from the demographic survey and combining covariates, I found that, while the lower Euclidean Distance Scores corresponded to predictors indicating more experience or exposure to various racial/ethnic groups, independent ANOVAs confirmed that only one factor exceeded the Bonferroni cutoff and was significant in predicting genomic ancestry estimation accuracy: ability to recognize Black faces (p=0.0006). Although not significant, the following demographic measures resulted in strong trends towards: location before ten (p=0.041), current neighborhood (p=0.031), high school European-American population (p=0.011). The geographic locations were divided among nationally recognized regions: west, northeast, south, and midwest, and the region that showed a significant difference was the south. One possible explanation for this result is that the weather may play a role in the amount of interaction between people. The nicer the weather, the more likely people are to be outside, at social events. Also, according to the 2010 census, more than half of African Americans live in the south and along the lines of more experience and exposure; it is likely that more people will develop a better sense of distinguishing faces if they see more. One interesting finding was the "ability to recognize Black faces" question (p= 0.0006). We would expect to see that raters who self responded as being "very good" would have a lower EDS than raters who responded as being "average", but the opposite was the case. This inverse relationship may have occurred because the question asks directly about ability and perhaps the pressure of "being very good" could have negatively affected a rater's performance. Conversely, those raters who claimed to be "average" may have responded modestly.

Although not statistically significant, the direction of ancestry estimation accuracy suggests that experience increases a participants' ability to rate genomic ancestry in human faces. For instance, when observing the neighborhood environment, there is an inverse relationship with

higher populated areas i.e. *urban* having the lowest EDS, and lower populated areas i.e. *rural*, having the highest EDS. This trend implies that areas with larger populations would result in better accuracy of rating genomic ancestry. As we explore the racial/ethnic breakdown of school, we see that EDS is also inversely related to diversity. That is, the more diverse the racial/ethnic breakdown of the school, the more of a variety of faces a person is exposed to, and thus the lower the EDS.

There are other ways to measure experience and exposure through different types of questions. For example, if I were interested in identity and how it might shape one's ability in perceiving genomic ancestry, I would ask questions that measured, specifically, factors of identity e.g. raters who were bi- or multiracial might have a different perspective than those who were not. In addition to asking different questions, there is an additional method that I considered, but did not employ, to measure estimation accuracy. The EDS gives a measure of discrepancy from the actual GE, and the difference provides a direction of discrepancy from the actual GE. Another method involves analyzing the β or slope values for each rater's estimates. This method would have also allowed me to evaluate the direction of estimation. I completed all three calculations for the dataset (See Appendix J), but only used the EDS values for the predictor analysis, and the difference values to evaluate over- and underestimations.

Additional Considerations

I realize that there are many additional analyses and questions to consider. For example, I presented in Appendix J melanin indices plotted against facial categorizations. Because I was looking at the face as a totality, skin was beyond the scope of this thesis but as stated above, is clearly important and needs to be studied in this context more. There are currently no studies that explore how skin pigmentation affects genomic ratings in human faces. Another consideration is the combination of this type of study with gaze tracking data to determine which specific facial features are being looked at when a participant is asked to rate facial ancestry. However, now that we know humans don't rate genomic ancestry randomly, but that they have some skill at it, specific facial components e.g. skin pigmentation and facial features should be explored in greater depth.

In addition to exploring specific effects of parts of the face, I think it is worth investigating whether or not perception can be modified, with the use of training videos involving faces with specific ancestry distributions. In linguistics, distributional learning in infants has been shown to affect how an infant discriminates sounds [Maye, 2002]. If this relationship is transferrable from sounds to faces, I'd expect to see better ancestry rating abilities after administering a pre-exposure rating task and a post-exposure rating task [from conversations with potential collaborator, Dan Weiss]. There should be a difference in how well participants perceive genomic ancestry.

CONCLUSION

The results and trends that I present in this thesis are preliminary and should be further investigated with larger sample sizes, as well as different population groups (e.g. Puerto Ricans). Larger sample sizes may provide the power needed to result in more of the factors being significant, or it may support my results of insignificance. Also, different combinations of population groups as both raters and stimulus faces are necessary to evaluate more completely patterns of genomic perception from faces. For example, population groups around the world should be used as both perceivers and stimulus to be perceived.

The trends found in my data were consistent with the hypothesis that experience and exposure yield better perception abilities. This information may provide very useful to the field of criminal justice, specifically eyewitness descriptions. This study may help provide accuracy to an eyewitness account to better draw a sketch of a suspect. For instance, a crime occurs and there are two main witnesses; one of them describes the perpetrator as "dark-skinned" and the other witness describes the suspect as "light-skinned". As we study and understand more about the intricacies of how people perceive faces, we can provide corrections to a sketch, using results from studies like this one. Information can be used from both eyewitness descriptions, but afterwards possibly enhanced according to each witness's demographics in order to make a more accurate suspect sketch. This type of procedure would be most useful if meaningful information were obtained with the least amount needed from the witness. Most likely, trauma is coupled with being an eyewitness and I'm sure that person would not want to be probed any more than necessary. Similarly, in the medical field, treatment is potentially influenced by the

perceptions of the doctor. It is important to train doctors to be mindful of both the advantages and disadvantages of their perceptions and how they may be biased in ways they didn't know.

The other-race effect, own-race bias, and cross-racial identification effect all posit that the "racial" group that one is a part of makes him or her better at recognizing that group based on having more experience and exposure. All of these theories can and should now be examined using genomic ancestry estimates rather than only memorization and categorizing tasks. We see that this relationship is complex and variable in that even the categorizing task yielded a variety in responses regarding the stimulus faces with varying ancestry. Because this study included a fairly small sample size (n=100), comparing both the categorical and genomic ratings might result in false conclusions. For example, we had some faces with more than 50% West African ancestry being categorized as "White". Whether this was a mistake or not is too difficult to tease out with such a small sample set. If this pattern is present with a sample size of power, we can validly conclude that something particular can explain this trend. In addition, because there were multiple predictors being analyzed, more participants are needed to increase power in comparison analysis.

APPENDIX A: Sampling Details

For the first sampling, Justin Whitt created a text filter in MicrosoftExcel which standardized all the answers for statistical analysis. This filter also reduces data entry time by allowing us to simply select the survey answer corresponding to that each participant from a drop down menu and move to the next, allowing for more time to evaluate for data entry errors. After Justin created the filter, he entered the data (n=53); I reviewed his data entry to ensure there were no errors. I entered the data for the second sampling (n=10) and cross referenced the surveys with the excel worksheet three times. Marquette Moore, Jeron Rowland, and I entered the data, as well as checked each other's entries for the last sampling (n=37). We recorded all data and stored it in MicrosoftExcel files on password protected computers to keep the identity of the participants secure.

APPENDIX B: Demographic Survey

ID#_____

Face Ancestry Analysis: Demographic Questionnaire

Please answer the following questions to the best of your knowledge.

Background Information

Age?

Sex?

- male
- female
- Other:

How do you identify yourself?

- Arab
- Asian/Pacific Islander
- Black/African/African American
- White/European/European American
- Hispanic
- Indigenous or aboriginal
- Latino
- Multiracial (please specify)
- Other:

How do others usually identify you?

- Arab
- Asian/Pacific Islander
- Black/African/African American
- White/European/European American
- Hispanic
- Indigenous or aboriginal
- Latino

- Multiracial (please specify)
- Other:

How do you identify your mother?

- Arab
- Asian/Pacific Islander
- Black/African/African American
- White/European/European American
- Hispanic
- Indigenous or aboriginal
- Latino
- Multiracial (please specify)
- Other:

How do you identify your father?

- Arab
- Asian/Pacific Islander
- Black/African/African American
- White/European/European American
- Hispanic
- Indigenous or aboriginal
- Latino
- Multiracial (please specify)
- Other:

Who raised you? *Mark all that apply.

- mother
- father
- both parents
- grandfather
- grandmother
- both grandparents

☐ other (please specify)

Where were you born?

○ United States

○ Other: _____

Before age 10, about how many people were in your household (including yourself)? _____

Before age 10, what were the ages of the people in your household? *Mark all that apply

☐ 0-5

☐ 6-10

☐ 11-15

☐ 16-20

☐ 21-up

☐ Other: _____

What city/county/state/zip code did you reside in for the majority of your life before age 10? _____

What city/county/state/zip code did you reside in for the majority of your life after age 10? _____

What city/county/state/zip code do you currently reside? _____

How long have you lived where you currently live? _____

Which of the following best describes the area in which you currently live?

○ Rural

○ Suburban

○ Urban

Before the age of 10, how many friends did you have that were European American (White)?

○ None (0%)

○ Almost a quarter (less than 25%)

○ About half (about 50%)

○ Majority (more than 75%)

○ All (100%)

What percentage of your Elementary School population was European American (White)?

- None (0%)
- Almost none (less than 25%)
- About half (about 50%)
- Mostly all (more than 75%)
- All (100%)

What percentage of your Middle School population was European American (White)?

- None (0%)
- Almost none (less than 25%)
- About half (about 50%)
- Mostly all (more than 75%)
- All (100%)

What percentage of your High School population was European American (White)?

- None (0%)
- Almost none (less than 25%)
- About half (about 50%)
- Mostly all (more than 75%)
- All (100%)

Currently, how many friends do you have that are European American (White)?

- None (0%)
- Almost none (less than 25%)
- About half (about 50%)
- Mostly all (more than 75%)
- All (100%)

What racial/ethnic group are you most attracted to? *Mark all that apply.

- Arab
- Asian/Pacific Islander
- Black/African/African American
- White/European/European American
- Hispanic

- [ ] Indigenous or aboriginal
- [ ] Latino
- [ ] Multiracial (please specify)
- [ ] Other:

What racial/ethnic group do you typically date? *Mark all that apply.

- [ ] Arab
- [ ] Asian/Pacific Islander
- [ ] Black/African/African American
- [ ] White/European/European American
- [ ] Hispanic
- [ ] Indigenous or aboriginal
- [ ] Latino
- [ ] Multiracial (please specify)
- [ ] Other:

What is the highest level of education you have completed?

- ○ Some high school
- ○ High school diploma
- ○ Some college
- ○ Bachelor's degree
- ○ Some graduate school
- ○ Graduate degree

What is your mother's highest level of education completed?

- ○ Some high school
- ○ High school diploma
- ○ Some college
- ○ Bachelor's degree
- ○ Some graduate school
- ○ Graduate degree

What is your father's highest level of education completed?

- Some high school
- High school diploma
- Some college
- Bachelor's degree
- Some graduate school
- Graduate degree

Which amount best describes your parents' combined (if applicable) income before the age of 10?

- Under $17,999
- $18,000-$64,999
- Over $65,000

Which amount best describes your parents' combined (if applicable) income after the age of 10?

- Under $17,999
- $18,000-$64,999
- Over $65,000

Face Recognition

How would you classify your ability to recognize Black faces?

- very good
- average
- very poor

How would you classify your ability to recognize White faces?

- very good
- average
- very poor

How would you classify your ability to recognize faces of actors on television?

- very good
- average
- very poor

Do you have a computer?

- yes
- no

How many hours a day do you spend on the computer?

- 0-2 hours
- 2-5 hours
- over 5 hours

Which website category would you prefer?

- news
- social networking
- blogs

Do you own a TV?

- yes
- no

How many hours of TV do you watch a day?

- 0-2 hours
- 2-5 hours
- over 5 hours

What is the name of your favorite TV show?

Give the racial/ethnic breakdown of the cast of your favorite TV show.

- all Black
- mostly Black
- mostly White
- all White
- racially diverse
- Other:

Have you ever experienced any head injuries?

- yes
- no

If you have experienced any head injuries, which side of your head (right or left)?

- left

- right

Describe your own skin tone.

- dark
- brown
- light
- White
- Other:

Estimate your proportion of African genomic ancestry from 0-100%

APPENDIX C: Matlab Scripts for Simulations

*To change the number of simulations to match 100, 1000, or 100000, simply change that value in the script (highlighted below).

Shuffle Data

```
% Import data from Excel
[data,txt,raw] = xlsread('FAA.xls','p1 and p1 ratings
sepaestimated','D4:D45');
[data2,txt,raw] = xlsread('FAA.xls','survey answers','AR2:AR101');
% Import actual EDS

%preassign eds variable with 100,000 arbitrary ones
eds=ones(100000,1);

% Call the SHAKE function which shuffles the values of data
% For loop calculates eds for 100,000 shuffles of the original data
for i=1:100000
    sdata=SHAKE(data); % sdata is the shuffled data
    diffsq=(data-sdata).^2; % Calculate the square difference b/
original and  shuffled data
    eds(i)=sqrt(sum(diffsq))/42; % Calculate eds
end
```

Statistical Data

```
figure
hist(eds,2:0.5:8) %Plot histogram of eds with xlimit of 2 through 8
in 0.5 increments
xlabel('EDS')
ylabel('Frequency')
title('Simulation Histogram')

figure
hist(data2,2:0.5:8) %Plot histogram of eds with xlimit of 2 through
8 in 0.5 increments
xlabel('EDS')
ylabel('Frequency')
title('EDS Histogram')

meds=mean(eds) % Calculate mean of eds
maxeds=max(eds) % Find maximum eds value
mineds=min(eds) % Find minimum eds value
stdveds=std(eds) % Find standard deviation of eds
mode_eds=mode(eds) % Find mode of eds
med_eds=median(eds) % Find median of eds
var_eds=var(eds) % Find variance of eds
range=maxeds-mineds % Calculate the range of eds
st_error=stdveds/sqrt(length(eds)) % Calculate the standard error
of eds
```

SHAKE Function

```
function [Y, I, J] = shake(X,dim)
% SHAKE - Randomize a matrix along a specific dimension
%   Y = SHAKE(X) randomizes the order of the elements in each
column of the
%   2D matrix. For N-D matrices it randomizes along the first non-
singleton
%   dimension.
%
%   SHAKE(X,DIM) randomizes along the dimension DIM.
%
%   [Y, I, J] = SHAKE(X) returns indices so that Y = X(I) and X =
Y(J).
%
%   Example:
%     A = [1 2 3 ; 4 5 6 ; 7 8 9 ; 10 11 12] ; % see <SLM> on the
FEX ...
%     Dim = 2 ;
%     B = shake(A,Dim)  % ->, e.g.
%      %  3     2     1
%      %  6     4     5
%      %  7     8     9
%      % 11    10    12%
%     C = sort(B,Dim) % -> equals A!
%
%     The function of SHAKE can be thought of as holding a matrix
and shake
%     in a particular direction (dimension), so that elements are
getting
%     shuffled within that direction only.
%
%   See also RAND, SORT, RANDPERM
%   and RANDSWAP on the File Exchange

% for Matlab R13
% version 4.1 (may 2008)
% (c) Jos van der Geest
% email: jos@jasen.nl

% History
% Created: dec 2005
% Revisions
% 1.1 : changed the meaning of the DIM. Now DIM==1 works along the
rows, preserving
% columns, like in <sort>.
% 2.0 (aug 2006) : randomize along any dimension
% 2.1 (aug 2006) : output indices argument
% 3.0 (oct 2006) : new & easier algorithm
% 4.0 (dec 2006) : fixed major error in 3.0
% 4.1 (may 2008) : fixed error for scalar input
```

SHAKE Script contd.

```
error(nargchk(1,2,nargin)) ;

if nargin==1,
    dim = min(find(size(X)>1)) ;
elseif (numel(dim) ~= 1) || (fix(dim) ~= dim) || (dim < 1),
    error('Shake:DimensionError','Dimension argument must be a
positive integer scalar.') ;
end

% we are shaking the indices
I = reshape(1:numel(X),size(X)) ;

if numel(X) < 2 || dim > ndims(X) || size(X,dim) < 2,
    % in some cases, do nothing
else
    % put the dimension of interest first
    [I,ndim] = shiftdim(I,dim-1) ;
    sz = size(I) ;
    % reshape it into a 2D matrix
    % we'll randomize along rows
    I = reshape(I,sz(1),[]) ;
    [ri,ri] = sort(rand(size(I)),1) ;  % get new row indices
    ci = repmat([1:size(I,2)],size(I,1),1) ; % but keep old column
indices
    I = I(sub2ind(size(I),ri,ci)) ; % retrieve values
    % restore the size and dimensions
    I = shiftdim(reshape(I,sz),ndim) ;
end

% re-index
Y = X(I) ;

if nargout==3,
    J = zeros(size(X)) ;
    J(I) = 1:numel(J) ;
end
```

APPENDIX D: *t* Test for Panel 1 and Panel 2

	Panel 1	*Panel 2*		
Mean			3.542351	3.451246
Variance			0.63832	0.646852
Observations	50	50		
Pearson Correlation			-0.12189	
Hypothesized Mean Difference			0	
df			49	
t Stat			0.5365	
P(T<=t) one-tail			0.29702	
t Critical one-tail			1.676551	
P(T<=t) two-tail			0.594041	
t Critical two-tail			2.009575	

APPENDIX E: Regression Analysis Panel 1 & 2 EDS per face

SUMMARY OUTPUT PANEL 1

Regression Statistics					
Multiple R	0.349256749				
R Square	0.121980277				
Adjusted R Square	0.100029784				
Standard Error	0.74403894				
Observations	42				
ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	3.076356516076356 51	3.076356516	5.5570631665570 63	0.023387087
Residual	40	22.14375777	0.553593944		
Total	41	25.22011429			
	Coefficients	***Standard Error***	***t Stat***		***P-value***
Intercept	2.769804731	0.205444005	13.48204216		1.83004E-16
WA Ancestry	0.009578763	0.004063374	2.357342395		0.023387087

SUMMARY OUTPUT PANEL 2

Regression Statistics					
Multiple R	0.008110649				
R Square	6.57826E-05				
Adjusted R Square	-0.024932573				
Standard Error	0.981884658				
Observations	42				
ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.002537002	0.002537002	0.002631478	0.959343432
Residual	40	38.56389928	0.964097482		
Total	41	38.56643629			
	Coefficients	*Standard Error*	*t Stat*	*P-value*	
Intercept	3.07892662	0.270607973	11.37781193	4.10823E-14	
WA Ancestry	0.000273856	0.005338538	0.051297938	0.959343432	

APPENDIX F: *ANOVA* Population Group Effects on EDS

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
African Americans	28	92.11	3.289643	1.007663
Brazilians	44	139.968	3.181091	0.466731
Europeans	12	30.92	2.576667	1.104133

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	4.502517	2	2.251258	3.068771	0.051919	3.109311
Within Groups	59.42181	81	0.733603			
Total	63.92433	83				

t-Test: Two-Sample Assuming Unequal Variances

	African Americans	*Europeans*
Mean	3.289643	2.576667
Variance	1.007663	1.104133
Observations	28	12
Hypothesized Mean Difference	0	
df	20	
t Stat	1.992836	
P(T<=t) one-tail	0.030052	
t Critical one-tail	1.724718	
P(T<=t) two-tail	0.060103	
t Critical two-tail	2.085963	

t-Test: Two-Sample Assuming Unequal Variances

	Brazilians	*Europeans*
Mean	3.181091	2.576667
Variance	0.466731	1.104133
Observations	44	12
Hypothesized Mean Difference	0	
df	14	
t Stat	1.886812	
P(T<=t) one-tail	0.040049	
t Critical one-tail	1.76131	
P(T<=t) two-tail	0.080098	
t Critical two-tail	2.144787	

APPENDIX G: *ANOVA* Codes

1. Age: 18-21 (0); 22-up (1)
2. Sex: male (0); female (1)
3. People in household before ten: <4 (0); 4 and up (1)
4. Location before 10: northeast (0); Midwest (1); south (2); west (3)
5. Current neighborhood: urban (0); rural (1); suburban (2)
6. European friends before ten: 0% (0); 25%(1); 50% (2); 75% (3); 100% (4)
7. Elementary school European American population: 0% (0); 25%(1); 50% (2); 75% (3); 100% (4)
8. Middle school European American population: 0% (0); 25%(1); 50% (2); 75% (3); 100% (4)
9. High school European American population: 0% (0); 25%(1); 50% (2); 75% (3); 100% (4)
10. Current European American friends: 0% (0); 25%(1); 50% (2); 75% (3); 100% (4)
11. Highest education level: some/grad school degree (0); some/high school diploma (1); some college (2); bachelors degree (3)
12. Mom highest education level: some/grad school degree (0); some/high school diploma (1); some college (2); bachelors degree (3)
13. Dad highest education level: some/grad school degree (0); some/high school diploma (1); some college (2); bachelors degree (3)
14. Parents income before ten: > $65,000 (0); <$17,000 (1); between $18,000 and $65,000 (2)
15. Ability to recognize Black faces: average (0); very good (1)
16. Ability to recognize White faces: average (0); very good (1)
17. Ability to recognize actors on TV: average (0); very good (1)
18. Type of websites frequented: social networking (0); news and blogs (1)
19. Hours of TV: 0-2 hours (0); 3-5 (1); over 5 (2)
20. Racial ethnic breakdown of favorite TV show: mostly White (0); all Black (1); all White (2); racially diverse (3); mostly Black (4)

Data Reduction Prior to Factor Analysis: From 20 predictors to 12

1. Sex: female (0); male (1)
2. People in household before ten: less than 4 (0); 4 and up (1)
3. Current neighborhood: urban (0); non-urban (1)
4. European friends before ten: 0% and 25%(0); 50%, 75%, 100% (1)
5. Current European Friends
6. Average Elem, Mid, High School Euro Population: 0% (0); more than 0% (1)
7. Rater's highest level of education
8. Mom and dad highest education level: high school diploma, some college (0); bachelors degree, some grad school, grad degree (1)
9. Parents income before ten: less than $65,000 (0); more than $65,000 (1)
10. Ability to recognize Black faces, White faces, and actors on TV: average (0); very good (1)
11. Type of websites frequented: social networking (0); news and blogs (1)
12. Hours TV * Racial/Ethnic Breakdown of favorite TV show

FACTOR ANALYSIS:

Five principal components were extracted

- PC1 parents income, dad ed, mom ed, household number
- PC2 presence of euro in school, euro friends before 10
- PC3 recognize blk faces, recognize actors
- PC4 sex, neighborhood
- PC5 websites

APPENDIX H: Predictor *t*-test & *ANOVA* Outputs

1. Age: 18-21 (0); 22-up (1)

t-Test: Two-Sample Assuming Equal Variances		
	18-21	*22-up*
Mean	3.501268	3.567931
Variance	0.991191	0.456153
Observations	71	29
Pooled Variance	0.838323	
Hypothesized Mean Difference	0	
df	98	
t Stat	-0.33038	
P(T<=t) one-tail	0.37091	
t Critical one-tail	1.660551	
P(T<=t) two-tail	0.74182	
t Critical two-tail	1.984467	

2. Sex: male (0); female (1)

t-Test: Two-Sample Assuming Unequal Variances		
	Male	*Female*
Mean	3.519672131	3.522051282
Variance	0.919229891	0.712979892
Observations	61	39
Hypothesized Mean Difference	0	
df	88	
t Stat	-0.013027718	
P(T<=t) one-tail	0.494817585	
t Critical one-tail	1.66235403	
P(T<=t) two-tail	0.989635169	
t Critical two-tail	1.987289823	

3. People in household before ten:

t-Test: Two-Sample Assuming Unequal Variances		
	Less than 4 in House	*More than 4*
Mean	3.575833333	3.43775
Variance	1.120936582	0.401392244
Observations	60	40
Hypothesized Mean Difference	0	
df	97	
t Stat	0.814837682	
P(T<=t) one-tail	0.208580669	
t Critical one-tail	1.660714611	
P(T<=t) two-tail	0.417161339	
t Critical two-tail	1.984723136	

4. Location before ten:

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Northeast	Between Groups	7.679	16	.480	.	.
	Within Groups	.000	0	.		
	Total	7.679	16			
Midwest	Between Groups	10.465	17	.616	.	.
	Within Groups	.000	0	.		
	Total	10.465	17			
South	Between Groups	59.560	48	1.241	9.973	.041
	Within Groups	.373	3	.124		
	Total	59.933	51			
West	Between Groups	1.173	4	.293	.	.
	Within Groups	.000	0	.		
	Total	1.173	4			

5. Current neighborhood:

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
urban	Between Groups	29.852	40	.746	12.161	.031
	Within Groups	.184	3	.061		
	Total	30.036	43			
rural	Between Groups	20.536	8	2.567	.	.
	Within Groups	.000	0	.		
	Total	20.536	8			
suburban	Between Groups	27.088	42	.645	.831	.680
	Within Groups	2.330	3	.777		
	Total	29.417	45			

6. European friends before ten:

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
eurofrnd100	Between Groups	2.002	2	1.001	.	.
	Within Groups	.000	0	.		
	Total	2.002	2			
eurofrnd0	Between Groups	20.242	19	1.065	.	.
	Within Groups	.000	0	.		
	Total	20.242	19			
eurofrnd25	Between Groups	10.464	33	.317	2.422	.335
	Within Groups	.262	2	.131		
	Total	10.726	35			
eurofrnd50	Between Groups	32.646	21	1.555	.	.
	Within Groups	.000	0	.		
	Total	32.646	21			
eurofrnd75	Between Groups	12.085	19	.636	.	.
	Within Groups	.000	0	.		
	Total	12.085	19			

7. Elementary school European American population:

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
element100	Between Groups	2.574	2	1.287	.	.
	Within Groups	.000	0	.		
	Total	2.574	2			
element0	Between Groups	12.958	14	.926	.	.
	Within Groups	.000	0	.		
	Total	12.958	14			
element25	Between Groups	20.989	23	.913	.	.
	Within Groups	.000	0	.		
	Total	20.989	23			
element50	Between Groups	30.057	26	1.156	231.207	.052
	Within Groups	.005	1	.005		
	Total	30.062	27			
element75	Between Groups	10.859	24	.452	.	.
	Within Groups	.000	0	.		
	Total	10.859	24			

8. Middle school European American population:

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
mid100	Between Groups	.192	1	.192	.	.
	Within Groups	.000	0	.		
	Total	.192	1			
mid0	Between Groups	3.538	11	.322	.	.
	Within Groups	.000	0	.		
	Total	3.538	11			
mid25	Between Groups	14.508	23	.631	.	.
	Within Groups	.000	0	.		
	Total	14.508	23			
mid50	Between Groups	36.173	29	1.247	.753	.741
	Within Groups	1.656	1	1.656		
	Total	37.829	30			
mid75	Between Groups	7.417	25	.297	.	.
	Within Groups	.000	0	.		
	Total	7.417	25			

9. High school European American population:

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
high0	Between Groups	10.448	9	1.161	.	.
	Within Groups	.000	0	.		
	Total	10.448	9			
high25	Between Groups	9.392	21	.447	.	.
	Within Groups	.000	0	.		
	Total	9.392	21			
high50	Between Groups	35.552	31	1.147	90.838	.011
	Within Groups	.025	2	.013		
	Total	35.577	33			
high75	Between Groups	18.095	25	.724	17.213	.189
	Within Groups	.042	1	.042		
	Total	18.137	26			

10. Current European American friends:

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
current0	Between Groups	5.750	8	.719	.	.
	Within Groups	.000	0	.		
	Total	5.750	8			
current25	Between Groups	34.883	46	.758	1.358	.464
	Within Groups	1.675	3	.558		
	Total	36.558	49			
current50	Between Groups	28.952	23	1.259	.	.
	Within Groups	.000	0	.		
	Total	28.952	23			
current75	Between Groups	4.718	10	.472	.	.
	Within Groups	.000	0	.		
	Total	4.718	10			

11. Highest education level

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
grad	Between Groups	5.278	13	.406	.	.
	Within Groups	.000	0	.		
	Total	5.278	13			
high	Between Groups	24.148	29	.833	12.850	.218
	Within Groups	.065	1	.065		
	Total	24.212	30			
somcoll	Between Groups	33.557	32	1.049	.237	.976
	Within Groups	8.847	2	4.423		
	Total	42.403	34			
bach	Between Groups	5.928	12	.494	.	.
	Within Groups	.000	0	.		
	Total	5.928	12			

12. Mom highest education level:

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
momgrad	Between Groups	16.955	30	.565	16.721	.192
	Within Groups	.034	1	.034		
	Total	16.989	31			
momhigh	Between Groups	11.901	14	.850	.	.
	Within Groups	.000	0	.		
	Total	11.901	14			
momsomcoll	Between Groups	21.707	16	1.357	.	.
	Within Groups	.000	0	.		
	Total	21.707	16			
mombach	Between Groups	18.979	29	.654	.086	.998
	Within Groups	7.644	1	7.644		
	Total	26.623	30			

13. Dad highest education level:

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
dadgrad	Between Groups	32.915	30	1.097	2.967	.434
	Within Groups	.370	1	.370		
	Total	33.285	31			
dadhigh	Between Groups	16.895	21	.805	.	.
	Within Groups	.000	0	.		
	Total	16.895	21			
dadsomcoll	Between Groups	14.223	17	.837	.	.
	Within Groups	.000	0	.		
	Total	14.223	17			
dadbach	Between Groups	11.888	20	.594	.	.
	Within Groups	.000	0	.		
	Total	11.888	20			

14. Parents income before ten:

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
below17	Between Groups	8.962	14	.640	.	.
	Within Groups	.000	0	.		
	Total	8.962	14			
btween18and65	Between Groups	44.851	39	1.150	.388	.932
	Within Groups	8.892	3	2.964		
	Total	53.743	42			
above65	Between Groups	8.105	30	.270	1.265	.537
	Within Groups	.427	2	.214		
	Total	8.532	32			

15. Ability to recognize Black faces:

t-Test: Two-Sample Assuming Unequal Variances		
	Very Good	*Average*
Mean	3.686818182	3.154848485
Variance	1.037006643	0.231094508
Observations	66	33
Hypothesized Mean Difference	0	
df	97	
t Stat	3.52963424	
P(T<=t) one-tail	0.000319097	
t Critical one-tail	1.660714611	
P(T<=t) two-tail	0.000638194	
t Critical two-tail	1.984723136	

16. Ability to recognize White faces: average (0); very good (1)

t-Test: Two-Sample Assuming Unequal Variances		
	Very Good	*Average*
Mean	3.635272727	3.374772727
Variance	0.903880943	0.738760412
Observations	55	44
Hypothesized Mean Difference	0	
df	96	
t Stat	1.429158534	
P(T<=t) one-tail	0.078102329	
t Critical one-tail	1.660881441	
P(T<=t) two-tail	0.156204657	
t Critical two-tail	1.984984263	

17. Ability to recognize actors on TV:

t-Test: Two-Sample Assuming Unequal Variances		
	Very Good	*Average*
Mean	3.431041667	3.651914894
Variance	0.630405275	1.062176688
Observations	48	47
Hypothesized Mean Difference	0	
df	86	
t Stat	-1.16844603	
P(T<=t) one-tail	0.122927871	
t Critical one-tail	1.66276545	
P(T<=t) two-tail	0.245855742	
t Critical two-tail	1.987934166	

18. Type of websites frequented:

t-Test: Two-Sample Assuming Unequal Variances		
	Social Networking	*News and Blogs*
Mean	3.633859649	3.305428571
Variance	1.066088409	0.360749076
Observations	57	35
Hypothesized Mean Difference	0	
df	90	
t Stat	1.928267242	
P(T<=t) one-tail	0.028487127	
t Critical one-tail	1.661961085	
P(T<=t) two-tail	0.056974255	
t Critical two-tail	1.986674497	

19. Hours of TV:

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
zerotwohours	Between Groups	49.075	57	.861	2.159	.238
	Within Groups	1.595	4	.399		
	Total	50.670	61			
threetofive	Between Groups	14.548	27	.539	.941	.688
	Within Groups	.572	1	.572		
	Total	15.120	28			
overfive	Between Groups	8.449	5	1.690	.	.
	Within Groups	.000	0	.		
	Total	8.449	5			

20. Racial ethnic breakdown of favorite TV show:

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
mostWhite	Between Groups	14.247	31	.460	.429	.886
	Within Groups	2.142	2	1.071		
	Total	16.390	33			
allBlack	Between Groups	22.770	13	1.752	.	.
	Within Groups	.000	0	.		
	Total	22.770	13			
raciallydiverse	Between Groups	20.442	25	.818	.	.
	Within Groups	.000	0	.		
	Total	20.442	25			
mostlyBlack	Between Groups	17.298	15	1.153	.	.
	Within Groups	.000	0	.		
	Total	17.298	15			

APPENDIX I: Sex Differences in EDS

Anova: Single Factor

SUMMARY

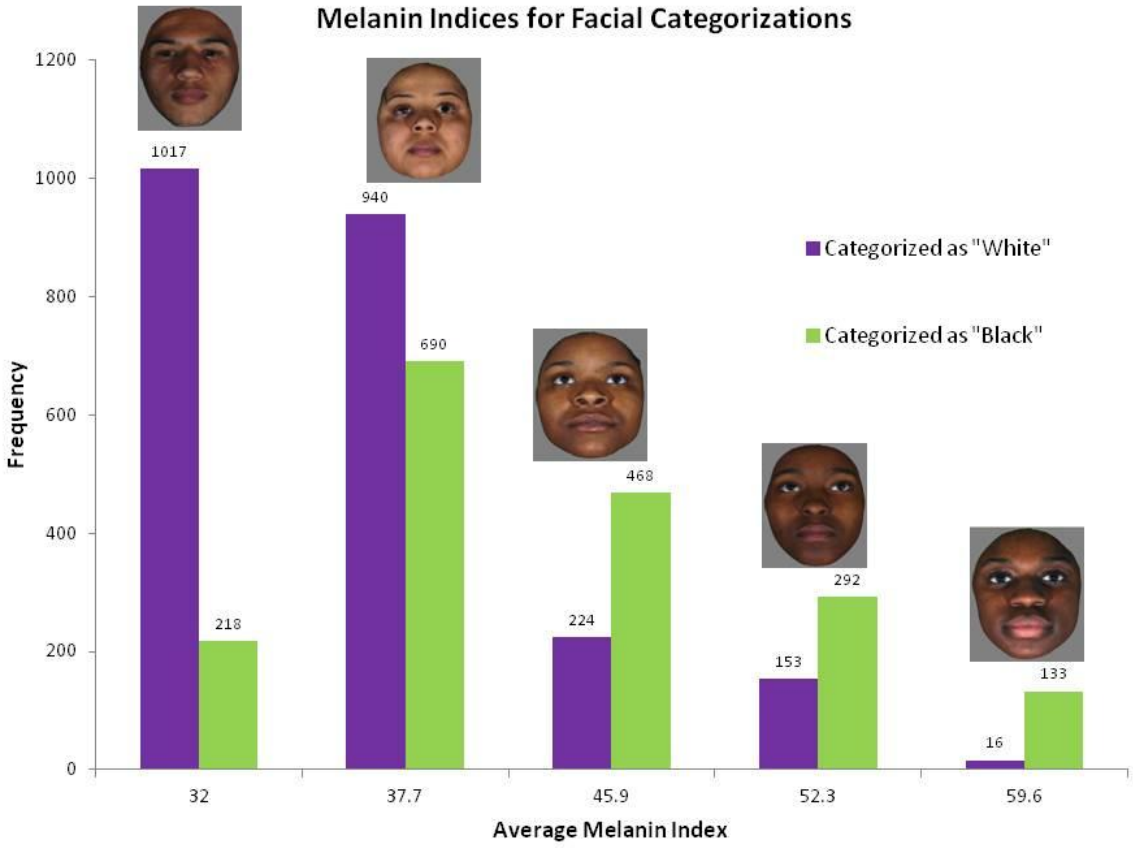
Groups	*Count*	*Sum*	*Average*	*Variance*
Men rating Men	61	362.203202	5.937757	3.081485
Men rating Women	61	259.209154	4.24933	1.097119
Women rating Women	39	164.8674884	4.227371	1.059265
Women rating Men	39	244.600317	6.271803	6.676394

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	169.6116704	3	56.53722	20.34492	1.61E-11	2.650677
Within Groups	544.6712762	196	2.778935			
Total	714.2829466	199				

APPENDIX J: Melanin Indices for Facial Categorizations

*X axis is in arbitrary categories.



APPENDIX K: (β values)

rater#	EDS	Beta	rater#	EDS	Beta
1	4.5138	0.9356	51	2.5870	0.8150
2	4.0610	0.9390	52	3.0004	1.1160
3	5.0889	0.9190	53	2.4118	0.8457
4	2.8949	0.9104	54	3.4891	0.5473
5	2.8564	0.6949	55	3.4574	1.2122
6	3.6077	0.6946	56	3.4078	1.0030
7	3.3969	0.4281	57	3.0004	1.1160
8	2.8337	0.8257	58	3.3545	1.0129
9	3.2607	1.2427	59	3.5747	0.6075
10	3.1484	0.9953	60	4.6230	0.6738
11	3.5892	0.9461	61	3.0184	0.7843
12	2.2694	0.8521	62	3.0514	0.8810
13	2.9919	1.0255	63	3.2843	0.7140
14	3.7432	0.8794	64	2.6759	1.0957
15	2.9408	0.9773	65	2.5153	0.8705
16	4.3223	0.7220	66	2.7191	1.0753
17	3.5352	0.6187	67	3.0907	0.9532
18	4.6852	1.0772	68	3.4191	0.8280
19	3.2514	0.8261	69	3.2766	1.0661
20	3.6878	1.1132	70	3.0820	0.7817
21	3.1757	0.7015	71	2.7058	0.9021
22	4.0432	0.9765	72	2.6970	0.9582
23	3.6950	1.0427	73	2.7100	0.8416
24	3.3848	1.2903	74	6.6118	0.0706
25	2.7989	0.9873	75	2.7064	0.7176
26	4.0898	1.4426	76	2.8223	0.7596
27	3.2513	0.5835	77	2.8823	0.5478
28	3.9537	1.0036	78	3.0419	1.1591
29	3.1331	1.0258	79	4.5476	1.0406
30	2.6212	0.9677	80	4.0489	0.3813
31	2.7164	1.1683	81	4.1151	1.3341
32	3.1642	1.3025	82	3.3500	1.1802
33	4.3223	1.0861	83	3.1124	0.7866
34	4.1335	0.8012	84	4.2470	0.9022
35	3.3610	0.9426	85	3.7008	0.6906
36	3.6338	0.8286	86	3.7479	0.7118
37	2.5696	0.9995	87	2.6753	0.7524
38	5.2174	0.5858	88	3.6740	0.9752
39	3.7662	0.8668	89	2.7049	1.0011
40	3.6595	0.8227	90	3.2460	0.5249
41	4.1857	0.7440	91	2.4464	0.8290
42	5.1115	0.9620	92	5.2356	0.9396
43	3.9994	0.6658	93	3.1344	0.7644
44	4.9705	1.0751	94	3.0755	0.5048
45	3.1284	0.9104	95	4.1016	1.0125
46	4.1216	1.0899	96	3.5698	0.5923
47	6.4076	-0.0107	97	3.1962	1.0338
48	3.7813	0.5511	98	3.1855	0.8890
49	2.4990	0.9879	99	3.3253	0.8157
50	2.7828	0.8038	100	3.6665	1.2141

REFERENCES

1. Ramon, Busigny, Rossion (2009). Impaired holistic processing of unfamiliar individual faces in acquired prosopagnosia. *Neuropsyhologia.* 48(4), 933-944.
2. Kennerknecht, Grueter, Welling, Wentzek, Horst, Edwards & Grueter (2006). First report of prevalence of non-syndromic hereditary prosopagnosia (HPA). *American Journal of Medical Genetics,* 140, 1617-1622.
3. Russell, Duchaine & Nakayama (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic Bulletin & Review,* 16, 252-257.
4. Little A, Apicella C, & Marlowe F (2007). Preferences for symmetry in human faces in two cultures: data from the UK and the Hadza, an isolated group of hunter-gathers. *Proceedings of the Royal Society.* 274 (1629) 3113-3117.
5. Malpass RS, Kravitz J (1969) Recognition for faces of own and other race. *Journal of Personal Social Psychology,* 13, 330-334.
6. de Heering, de Liedekerke, Deboni &Rossion (2009). The role of experience during childhood in shaping the other-race effect. *Developmental Science,* DOI: 10.1111/j.1467-7687.2009.00876.x.
7. Steffensmeier S & Demuth S (2006). Ethnicity and Judges' Sentencing Decisions: Hispanic-Black-White Comparisons. *Criminology,* 39(1) 145-178.
8. Sporer, S.L. (2001).Recognizing faces of other ethnic groups: An interaction of theories. *Psychology Public Policy, and Law,* 7(1), 36–97.
9. Hehman E, Mania EW, & Gaertner SL (2010). Where the Division Lies: Common Ingroup Identity Moderates the Cross-Race Facial-Recognition Effect. *Journal of Experimental Social Psychology,* 46(2), 445-448.
10. Meissner & Brigham (2001). Thirty Years of Investigating the Own-Race Bias in memory for Faces A Meta-Analytic Review. *Psychology, Public Policy and Law,* 7, 3-35.
11. Sangrigoli, De Schonen (2004). Recognition of own-race and other-race faces by three-month-old infants. *Journal of Child Psychology Psychiatry,* 45, 1219-1227.
12. Horry, R., Wright, D.B., Tredoux, C.G. (2010). Recognition and context memory for faces from own and other ethnic groups: A remember-know investigation. *Memory & Cognition,* 38 (2), 134-141.
13. Sangrigoli, Pallier, Argenti, Ventureyra & de Schonen (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science,* 16, 440-444.
14. de Heering, Houthuys, Rossion (2007). Holistic face processing is mature at 4 years of age: Evidence from the composite face effect. *Journal of Experimental Child Psychology.* 2001; 80:347-371.
15. Lebrecht, Pierce, Tarr, Tanaka (2009). Perceptual Other-Race Training Reduces Implicit Racial Bios. *Plos One*, 4, 1: 1-7.
16. Greenwald & Krieger (2006). Implicit Bias: Scientific Foundations. *California Law Review,* 94, 4: 945-967.

17. Levin & Beale (2000). Categorical perception occurs in newly learned faces, other-race faces, and inverted faces. *Perception and Psycholphysics,* 62, 2: 386-401.
18. Calder, Young, Perrett, Etcoff, & Rowland (1996). Categorical Perception of Morphed Facial Expressions. *Visual Cognition,* 3, 2: 81-117.
19. Levin & Angelone (2001). Categorical Perception of Race. *Perception,* 31: 567-578.
20. Klimentidis & Shriver (2009). Estimating Genetic Ancestry Proportions from Faces. *PLoS ONE,* 4, e4460.
21. Pfaff, Parra, Bonilla, Heister, McKeigue, Kamboh, Hutchinson, Ferrell, Boerwinkle & Shriver (2001). Population Structure in Admixed Populations: Effect of Admixture Dynamics on the pattern of Linkage Disequilibrium. *American Journal of Human Genetics,* 68, 198-297.
22. Parra, Kittles & Shriver (2004). Implications of correlations between skin color and genetic ancestry for biomedical research. *Nature Genetics Supplement,* 36, S54-S60.
23. Cann, Stoneking, & Wilson (1987). Mitochondrial DNA and human evolution. *Nature*, 325: 31-36.
24. Lieberman, McBratney, Krovitz (2002). The evolution and development of crania form in *Homo sapiens*. *Proceedings National Academy of Sciences,* USA, 99: 1134-1139.
25. White, Asfaw, DeGusta (2003). Pleistocene *Homo sapiens* from Middle Awash, Ethiopia. *Nature*, 423: 742-747.
26. Jobling, Tyler-Smith (2003). The human Y chromosome: an evolutionary marker comes of age. *Nature Review Genetics*. 4: 598-612.
27. Graves (2004). *The Race Myth: Why We Pretend Race Exists in America*. New York. The Penguin Group.
28. Jobling, Hurles, & Tyler Smith (2004). *Human Evolutionary Genetics: Origins, People, & Diseases.* New York. Garland Publishing.
29. Glass, Li (1953). The Dynamics of Racial Intermixture-an Analysis Based on the American Negro. *The American Journal of Human Genetics,* 5, 1: 1-20.
30. Parra, Marcini, Akey, Martinson, Batzer, Cooper, Forrester, Allison, Deka, Ferrell & Shriver (1998). Estimating African American Admixture Proportions by Use of Poulation Specific Alleles. *American Journal of Human Genetics.* 63: 1839-1851.
31. Neel (1973). “Private” Genetic Variants and the Frequency of Mutation Among South American Indians. *Proc. Nat. Acad. Sci, USA*, 70, 12: 3311-3315.
32. Reed (1973). Number of gene loci required for accurate estimation of ancestral population proportions in individual human hybrids. *Nature*, 244: 575-576.
33. Chakraborty, Kamboh, Ferrell (1991). “Unique” alleles in admixed populations: a strategy for determining “hereditary” population differences of disease frequencies. *Ethnic Discussions,* 1, 3: 245-256.

34. Shriver, Smith, Jin, Marcini, Akey, Deka, Ferrell (1997). Ethnic-Affiliation Estimation by Use of Poulation-Specific DNA Markers. *American Journal of Human Genetics.* 60: 957-964.
35. Pfaff, Parra, Bonilla, Hiester, McKeigue, Kamboh, Hutchinson, Ferrell, Boerwinkle, Shriver (2001). Population Structure in Admixed Populations: Effect of Admixture Dynamics on the pattern of Linkage Disequilibrium. *American Journal of Human Genetics.* 68: 198-207.
36. Matthes C (2008). A Preliminary Genetic Investigation of Normal Variation in Facial Features. Penn State University Master's Thesis. Mintz, S (2007). *Digital History.* Retrieved 7/13/2011 from http://www.digitalhistory.uh.edu
37. Parra F, Amada R, Lambertucci J, Rocha J, Antunes C, & Pena S (2003). Color and Genomic Ancestry in Brazilians. *PNAS,* 100(1) 177-182.
38. Doniger, Wendy. (2004). The Mythology of Self-Imitation in Passing: Race, Gender and Politics [PDF document]. Received from: http://divinity.uchicago.edu/martycenter/publications/webforum/122004/doniger.pdf
39. Kennedy, Randall. (2001). Racial Passing. *Ohio State Law Journal.* 62: 1-28.
40. Stuckert, Robert P. (1958). African Ancestry of the White American Population. *The Ohio Journal of Science*, 53(3), 155-160.
41. Crary, David. "Passing for White." *South Florida Sun-Sentinel.* 1 November 2003. Web 5 September 2010.
42. Ellison, Ralph. *Shadow and Act.* 1964. Print. Hardcover.
43. Robinson, Amy (1994). It Takes One to Know One: Passing and Communities of Common Interest. *Critical Inquiry,* 20 (4), 715-736.
44. Halder, Yang, Kranzler, Stein, Shriver & Gelernter (2009). Measurement of Admixture Proportions and Description of Admixture Structure in Different U.S. Populations. *Human Mutation,* 30, 1299-1309.
45. Bothwell, Brigham & Malpass (1989). Cross-Racial Identification. *Personality and Social Psychology Bulletin,* 15, 19-25.
46. Sinha M, Larkin EK, Elston RC, Redline S (2006) Self-reported race and genetic admixture. N Engl J Med 354(4): 421–2.
47. Chiao, Heck, Nakayama, & Ambady (2006). Priming Race in Biracial Observers Affects Visual Search for Black and White Faces. *Psychological Science,* 17, 5: 387-392.
48. Beale & Keil (1995). Categorical effects in the perception of faces. *Cognition*, 57: 217-239.
49. Gauthier, I & Nelson, C.A. (2001). The Development of Face Expertise. *Current Opinion in Neurobiology*, 11:219-224.
50. Hyde, Fennema, & Lamon (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin,* 107, 2: 139-155.

51. Byrnes, Miller, & Schafer (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin,* 125, 3: 367-383.
52. Nesse, R.M., Silverman, A., Bortz, A. (1990). Sex Difference in Ability to Recognize Family Resemblance. *Ethology and Sociobiology,* 11, 11-21.
53. Maye, Werker, Gerken (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition,* 82, B101-111.
54. Chance, Goldstein & McBride (1975). Differential experience and recognition memory of faces. *Journal of Social Psychology,* 97, 243-253.
55. Barkowitz & Brigham [study 1] (1982). Recognition of faces: Own-race bias, incentive, and time delay. *Journal of Applied Social Psychology,* 12, 255-268.
56. Malpass, Lavigueur & Weldon (1973). Verbal and visual training in face recognition. *Perception and Psychophysics,* 14, 285-292.
57. Galper (1973). Functional race membership and recognition of faces. *Perceptual and Motor Skills,* 37, 455-462.
58. Brigham & Barkowitz (1978). Do "they all look alike?" The effect of race, sex, experience, and attitudes on the ability to recognize faces. *Journal of Applied Psychology,* 8, 306-318.
59. Brigham & Williamson (1979). Cross-racial recognition and age: When you're over 60, do they still "all look alike?" *Personality and Social Psychology Bulletin,* 5, 218-222.
60. Ellis & Deregowski (1981). Within-race and between-race recognition of transformed and untransformed faces. *American Journal of Psychology,* 94, 27-35.
61. Devine & Malpass (1985). *Differential face recognition: A less than convincing explanation.* Paper presented at the meetings of the Psychonomic Society, Philadelphia.
62. Pinkham, Sasson, Calkins, Richard, Hughett, Gur & Gur (2008). The other-race effect in face processing among African American and Caucasian individuals with schizophrenia. *American Journal of Psychiatry,* 165, 639-645.
63. Lindsay, Jack & Christian (1991). Other-race face perception. *Journal of Applied Psychology,* 74, 587-589.
64. Ferguson, Rhodes and Lee & Sriram (2001). 'They all look alike to me': Prejudice and cross-race face recognition. *British Journal of Psychology,* 92, 567-577.
65. Tanaka, Kiefer & Bukach (2004). A holistic account of the own-race effect in face recognition: evidence from a cross-cultural study. *Cognition,* 93, B1-B9.
66. Fiset, Blais, Tanaka, Arguin, Bub, & Gosselin (2009). The Information subtending the other-race effect. *Journal of Vision*, *9*, 508 508a, http://journalofvision.org/9/8/508/, doi:10.1167/9.8.508.
67. Pascalis, de Haan & Nelson (2002). Is Face Processing Species-Specific During the First year of Life? *Science,* 296, 1321-1322.
68. Saffran, Johnson, Aslin, Newport (1999). Statistical learning of tone sequences by human infants ad adults. *Cognition,* 70, 27-52.

69. Chiroro, P.M., Tredoux, C.G., Radaelli, S., Meissner, C.A. (2008). Recognizing faces across continents: The effect of within-race variations on the own-race bias in face recognition. *Psychonomic Bulletin & Review,* 15 (6), 1089-1092.