

# UC Irvine

## UC Irvine Electronic Theses and Dissertations

### Title

Reconsidering Facial Attractiveness: A Systematic Multivariate Approach to Identifying the Ethnicity-Specific Cues That Define Beauty

### Permalink

<https://escholarship.org/uc/item/6cq0g34b>

### Author

Banyard, Derek A.

### Publication Date

2018

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,  
IRVINE

Reconsidering Facial Attractiveness: A Systematic Multivariate Approach to Identifying the  
Ethnicity-Specific Cues That Define Beauty

THESIS

submitted in partial satisfaction of the requirements  
for the degree of

MASTER OF SCIENCES

in Biomedical & Translational Sciences

by

Derek Anthony Banyard

Thesis Committee:  
Professor Gregory Evans, Chair  
Professor Donald Hoffman  
Assistant Professor John Billimek

2018



## DEDICATION

To

my dearest wife, Angela, and my precious daughter, Ella

in appreciation for loving me and bearing with me through my cacophony of duties as I transition to this final phase of my training.

A Reminder

If there is no struggle, there is no progress.

Frederick Douglass

An Inspiration

Out of the huts of history's shame  
I rise  
Up from a past that's rooted in pain  
I rise  
I'm a black ocean, leaping and wide,  
Welling and swelling I bear in the tide.

Leaving behind nights of terror and fear  
I rise  
Into a daybreak that's wondrously clear  
I rise  
Bringing the gifts that my ancestors gave,  
I am the dream and the hope of the slave.  
I rise  
I rise  
I rise.

Maya Angelou  
"Still I Rise"

I Love You Both.

# TABLE OF CONTENTS

	Page
LIST OF FIGURES	v
LIST OF TABLES	vi
ACKNOWLEDGMENTS	vii
ABSTRACT OF THE THESIS	viii
INTRODUCTION	1
CHAPTER 1: Historical Tools of Aesthetic Training and Assessment	7
Neoclassical Canons	7
Phi or 'The Golden Ratio'	8
CHAPTER 2: Validation of Historical Tools in Aesthetics	10
Review of the Literature	10
Model Papers	15
CHAPTER 3: Components of Attractiveness	32
Averageness (Koinophilia)	33
Symmetry	36
Sexual Dimorphism	38
Other Considerations	40
CHAPTER 4: Methodology to Objectively Quantify Facial Attractiveness	43
Subject Recruitment	45
Photographic Sample Acquisition	47
Image Grouping and Analysis	48
Machine Learning	57
Rater Recruitment	58
Subjective and Objective Rating Systems	59
Statistical Analysis	60
Anticipated Results and Next Steps	62
CHAPTER 5: Summary and Conclusions	64
REFERENCES	66

## LIST OF FIGURES

		Page
Figure 1	Anthropometric Landmarks for proportions assessment (frontal view)	49
Figure 2	Anthropometric Landmarks for proportions assessment (lateral view)	50
Figure 3	Anthropometric Angles for proportions assessment (lateral view)	52
Figure 4	Anthropometric proportions for symmetry assessment (frontal view)	54
Figure 5	Anthropometric proportions associated with aging (frontal view)	55
Figure 6	Facial proportions assessment for assessment of sexual dimorphism	56

## LIST OF TABLES

		Page
Table 1	Essential Steps in Objectively Quantifying Aesthetic Surgical Results	6
Table 2	The neoclassical Canons	8
Table 3	Attractiveness Quotients (Bashour)	19
Table 4	Multiple linear regression of nodal point versus attractiveness quotient (Bashour)	24
Table 5	Neoclassical canons assessed from nodal points (Schmid <i>et al.</i> )	25
Table 6	Golden ratios analyzed (Schmid <i>et al.</i> )	27
Table 7	Models used after stepwise regression and variable selection (Schmid <i>et al.</i> )	28
Table 8	Canon formulas, symmetry pairs, and golden ratios in the final models (Schmid <i>et al.</i> )	29
Table 9	Anthropometric landmarks for proportion assessment	51

## ACKNOWLEDGMENTS

I would like to express deep appreciation to my committee chair, Professor Gregory Evans, whose guidance was essential to helping me maintain a clinically relevant, stepwise approach to an incredibly difficult research question.

I would also like to thank my thesis committee members, Professor John Billimek and Professor Donald Hoffman. Dr. Billimek has been an exceptional teacher and was critical to me maintaining focus on the right methodological approach. Dr. Hoffman has helped to reignite the cognitive science curiosity I engendered during my undergraduate studies.

I would like to offer a special thanks to Dr. Val Lambros who provided permission to include copyrighted composite photographs as part of my thesis and who has eagerly provided his wisdom and expertise regarding facial aging and the intricacies of taking high quality three-dimensional medical photographs.

I would like to thank Professor Charless Fowlkes for providing assistance and education regarding reasonable approaches to machine learning and computational vision. I'd also like to thank Professor Craig Stark for providing expertise on objectively measuring facial attractiveness.

Thank you, Professor Alan Widgerow, for always pushing me and helping me to keep things in perspective.

In addition, I would like to thank Rollin Read and Canfield Scientific, Inc. for their support of my project and future collaborations.

Finally, I would like to thank Professor Al Aly, past president of the Aesthetic Surgery Education and Research Foundation, for planting the seed of this project in my mind three years ago, and for his continued encouragement and support of me to continue striving to overcome the most significant of challenges.



## **ABSTRACT OF THE THESIS**

Reconsidering Facial Attractiveness: A Systematic Multivariate Approach to Identifying the Ethnicity-Specific Cues That Define Beauty

By

Derek Anthony Banyard

Master of Science in Biomedical & Translational Sciences

University of California, Irvine, 2018

Professor Gregory Evans, Chair

Historically, education for practitioners who deal in aesthetics has been rooted in outdated concepts like the 'Golden Ratio,' or phi, and the neoclassical canons. The field of facial attractiveness research is extensive and includes data that invalidates the concepts of phi and the neoclassical canons as tools that can be used to assess attractiveness in various ethnicities. Here, I provide analysis of the facial attractiveness research and propose a novel methodology and statistical model to objectively quantify ethnicity-specific beauty, which includes, but is not limited to, the components of averageness (koinophilia), symmetry, sexual dimorphism, youthfulness, and skin tone. Additionally, consideration is given to the perception of beauty as shaped by the age, gender and ethnicity of the subject as well as the observer. An objective tool for the classification of attractiveness is extremely complex, therefore subjective and objective ratings systems will be employed to glean meaningful data that may one day elucidate the factors that define the beauty gestalt.

## INTRODUCTION

Facial attractiveness is an area of intense interest for the scientific community at large. A simple *Google Scholar* search on “facial attractiveness” yields approximately 12,800 articles during the 20-year period from 1997 to 2017; this is approximately a 14,000% increase over the preceding 20 years. A field that was once dominated by psychologists is now piquing the interest of the medical community, and primarily, surgeons who are focused on aesthetically pleasing outcomes.

The oldest *Pubmed* record on the study of facial attractiveness was an odontologic case study on the pathogenesis of facial asymmetry presented at the Proceedings of the Royal Society of Medicine in 1911.<sup>1</sup> Since that time, the fields of dermatology, oromaxillofacial surgery, and plastic surgery have increased their research on the topic in hopes of improving results for their patients.

In particular, the significant rise of interest observed in the field of plastic and reconstructive surgery is two-fold. First, plastic surgeons have always been committed to optimizing aesthetic results for their patients, but are cognizant to the fact that postsurgical outcomes are primarily judged subjectively by the patient, which are often guided by the subjective opinion of the surgeon.<sup>2</sup> Second, and more broadly, the field of medicine is experiencing an increased emphasis on the importance of evidence-based medicine and therefore pursuing treatment strategies that yield optimal results for the patient that can be quantified objectively. Unfortunately, there are very few tools at the plastic surgeon’s armamentarium that facilitate surgical planning or allow for the assessment postsurgical outcomes in an objective fashion.<sup>2,3</sup>

The importance of facial attractiveness research cannot be overstated. The most influential pioneer in this field, Dr. Judith Langlois, became interested in the science behind attractiveness when she validated a study demonstrating that infants aged 3 to 6 months preferred to stare at more attractive faces over less attractive ones.<sup>4</sup> Research has since demonstrated that less attractive individuals are judged as being less well-adjusted, socially appealing, and academically and interpersonally competent when compared to their attractive counterparts.<sup>5</sup> Infants deemed attractive receive more positive maternal attention than do infants deemed unattractive.<sup>5</sup> And perhaps not surprisingly, the level of an individual's attractiveness can be directly correlated with upward economic mobility, the likelihood of being hired for a job, and the probability of gaining a promotion.<sup>6</sup>

However, the primary driving force behind attractiveness research, or beauty, is mate selection, or the Darwinian term, "survival of the fittest." Evolutionary psychologists posit that attractiveness is directly correlated to superior physical and reproductive traits, which connote a survival benefit that is necessary for the propagation of one's lineage and species. For instance, researchers have demonstrated a link between secondary sexual traits and parasite resistance,<sup>7,8</sup> while others have linked the degree of male masculinity to semen quality.<sup>9</sup> Interestingly, research has demonstrated that females rate attractiveness differently when they are ovulating versus other times in their reproductive cycle.

As the fields of psychology and biology have progressively intertwined, and specifically the fields of cognitive psychology, evolutionary psychology, neuroscience, and face research, theories have emerged that offer more scientifically-based explanations for the components that constitute the aesthetic ideal. An article at the turn of the century in *Discover Magazine* entitled "Do You Love This Face?" summarized interviews with thought-

leaders in attractiveness research and ultimately concluded that the key tenets of an attractive face include symmetry, youthfulness, sexual dimorphism and averageness.<sup>10</sup> Since that time, conflicting data has emerged with regards to symmetry,<sup>9,11</sup> and sexual dimorphism.<sup>9,12</sup> The body of literature is limited with regards to the impact of youthfulness on facial attractiveness, but it is no secret that individuals seeking aesthetic enhancement are often pursuing a more youthful appearance. Additionally, a feature that has recently been called into question is the relationship that skin tone implies on the overall perception of attractiveness.<sup>9,13</sup>

One facet that has not been challenged, however, is the contribution that averageness, or koinophilia, which is an evolutionary biology term for *the inclination towards the average*, impacts the overall gestalt of beauty perception. The idea that a composite of faces, or averaged face, is more attractive than an individual face is widely attributed to Sir Francis Galton who projected a series of faces onto a single photographic film and generated a face that was more attractive than any of the individual faces used to construct this composite.<sup>14</sup> In what is considered to be the landmark paper on the topic, “Attractive Faces Are Only Average,” Langlois and Roggman revealed, for the first time, that when observers view a composite headshot photo comprised of the mathematical average of individual headshot photos, the observers universally rate the composite photo as more attractive than any of the individual headshots.<sup>15</sup> To this day, researchers often employ averaging techniques as the gold standard when studying facial attractiveness.

The principle of averageness is not the only objective component that can describe attractiveness, however. For example, Perrett *et al.* demonstrated that a composite of 15 attractive faces is perceived as more attractive than a 60-face composite derived from the

same sample of patients that included the 15 attractive faces.<sup>16</sup> This group went further, and demonstrated that when features of the attractive composites were exaggerated, the faces could be made even more attractive.<sup>16</sup>

The principles of the aesthetic education curriculum for plastic surgeon trainees are limited with regards to ethnicity-specific norms of facial proportions. While plastic surgeons are cognizant of some of the differences common to specific ethnicities, there are no empirically-based guidelines or tools that allow for the rapid pre-planning or post-analysis of facial proportions specific to that ethnicity. The rare discussion of optimal facial proportions in the principal text for plastic surgery education, *Grabb and Smith's Plastic Surgery*, concedes that these proportions are ethnicity-specific, yet the authors go on to provide an extensive discussion based on the neoclassical canons derived from Caucasian females in describing the tenets of nasofacial analysis.<sup>17</sup> Later, the text concedes that the “neoclassical canons describing ideal facial proportions have a limited role in surgical evaluation and planning because they are arbitrary.”<sup>17</sup>

More broadly, society is trending towards an environment where young individuals are increasingly self-conscious about their looks. This so-called “selfie generation” is often fixated on their physical appearance, which can lead to an unhealthy compulsion towards constant modification of physical attributes, including plastic surgery.<sup>18</sup> A recent systematic review on body dysmorphic disorder (BDD) revealed that while its prevalence constitutes 1.9% of adults in the community, 13.2% and 20% of cosmetic surgery and rhinoplasty surgery patients are troubled by this diagnosis, respectively.<sup>19</sup> Unfortunately, this study did not stratify findings by race and ethnicity, but it does highlight the large

proportion of an increasing demographic of patients seeking aesthetic surgery who may benefit from the reassurance of their current aesthetic composition.

One of the most significant challenges in creating a tool to objectively measure beauty rests in the conundrum of the 'universality of beauty.' Previous work has demonstrated that females are typically rated more attractive than males,<sup>20</sup> and female observers rate female photos more harshly than their male counterparts.<sup>20,21</sup> Perhaps even more obvious than gender differences are the societal and cultural influences that effect the perception of attractiveness. In one study, researchers examined the effects of sexual dimorphism cues on perceived facial attractiveness.<sup>12</sup> When stratified by degree of urbanization, they found that more primitive societies preferred more masculinized faces, and that masculinity is far less associated with aggressiveness when compared to the perceptions of more urbanized groups.<sup>12</sup> In another study, researchers found that Black South Africans and White Scottish individuals demonstrated good agreement on what is attractive when viewing white faces, but this relationship is far less significant when viewing black faces.<sup>22</sup> The authors believe that it is the lack of exposure to the other race that led the Scottish Whites to judge black faces more harshly.<sup>22</sup> Thus, it is clear that environment shapes our perception of beauty, and therefore must be taken into account in such an endeavor.

To date, a validated surgical planning and outcomes assessment tool for facial attractiveness is lacking. Meanwhile, other medical disciplines are becoming entrenched in evidence-based methodologies that are either altering or further solidifying their standards of care. Recently, a group of plastic surgeons acknowledged that personal life experiences shape what is considered aesthetically pleasing,<sup>2</sup> and this is a concept that will be explored

in greater detail later in this paper. Ultimately, this same group challenged the plastic surgery community to continuously and critically identify means of objectively quantifying surgical outcomes via four steps (Table 1).<sup>2</sup> However, it cannot be understated, nor should it ever be underestimated, the degree of complexity and difficulty involved in an attempt to create a system that objectively quantifies facial beauty.

Table 1. Essential Steps in Objectively Quantifying Aesthetic Surgical Results<sup>2</sup>

1.	Measure, calibrate, and average specific parameters or data points (usually through photographs).
2.	Define an “ideal.”
3.	Measure and compare the same parameters preoperatively and postoperatively.
4.	Compare the postoperative results to the previously established “ideal.”

## **CHAPTER 1: HISTORICAL TOOLS OF AESTHETIC TRAINING AND ASSESSMENT**

Aesthetics training for physicians is rooted in fundamentals such as the 'golden ratio' or phi ( $\phi$ ), an observed symmetry found ubiquitously in nature, and the neoclassical canons, a series of facial proportions which are widely attributed to the Greek sculptor Polycleitus. Over the past several decades, psychologists and physician-researchers concerned with facial aesthetics have explored the degree to which these concepts actually correlate with the anthropometric facial features of various geographically-based cohorts of individuals.

### **NEOCLASSICAL CANONS**

The neoclassical canons of facial attractiveness are thought to first be described by the Greek sculptor Polycleitus (circa 450 to 420 BC) who employed Egyptian principles to identify 11 facial proportions that constitute the aesthetic ideal (Table 2).<sup>23</sup> These proportions were often used as guides for artist to follow in creating human likenesses, much like Polycleitus did in his creation of the famous statue Doryphorous.<sup>23</sup> The concept of the neoclassical canons is also documented by various Renaissance artists including Alberti, Francesca, Pacioli, and da Vinci and was propagated by artists who specialized in documenting and creating anatomical renditions of the human body from the seventeenth to nineteenth centuries.<sup>23</sup>



Table 2. The Neoclassical Canons<sup>23</sup>

1.	The head can be divided into equal halves at a horizontal line through the eyes.
2.	The face can be divided into equal thirds, with the nose occupying the middle third.
3.	The head can be divided into equal quarters, with the middle quarters being the forehead and nose, respectively.
4.	The length of the ear is equal to the length of the nose.
5.	The distance between the eyes is equal to the width of the nose.
6.	The distance between the eyes is equal to the width of each eye (the face width here can be thus divided into equal fifths).
7.	The width of the mouth is one and one-half times the width of the nose.
8.	The width of the nose is one-fourth the width of the face.
9.	The nasal bridge inclination is the same as the ear inclination.
10.	The lower face can be divided into equal thirds.
11.	The lower face can be divided into equal quarters.

It wasn't until Dr. Leslie Farkas, a plastic surgeon-scientist at the Hospital for Sick Children, began to publish his findings on craniofacial morphology in the 1970s, that the medical community began to recognize that multiple components of the neoclassical canons differ greatly among individuals of none white-European decent.<sup>24</sup> Since that time, the number of studies that have attempted to correlate the neoclassical canons to the average measurements of various ethnic cohorts has increased significantly.<sup>25-34</sup>

### **PHI OR 'THE GOLDEN RATIO'**

Phi ( $\varphi$ ) is the ratio obtained when a line ABC is cut such that  $AB/AC = BC/AB$  or roughly  $\varphi = 1.618$ .<sup>23</sup> The earliest work linking the golden ratio to aesthetics was a mathematics book written by Luca Pacioli and illustrated by Leonardo da Vinci in 1509.<sup>35</sup> *De Divina Proportione*, or "on the divine proportion," focused on the golden ratio's application to geometry, architecture, visual art and the ideal human aesthetic. Surprisingly, modern-day aesthetic education has not strayed significantly from this original line of thinking. For instance, *High Definition Body Sculpting*, a well-respected text

edited by Drs. Alfredo Hoyos and Peter Prendergast, goes into extensive detail demonstrating the various occurrences of phi found throughout the human form.<sup>36</sup>

Researchers have gone to great lengths to validate phi as a key mathematical feature of facial attractiveness but have only done so in limited cohorts.<sup>20,37</sup> This sampling problem is a theme that seemingly recurs in all studies that have validated phi or the neoclassical canons as key components of facial attractiveness. This notion was made increasingly evident in a systematic review published by Fang *et al.* where they examined the racial and ethnic differences in face measurements. This group discovered significant proportional differences in the measurements of forehead height, as well as measurements of the eyes, nose, and mouth among various ethnicities.<sup>24</sup>

## CHAPTER 2: VALIDATION OF HISTORICAL TOOLS IN AESTHETICS

Recently, plastic surgeon Dr. Stephen Marquardt developed a digital device designed to function as a surgical planning tool to improve facial attractiveness. This 'phi mask' is made up of various polygonal shapes derived from  $\varphi$  and when applied to a two-dimensional (2D) image of a face, can be used to modify key features that result in a perceived improvement of the overall facial aesthetic.<sup>38</sup> The 'phi mask' has undergone a significant amount of criticism for a number of reasons, most importantly in that Dr. Marquardt based the anchor points of his mask on images of white female models of European decent, and therefore, the mask doesn't necessarily apply to the male sex or individuals of other ethnicities.<sup>39,40</sup> Nonetheless, tremendous research has examined the role that the neoclassical canons and phi contribute to the aesthetic ideal.

### REVIEW OF THE LITERATURE

Dr. Mounir Bashour, an oculoplastic surgeon from Toronto, Canada, has performed the most extensive study to validate the presence of phi in a population to date.<sup>20</sup> His research will be reviewed in great detail later in the text, but there are other studies that are worth review.

Pallett *et al.* used 2D photographs of female faces and altered various horizontal and vertical proportions using photo editing software. Particularly, they altered features like eye-to-mouth distance and the horizontal distance between the eyes, and then queried college student raters to determine the ratios that resulted in the most attractive proportions.<sup>37</sup> They determined that a vertical distance between the eyes and mouth that is 36% the length of the face, and a horizontal distance between the eyes that is 46% the

width of the face, are the proportions that represent the ideal ratio of these features with regards to perceived facial attractiveness. The authors found these ratios to be significantly different than phi;<sup>37</sup> however, further analysis determined that careful measurement reveals the presence of phi using these new ratios.<sup>41</sup>

In another study, a group performed anthropometric analysis of the relation of the malar prominence and its vertical relation to other facial features.<sup>42</sup> A convenience sample of images from 67 patients being evaluated for rhinoplasty were acquired from an otolaryngology practice and measurements were performed to determine the relationship of the malar prominence to the lateral canthi, chin, and pupils. Among a cohort that was fairly evenly split between males and females, they found no significant differences in measurements among the various ethnic groups that was mostly comprised of Caucasians (43%) and those with race unspecified (30%). Interestingly, the average chin-to-malar prominence distance compared to the chin-to-eye canthus distance approximated the golden ratio.<sup>42</sup>

Anand *et al.* prospectively collected the photos of 50 females and 50 males to determine the prevalence of the golden ratio in a cohort of individuals from Northern India.<sup>43</sup> The horizontal measurements include the intercanthal, interalae, and intertemporal distances. The vertical measurement points included the stomion of the lips, pupils, and chin. Not surprisingly, the researchers found no correlation between the proportions measured and phi. This relationship held true for both males and females.<sup>43</sup> Another study compared the measurements of photographs of attractive Indian-American women to Caucasian women and uncovered significantly different proportions between the two groups.<sup>27</sup>

In a more unique approach, a group analyzed the facial proportions of contestants who participated in the 2012 Miss Korea contest.<sup>44</sup> This population, which was considered attractive, was compared to a group of nursing students of similar age. Three-dimensional (3D) images were prospectively collected and analyzed for typical anthropometric facial proportions. The researchers found significant differences in the measurements between the two groups, but perhaps more interestingly, the golden ratio was not found in either group.<sup>44</sup> While the body of literature exploring the presence of phi in various populations may be lacking, there are far more studies that seek the presence of the neoclassical canons in various cohorts.

A group from the University of Michigan performed a fairly rigorous systematic review on research papers focused on facial anthropometry and ethnicity. This group was able to collect 11 measurements for 27 ethnic groups including African-Americans, Bulgarians, Singaporean Chinese, and Zulu, to name a few.<sup>24</sup> The quotient for each facial proportion was calculated to yield a coefficient of variation of each ethnic group. Interestingly, the researchers found a set of anthropometric measures that were extremely similar between groups, as well as a set that differed. In particular, it was revealed that zygion-zygion, exocanthion-exocanthion, and gonion-gonion measurements demonstrated the lowest level of variation among groups. In contrast, forehead height and endocanthion-endocanthion measurements showed the greatest level of variability. Other measurements such as subnasale-gnathion, cheilion-cheilion, and alare-alaré also demonstrated observed differences among the ethnic groups, but to varying degrees.<sup>24</sup> The authors came to the conclusion that efforts need to be made to collect the measurements that represent the “vast ethnic spectrum” and additional studies to corroborate this disparity.

There exist multiple studies that have analyzed the presence of the neoclassical canons in the Chinese. Wang *et al.* queried the presence of four neoclassical canons in a large subset of individuals from the Han ethnic group and compared these findings to young North American Caucasian adults.<sup>33</sup> These researchers uncovered that the Han sometimes exhibit features that are more in line with the neoclassical canons than North American Caucasians, while there are other features that are less prevalent. This study did not use averaged features to determine overall deviation, but it exhibits the wide variability found within and across two distinct ethnic groups.<sup>33</sup> In a similar study, a group examined 3D photos from a large cohort from Southern China (Hong Kong). Again, the authors did not average the photos of the 51 male or 52 female subjects, but similar to the previously mentioned study, there was little to no association with the neoclassical canons.<sup>28</sup>

A study by Zhao *et al.* went a step further to examine the presence of neoclassical canons in an attractive female cohort of Han Chinese. The researchers classified 450 two-dimensional face photographs based on eight face types and had aesthetic 'experts' rate them based on level of attractiveness. The faces deemed attractive were averaged, and their measurements subsequently compared to the neoclassical canons. They found that as the temporal width and pogonion-gonion distance increased, so did attractiveness. They also found an inverse relationship between bizygomatic and bigonial widths and attractiveness. Most notably, the researchers did not find the presence of the neoclassical canons in what is the world's largest ethnic group.<sup>34</sup> In a similar study, a research group collected 3D photos of Han Chinese women from Shanghai, averaged them, and compared their measurements to that of French Caucasian women.<sup>45</sup> Not only did they find significant differences between the measurements of the two groups, they also corroborated the

findings that the average anthropometric measurements of the Han Chinese do not mirror the neoclassical canons.<sup>45</sup>

Porter *et al.* decided to explore the anthropometric measures of African-American (AA) females and males and compare these findings to their Caucasian counterparts, as well as assess the presence of neoclassical canons in a series of papers.<sup>31,32</sup> Two-dimensional photos of young AA males and females were prospectively collected and the measurements of the individual photos were compared to the neoclassical canons as well as the average measurements of Caucasians photos. Not surprisingly, this group found a number of differences in facial proportions and measurements. Most notably, AA women have longer foreheads, shorter nose length, lengthened lower face height, and wider measurements for most horizontal measures.<sup>32</sup> African-American men tend to have shorter nose length, wider alar width, shorter nasal tip protrusion, shorter columella, a less inclined nasal bridge, and more acute nasolabial angle when compared to Caucasian men. In general, the neoclassical canons were mostly absent in the African-American men and women included in these studies.<sup>31,32</sup>

The absence of neoclassical findings is widespread in the anthropometric literature of various ethnicities. Al-Sebaei revealed that the neoclassical canons were not found in a large cohort of Saudi Arabian dental students.<sup>25</sup> Saad *et al.* examined the presence of nasofacial canons among ethnic beauty pageant winners and found that the width of the eye position and mouth size were larger, and nose smaller in Eastern Mediterranean and European ethnic groups compared to others.<sup>46</sup> Kusugal used a 3D scanner to study the measures of Indian and Malaysian women, and found that the orbital, orbitonasal, and nasoaural canons were far more prevalent in Malaysian women when compared to Indian

women.<sup>30</sup> And finally, a recent study revealed that the neoclassical canons were present in no more than 33% of the large cohort of Turkish males or females.<sup>29</sup>

The absence of neoclassical canons isn't limited to the various ethnic minority groups studied. Pavlic *et al.* evaluated the presence of neoclassical canons in adolescent and young Croatian (Caucasian) adults, and found them to be present in 55 to 65% of the cohort.<sup>47</sup> These researchers also found that the deviation most concerning to the subjects, on a psychosocial level, was that of the vertical facial proportions.<sup>47</sup> Similarly, Cvicelova *et al.* found that most of the canons were absent or weakly correlated with a group of adolescent and college-aged Caucasian males.<sup>26</sup> The consistent and pervasive findings that phi and the neoclassical canons, tools that are still used today to teach aesthetic concepts, don't necessarily correlate with the vast anthropometric measures of the face nor correlate with attractiveness throughout the world and suggest that the discipline would benefit from a new approach to objectively quantifying the aesthetic ideal.

## **MODEL PAPERS**

Bashour, an oculoplastic surgeon in Toronto, Canada, took the first major step in performing an extensive analysis of phi and its role in facial attractiveness measures using multiple linear regression.<sup>20</sup> He studied the validity of a mathematical model system, the 'phi mask,' on a cohort of college students at his institution. A few years later, Schmid *et al.* developed an index of facial attractiveness based on the neoclassical canons, symmetry and the golden ratio.<sup>21,48</sup> Here, I present a detailed analysis of these two papers that serve as the foundation for my proposed statistical model for quantifying facial attractiveness. This tool will be integrated into a complex methodology that will allow for the objective



determination of attractiveness among various ethnic groups and determine the level to which attractiveness conforms to the neoclassical canons and phi.

#### *ANALYSIS OF PHOTOGRAPHED SUBJECTS*

Bashour recruited 37 male and 35 female students from the University of Toronto with an age range of 18 to 30 years. The males were  $22.8 \pm 3.27$  years old and the females were  $21.2 \pm 2.92$  years old. All subjects were of “white European extraction.” Photos were taken in high resolution using a Kodak DCS-560 digital camera. The photographs were standardized as follows: the subjects were seated with standardized lighting against a common background, all faces were kept at their respective relative sizes and not normalized, and the images were cropped to reduce visibility of all hair, ears, and neck. Additionally, each subject’s appearance was standardized as much as possible with the following: the subjects wore no makeup or adornments, male subjects were clean shaven, hair was worn off the forehead, and finally, photos were captured with the head in a standardized position with neutral facial expressions and a closed mouth.<sup>20</sup>

After the photos were taken, a series of composites were created. Major facial features were manually marked with 224 predefined feature points. A series of composites for each gender were then generated: sixteen 2-face (Av2), eight 4-face (Av4), four 8-face (Av8), two 16 face (Av16) and one 32-face (Av32) were created to bring the total count of male faces to 68 and female faces to 66.<sup>20</sup>

Schmid *et al.* gathered the majority of photos used in their study from the Facial Recognition Technology (FERET) database, a repository of photos the U.S. federal government uses to sponsor the development of facial recognition algorithms. From this

database, they selected images that only included full frontal views of Caucasian faces with little or no facial expression. This group collected 420 unique images (210 male, 210 female) who were presumed to be of average attractiveness. No data was provided about the age of the subjects as this was likely impossible to collect. Additionally, this group gathered 32 images of popular movie personalities ranging from the 1930s until the time of the study. This cohort, which included 15 males and 15 females, was selected because they were assumed to be more attractive than the normal population.<sup>21</sup>

Both Bashour and Schmid *et al.* employed very little, if any, sampling controls to avoid selection bias or ensure that they selected subjects that are representative of the general population. Bashour's cohort was young (18-30 years old) and of white, European decent, therefore, it should be inferred that his statistical model should only apply to individuals with similar characteristics. Schmid *et al.* were inherently limited in their ability to select images representative of a larger population since the FERET images don't provide any specific demographic information related to age and ethnicity. Additionally, their judgment of who is of average attractiveness versus above average attractiveness was based on the subjective judgement of the researchers, with the assumption that the 'famous' individuals they selected were of above average attractiveness compared to the general population. In this regard, Bashour's study was superior in that he employed a sort of gold standard with the 'phi mask' for which he could normalize the deviation from an ideal.<sup>20,21</sup>

#### *ANALYSIS OF RATERS OF ATTRACTIVENESS*

Bashour utilized two rating systems for his study. In the Survey Arm, he recruited students at the University of Toronto as well as patients at his ophthalmology clinic, Lasik M.D., to rate 2D images that he collected. These raters included 25 males and 25 females, with a mean age of  $25.8 \pm 10.8$  years and a range of 10-52 years. In the Internet Arm, Bashour recruited random internet users to rate faces. Similar to the survey arm, there were 25 males and 25 females raters, with a mean age of  $21.6 \pm 9.8$  years and a range of 10-52 years. When examining the groups combined, there were 100 total raters (50 male, 50 female), with a mean age of  $21.6 \pm 9.8$  years, and a range of 10-52 years.<sup>20</sup>

The raters were asked to perform the ratings on a computer using a 1 to 10 scale where 1 reflects the least attractive, and 10 reflects the most attractive. This rating interface was also ported to the web for random internet users. Finally, attractiveness quotients were calculated by finding the means of the rating scores for each photo that were stratified by the different survey groups. The author does not provide any information about the race or ethnicity of the raters, but based on the subjects recruited for the photographs, an assumption can be made that the majority of the raters were of white, European decent.<sup>20</sup>

Bashour performed a Cronbach's coefficient alpha analysis to measure the reliability of the attractiveness ratings. He found excellent reliability for all raters and sexes. For example, in the survey arm, the coefficient was 0.97 for male faces, 0.96 for female faces and 0.98 when the gender of the face images was combined into a single group. When the raters were stratified by gender, the coefficient was found to be 0.97 for male raters and 0.96 for female raters. Similarly, Bashour found good reliability among the internet arm. The Cronbach's alpha coefficient was found to be 0.91 for male faces, 0.96 for female faces,

and 0.85 when the faces were combined. Male raters were slightly less reliable than female raters (0.90 versus 0.95, respectively).<sup>20</sup>

He then took the mean ratings for each face and calculated the individual versus combined scores of the different arms to create attractiveness quotients. The mean attractiveness scores can be found in table 3. On average, female faces were rated higher than male faces in both the internet and survey arms. Additionally, raters in the survey arm rated pictures higher than those in the internet arm, and the author attributes this difference to the anonymity that is tied to the internet.<sup>20</sup>

Table 3. Attractiveness Quotients (Bashour)<sup>20</sup>

	Attractiveness Quotient (SD)	T-statistic (p-value)	Linear regression	
			r	F statistic (p-value)
<b>Combined faces</b>		15.39 (p < 0.0001)	0.91	626.56 (< 0.0001)
<b>Internet</b>	4.05 (± 1.08)			
<b>Survey</b>	4.76 (± 1.27)			
<b>Female faces</b>		13.70 (p < 0.0001)	0.91	302 (< 0.0001)
<b>Internet</b>	4.30 (± 1.22)			
<b>Survey</b>	4.80 (± 1.36)			
<b>Male faces</b>		9.59 (p < 0.0001)	0.95	631 (< 0.0001)
<b>Internet</b>	3.80 (± 0.86)			
<b>Survey</b>	4.73 (± 1.20)			

Schmid *et al.* also recruited students and employees from their home institution, the University of Nebraska-Lincoln. The raters included 18 males and 18 females. Their average age was not provided, but the range was 19 – 61 years. The subjects were asked to rate the attractiveness on a 1 to 10 scale similar to that delivered by Bashour. Additionally, this group recorded the amount of time taken to rate each face. An additional factor this group wanted to consider was the raters' opinion of self-attractiveness, also assessed by the 10-point scale. The authors do not provide any information about the race or ethnicity

of the raters, but based on the demographic profile of the institution, it can be inferred that the majority of the raters were white.<sup>21</sup>

Schmid *et al.* employed a ‘partially balanced incomplete block design’ whereby the images presented to the raters were split into 6 groups of 70 images each (35 male, 35 female). Each group also had 30 duplicate images (15 male, 15 female) for a total of 100 images per group. Each rater was asked to rate two of the groups. Duplicate images were included to check for interrater consistency. Raters also rated 32 faces of ‘famous’ individuals (16 male, 16 female) for a total of 232 images per rater.<sup>21</sup>

In contrast to Bashour, Schmid *et al.* conducted a detailed analysis to answer the question of whether men and women rate faces differently. They performed a one-way analysis of variance (ANOVA) where the average rating of each image served as the dependent variable. The ratings of duplicate images were averaged for each rater. The following statistical model was used:

$$ARijkl = Si + P(S)ij + Gk + I(G)kl + (S * G)ik + eijkl, i = 1, 2, j = 1, \dots, 18, k = 1, 2, l = 1, \dots, 116$$

where  $S$  is the effect due to the gender of participant,  $P(S)$  is the random effect due to participant,  $G$  is the effect of the image gender,  $I(G)$  is the random effect due to image,  $S * G$  is the interaction effect due to the gender of the participant and gender of the image and  $e$  is the residual error. The authors observed a “slight” difference in how men and women rated faces, however, this assessment may be considered an understatement as the difference was approaching statistical significance. In general, the authors found that males rate female faces higher ( $p = 0.0571$ ), and that female faces are rated significantly higher than male faces by both male and female raters ( $p = 0.0004$ ).<sup>21</sup>

These researchers determined there was agreement between male and females when rating known 'famous' faces, but this agreement was lost when rating unknown faces. Interestingly, they found that raters rated 'famous' males and females equally, whereas in the unknown cohort, females were rated higher.<sup>21</sup>

Next, this group wanted to answer the question of whether male and female raters exhibit the same variability when rating faces. They compiled a dataset of faces that were rated twice by the same rater (120 ratings per participant) and computed the variance for each rater and each face gender. The variance in ratings of the same face was compounded over the 30 sets of duplicate faces resulting in two variances per rater (one for male faces, one for female faces). Upon plotting the variances, a lognormal distribution was found, therefore a procedure known as GLIMMIX in Statistical Analysis System (SAS) software was performed. They found that male and female rater variability did not differ depending on the gender of the face being judged. Females demonstrated slightly higher variability ( $\sigma^2_M = 0.8318$ ) in their ratings when compared to males ( $\sigma^2_M = 0.8318$ ), but this was not statistically significant ( $p = 0.1658$ ).<sup>21</sup>

Interestingly, this group found that there was a positive correlation between self-rating and the average rating of others (intercept= 2.898,  $b = 0.38$ ,  $p = 0.0041$ ,  $R^2 = 0.3437$ ), and this was true for both sexes. However, males tended to rate themselves higher than females, but this finding was not statistically significant. They also investigated the relationship of the speed of rating on the rating score, which was found to be divergent between the genders ( $p = 0.0016$ ). For instance, for each additional second a female spent rating an image, the rating decreased by 0.0135 points ( $p = 0.3194$ ). Conversely, for each

additional second a male spent rating an image, the rating increased by 0.0408 points ( $p=0.0072$ ). This trend did not depend on the gender of the face in the image.<sup>21</sup>

Overall, both studies provide validation for their method of rating attractiveness. Both studies recruited raters to judge photos on a 10-point scale. Bashour used Cronbach's alpha analysis to demonstrate high reliability among survey and internet raters, with little difference observed when stratifying raters based on gender. Bashour also demonstrated that while the survey arm was superior in terms of reliability, the internet arm exhibits sufficient reliability, thus introducing a recruitment tool that harnesses the large reach of the internet to gather reliable attractiveness ratings. One major limitation of this approach, however, is the inability to verify the age, sex and other demographic details of the raters.<sup>20</sup>

Schmid *et al.* performed a more detailed analysis of raters that examined rater reliability by exposing raters to duplicated images. They found no statistically significant differences between the variance of male and female raters. They also uncovered that female faces are rated higher than male faces, which is consistent with the findings of Bashour. It is probably not a surprise that raters who rate themselves highly, also rate others highly, but the finding that time spent to rate a photo yields an opposite effect for males versus females indicates an intriguing question related to gender and perception.<sup>21</sup>

#### *ANALYSIS OF STATISTICAL METHDOLOGIES*

Bashour employed the 'phi mask' as a template for ideal facial attractiveness. This mask contains geometric figures based on the 'golden ratio' and is inherently symmetrical, thus enabling the assessment of two components thought to contribute to facial attractiveness. The mask was sized and placed using only the interpupillary distance as reference line.

Thirty-seven nodal points on each facial image were selected that were also present on the mask. Distances were measured by hand and using computer software (Image Pro Plus 4.5). The deviation of mask to face nodal points was measured in centimeters and from this ratio, the researchers calculated a mask deviation score. Additionally, Bashour performed multiple linear regressions on the various nodal points to determine a weightings system.<sup>20</sup>

A linear regression was performed to assess the relationship between the number of faces in the composite photos and the attractiveness score. Table 3 indicates a strong 'goodness of fit' and high correlation between the *internet* and *survey arms*. The mean attractiveness quotients for individual versus composite faces were also calculated using ANOVA. This analysis revealed a significant effect of the number of faces:  $F(5,67) = 17.78$  (internet), 25.28 (survey), and 24.61 (combined), and this relationship was statistically significant. ( $p < 0.0001$ ). The relationship comparing composite faces versus individuals held true when just comparing female faces [ $F(5,65) = 15.20$  (internet), 19.41 (survey), and 18.03 (combined),  $p < 0.0001$ ], male faces [ $F(5,67) = 17.78$  (internet), 25.28 (survey), and 24.61 (combined),  $p < 0.0001$ ], as well as faces when the genders were combined [ $F(5,65) = 16.34$  (internet), 45.36 (survey), and 21.63 (combined),  $p < 0.0001$ ].<sup>20</sup>

Next, Bashour performed an ANOVA to compare the variance of mask deviation score with gender and the number of composites as covariates. Analysis revealed that there was a significant effect by gender (male faces deviated further from the mask) as well as number of faces in the composites (the greater the number of faces, the greater the deviation from the mask),  $F = 13.84$  and  $F = 7.98$ , respectively ( $p < 0.0001$ ).<sup>20</sup>

Bashour then performed Pearson's correlations to determine the relationship between mask deviation score and the attractiveness quotients, and found that for



combined faces, male faces, and female faces, there was a negative and significant correlation that ranged between -0.46 and -0.51, ( $p < 0.0001$ ). This finding demonstrates that faces that more closely resembled the measurements of the phi mask were deemed as more attractive.<sup>20</sup>

Finally, Bashour performed multiple linear regression analysis of deviations from the nodal points selected as weighted by the regression analysis discussed in the first section of this paper. Nodal point deviations were selected as the independent variables, and attractiveness quotient served as the dependent variables (table 2).<sup>20</sup>

Table 4. Multiple linear regression of nodal point versus attractiveness quotient (Bashour)<sup>20</sup>

	<b>R</b>	<b>R<sup>2</sup></b>	<b>R<sup>2</sup> adjusted</b>	<b>F</b>	<b>p-value</b>
<b>On both faces</b>					
<b>Internet</b>	0.74	0.55	0.39	2.44	< 0.0001
<b>Survey</b>	0.73	0.53	0.36	3.28	< 0.0001
<b>Combined</b>	0.74	0.55	0.38	2.76	< 0.0001
<b>On female faces</b>					
<b>Internet</b>	0.93	0.87	0.71	5.59	< 0.0001
<b>Survey</b>	0.93	0.87	0.72	5.75	< 0.0001
<b>Combined</b>	0.94	0.88	0.74	6.35	< 0.0001
<b>On male faces</b>					
<b>Internet</b>	0.78	0.60	0.16	1.38	0.1825
<b>Survey</b>	0.79	0.63	0.22	1.54	1.087
<b>Combined</b>	0.79	0.62	0.21	1.50	0.1265

Schmid *et al.* used a graphical user interface to manually identify 29 landmarks in each image selected. They then created geometric model to reduce these various measurements into a score,  $F_i$ :

$$F_i = \{f_{i1}, f_{i2}, \dots, f_{im}\}$$

where each feature point,  $f_{ij} = (x_{ij}, y_{ij})$ ,  $1 \leq i \leq n$ ,  $1 \leq j \leq m$ , is represented by its 2D spatial coordinates in the face. Ultimately, their goal was to determine a function  $A$  that maps a face to an attractiveness score.<sup>21</sup>

$$A(Fi) \rightarrow [1,10]$$

This group used three predictors to compute facial attractiveness: the neoclassical canons, facial symmetry and the golden ratios. A number of neoclassical canons have been described in the literature (table 2),<sup>23</sup> but Schmid *et al.* were only able to test 6 based on the 29 facial landmarks chosen from Dr. Farkas' previous work (table 5). In order to ensure the canon ratios adhered to the number of features, the authors used the coefficient of variation which is the ratio of the standard deviation of the distances to the mean of the distances. This enabled them to incorporate all the distances into one value. Therefore, a value of zero indicates no variation in the distances, and, the larger the value, the more the face differs from the canon.<sup>21</sup>

Table 5. Neoclassical canons assessed from nodal points (Schmid *et al.*)<sup>21</sup>

Formula number	Description
2	Forehead height = nose length = lower face height
4	Nose length = ear length
5	Interocular distance = nose width
6	Interocular distance = right or left eye fissure width
7	Mouth width = 1.5 × nose width
8	Face width = 4 × nose width

Based on their analysis, Schmid *et al.* discovered that five canons had a significant relationship with attractiveness. These relationships include (1) forehead height equals nose length equals lower face height; (2) nose length equals ear length; (3) interocular distance equals nose width; (4) interocular distance equals right or left eye fissure width; and (5) face width equals four times the nose width. As the measurements deviated from the above-mentioned canons (increase in coefficient of variation), the attractiveness scores decreased significantly ( $p < 0.006$ ). Interestingly, they found that as the coefficient of variation increased for females, so did the attractiveness scores (the opposite finding was

found for males,  $p = 0.0028$ ), suggesting that females are viewed more attractive when they have smaller noses and/or a larger distance between the eyes, which runs contrary to notion that universal adherence to the canons represents the aesthetic ideal.<sup>21</sup>

To measure symmetry, the group fitted the least squares regression line through the seven points 1, 3, 19, 23, 26, 28 and 29, and assessed the measurements of the following pairs: (1) eyebrows (Points 2 and 4; Points 7 and 8); (2) eyes (Points 11 and 14; Points 12 and 13; Points 15 and 16); (3) nose (Points 18 and 20); (4) ears (Points 5 and 10; Points 17 and 21); (5) lips (Points 22 and 24; Points 25 and 27); (6) face (Points 6 and 9). To determine the symmetry of the face, a formula was created to give overall facial symmetry measures (FSMs). The equations are as follows:

$$(1) \text{ Difference: } FSM_{Diff}(d) = d_L - d_R$$

$$(2) \text{ Ratio: } FSM_{Ratio}(d) = d_L \div d_R$$

$$(3) \text{ LN(Ratio): } FSM_{LNRatio}(d) = \ln(d_L \div d_R)$$

$$(4) \text{ Adjusted Difference: } FSM_{AdjDiff}(d) = (d_L - d_R) / [(d_L + d_R) / 2]$$

For equations 1, 3 and 4, a finding of zero implies symmetry; for equation 2, a finding of one implies symmetry.<sup>21</sup>

This group found that the difference symmetry measure (equation 4), which measures the difference in distances from the symmetric points to the line of symmetry, is the strongest symmetry measure associated with attractiveness. To determine the contribution of symmetry pairs towards the attractiveness of the face, the group performed a stepwise regression to reduce the number of variables in the model. They found that the symmetry of the nose (points 18 & 20) and mouth (points 25 & 27) were important for

attractiveness scores of males faces ( $p = 0.0025$ ), but not female faces ( $p = 0.604$ ). They also determined that the symmetry of the upper tips of the lips was also important.<sup>21</sup>

Next, Schmid *et al.* were able to use 17 ratios derived from the 29 nodal points they initially selected to test the hypothesis that the presence of golden ratios is associated with facial attractiveness (table 6). Analysis revealed that only 6 of the 17 ratios tested were predictors of attractiveness: ear length to nose width, mouth width to interocular distance, lips-chin distance to interocular distance, lips-chin distance to nose width, length of face to width of face, and nose width to nose-mouth distance.<sup>21</sup>

Table 6. Golden ratios analyzed (Schmid *et al.*)<sup>21</sup>

Ratio Number	Numerator points	Denominator points	Description
1	y10-y21	x12-x13	Ear length to interocular distance
2	y10-y21	x18-x20	Ear length to nose width
3	x15-x16	x12-x13	Mideye distance to interocular distance
4	x15-x16	x18-x20	Mideye distance to nose width
5	x25-x27	x12-x13	Mouth width to interocular distance
6	y23-y29	x12-x13	Lips-chin distance to interocular distance
7	y23-y29	x18-x20	Lips-chin distance to nose width
8	x12-x13	x12-x11	Interocular distance to eye fissure width
9	x12-x13	y23-y28	Interocular distance to lip height
10	x18-x20	x12-x11	Nose width to eye fissure width
11	x18-x20	y23-y28	Nose width to lip height
12	x18-x20	y19-y26	Eye fissure width to nose-mouth distance
13	x12-x11	y19-y26	Lip height to nose-mouth distance
14	y23-y28	x17-x21	Length of face to width of face
15	y1-y29	y26-y29	Nose-chin distance to lip-chin distance
16	x18-x20	y19-y26	Nose width to nose-mouth distance
17	x25-x27	x18-x20	Nose width to nose-mouth distance

Interestingly, they found that ratings given by females decreased more so than ratings given by males when ratio 2, ear length to nose width, deviates from the golden ratio ( $p = 0.004$ ). This relationship held constant for ratios 7 (lips-chin distance to nose width,  $p = 0.0151$ ) and 17 (nose width to nose-mouth distance,  $p = 0.003$ ). Concordance with the

golden ratio was found with ratios 5 (mouth width to interocular distance,  $p = 0.002$ ), 6 (lip to chin distance to interocular distance,  $p < 0.0001$ ), and 14 (length of face to width of face,  $p = 0.0077$ ). Additionally, images of females that deviate from the golden ratio are rated more harshly by female than male raters. This relationship does not hold for male images, which are viewed equally by the two genders.<sup>21</sup>

Finally, Schmid *et al.* performed a stepwise linear regression, combining all analyzed factors, to determine their contribution to facial attractiveness. Among the neoclassical canons, symmetry, and ‘golden ratio,’ this group compiled a model that included 78 variables: 6 canons, 55 symmetries and 17 golden ratios, which yielded an  $R^2$  of 0.2433. The stepwise linear regression, which was implemented to minimize variance, resulted in a reduced model of 16 predictor variables, which yielded an  $R^2$  of 0.1923. Due to earlier analysis that male and female raters judge images differently, Schmid *et al.* created separate models for the different gender combinations, which revealed  $R^2$  values that were greater than their reduced  $R^2$  values found when the model was generally applied (table 7).<sup>21</sup>

Table 7. Models used after stepwise regression and variable selection (Schmid *et al.*)<sup>21</sup>

<b>Rater/face</b>	<b>R<sup>2</sup> (optimized)</b>	<b>R<sup>2</sup> (reduced)</b>	<b>No. variables in the model</b>
Female/female	0.2378	0.2335	8
Female/male	0.2162	0.2097	8
Male/female	0.2106	0.2088	11
Male/male	0.2053	0.2013	10

Table 8 reveals the canons that were selected for each gender pairing based on the previous analysis. Even though each gender combination resulted in a different model to yield an optimized result, there were specific components that were common to each

model that include canon 6 (interocular distance equals right or left eye fissure width), symmetry pair 22-24 (tips of upper lip) and golden ratios 5 (mouth width to interocular distance) and 7 (Lips-chin distance to nose width).<sup>21</sup>

Table 8. Canon formulas, symmetry pairs, and golden ratios in the final models (Schmid *et al.*)<sup>21</sup>

<b>Rater/face</b>	<b>Canon formulas</b>	<b>Symmetry pairs</b>	<b>Ratio nos.</b>
Female/female	6, 8	22-24	5, 6, 7, 14, 17
Female/male	2, 6	7-8, 18-20, 22-24	5, 6, 7
Male/female	2, 4, 5, 6, 8	22-24	2, 5, 7, 14, 17
Male/male	2, 4, 6, 8	18-20, 22-24, 25-27	5, 6, 7

Bashour demonstrated that averaged, or composite photos, are increasingly rated attractive as more images are added to the composites, and that the ‘phi mask’ closely correlates with the most attractive images rated. His weighted multiple linear regression approached demonstrated very good best fit models for female faces, with the lowest adjusted R<sup>2</sup> values in the range of 0.71 – 0.74. These adjusted R<sup>2</sup> values decline fairly significantly, however, when applied to male images (0.16 – 0.21), indicating that the weighted approach of the ‘phi mask’ is not a good fit for male faces. Overall, he demonstrated very good results, validating the phi mask in the population he studied.<sup>20</sup>

Schmid *et al.* took a very nuanced approach to analyzing male and female raters of facial attractiveness and incorporating these findings into their stepwise linear regression. The R<sup>2</sup> values decreased after reduction, but this can partially be explained by the concept that the fewer variables in a model, the lower the expected R<sup>2</sup> value. Still, one might expect a much better ‘goodness of fit,’ when combining so many variables to attempt to predict the factors that constitute facial attractiveness. Nonetheless, they provide a very detailed and sound approach to improving on this model in future studies.<sup>21</sup>

## ANALYSIS

Bashour hypothesized that “an objective quantitative system should be devisable that would have at least the same correlation with measures by various panels of judges (at least an  $r > 0.80$  and preferably an  $r > 0.90$ ).” Accordingly, he felt the need to distinguish the difference between the definition of facial attractiveness, “*The visual properties of a face that are pleasing to the visual sense of an observer,*” and the definition of beauty, “*the assemblage of graces or properties pleasing to the eye, the ear, any or all of the senses, the intellect, the aesthetic faculty, and/or the moral sense.*” Bashour eventually came to the conclusion that his study more precisely measures the *beauty gestalt*, “*full frontal repose static two-dimensional photographic facial attractiveness*. This can be precisely defined as *the time-static visual properties of a face in a photographic two-dimensional frontal repose image that are pleasing to the visual sense of an observer.*”<sup>20</sup>

Bashour has essentially validated the ‘phi mask’ to qualify as an objective system for quantifying facial attractiveness in young males and females of white, European decent. His definition of a *beauty gestalt* demonstrates the need for an improved system of objectively quantifying attractiveness based on the fact that we, as humans, perceive other humans in three dimensions. Additionally, while his study validates the ‘phi mask,’ it only does so for a very select cohort. The finding that his weighted model results in a loss of ‘goodness of fit’ compared to the unweighted model demonstrates that there are other factors that contribute to facial attractiveness for which his model does not account.

Schmid *et al.* took a more holistic approach by combining the components of the neoclassical canons, the ‘golden ratio,’ and symmetry in a stepwise regression approach to

identify which components truly impact the perception of facial attractiveness. While their statistical methodology was very comprehensive, their methodology was based on what I believe is inherently flawed hypotheses, which are that the golden ratio and neoclassical canons, very antiquated ideas, are the true models for attractiveness. This study would have been far more complete if this group started with the assumption that the components of facial attractiveness are unknown, and therefore taken a more *tabula rasa* approach to identifying those factors.



## CHAPTER 3: COMPONENTS OF ATTRACTIVENESS

The foundation for this research proposal came into focus after one of my mentors introduced me to a 1990 article in *Discover Magazine* entitled, “Do you love this face?” This article discusses the components that make up attractiveness according to the leading evolutionary, adolescent and cognitive psychologists and psychobiologists of that time.<sup>10</sup> Ultimately, the experts came to the conclusion that facial attractiveness can be broken down into four key components: averageness, symmetry, youthfulness, and sexual dimorphism, or the tendency towards more feminine or masculine features.<sup>10</sup> It is also important to note that there are other features, such as skin quality, familiarity with facial features, and exceptions, such as extreme or uncommon features which also must be considered.<sup>10</sup>

One of the most common criticisms of the above biological explanation of facial attractiveness is the notion that it is the preferences we form in early childhood and beyond, influenced by the environment in which we exist, that is the primary shaper of what we find attractive. Judith Langlois became convinced, however, that we possess innate ability to perceive attractive faces when she discovered that infants aged 3 to 6 months prefer more attractive faces over less attractive faces.<sup>4</sup> Dr. Gillian Rhodes performed a meta-analysis of facial attractiveness research and came to the conclusion that averageness, symmetry, and sexual dimorphism all, in fact, contribute significantly to our perception of attractiveness.<sup>49</sup> In this review, she also states that for women, femininity is preferred to averageness and that women prefer more masculine features during the fertile period of the menstrual cycle.<sup>49</sup> She does not deny, however, that preferences also influence what we find attractive.

Little *et al.* went further in their review of facial attractiveness research.<sup>50</sup> In addition to those components mentioned above, they also discuss the contributions of factors like skin color and texture, cues to a subject's personality, as well as the differences in face preferences as they are related to one's own hormone levels, fertility, perceived self-attractiveness, familiarity and imprinting, and social learning.<sup>50</sup> At the center of much of the modern attractiveness research is the evolutionary question of whether the innate ability to perceive attractiveness is related to the positive predictive ability of optimal mate selection.<sup>9</sup> Nonetheless, it is clear that the study of facial attractiveness is an incredibly complex topic and I endeavor with this proposal to create a model that may get the scientific community a little bit closer to an all-encompassing explanation of what defines beauty.

### **AVERAGENESS (KOINOPHILIA)**

Dr. Judith Langlois, a pioneering developmental psychologist at the University of Texas, Austin, disrupted the field of attractiveness research with her 1980 publication, "Attractive faces are only average."<sup>15</sup> Langlois and Roggman argue that, prior to this study, attempts at answering the question of 'what defines beauty' were approached with analysis of specific facial feature assessments, with little explanation of why one feature contributed to beauty more than others. They hypothesized that, based on various examples in the space of evolutionary and cognitive psychology, the average, or mean value of attributes, represents the prototype of any given dataset. To test this theory, digitized facial photos were used to create a progressive set of two, four, eight, sixteen and thirty-two individuals' photos that were each morphed into composite individual photos. Raters were then asked

to judge the level of attractiveness of the individual and composite photos. Inter-rater reliability was high and for the first time, it was demonstrated that averaged faces were rated statistically significantly higher than individual photos.<sup>15</sup>

After reviewing multiple editorials and commentaries, much of which was critical to their approach and offered alternative explanations for their findings, Langlois *et al.* followed up with a second publication on the topic.<sup>51</sup> The first point they addressed is the confusion around the term “average.” It would seem that the most common perception of this word is what *Merriam-Webster’s* thesaurus lists as secondary synonyms to the word: common, normal, regular, typical, etc.<sup>52</sup> However, it is actually the true definition of the word, “an estimation of an arithmetic mean,”<sup>53</sup> that Langlois *et al.* wanted to emphasize here.<sup>51</sup> To try and reduce further confusion with this concept, I will use the term koinophilia moving forward.

The concept of koinophilia and its role in attractiveness have been frequently studied and validated over the past several decades.<sup>23,54-56</sup> The evolutionary biology explanation for this theory is such that, for any given population, extreme characteristics tend to fall away in favor of average ones.<sup>10</sup> Langlois theorizes that humans are innately cognitive averagers, and by the time a child reaches adolescence, they have seen thousands of faces and already subconsciously formed an average from them which are inherently used for comparison.<sup>10</sup> And in fact, Little *et al.* confirmed that this phenomenon is not specific to one racial group.<sup>57</sup>

One proposed biological explanation for the positive correlation between koinophilia and attractiveness is an idea termed ‘processing fluency.’ Langlois’s group used electroencephalograms and observed reduced work of the posterior N170 region of the

brain when observers viewed more attractive faces. This improved processing fluency theory posits that that which is easier and quicker for the brain to process results in a more rewarding and desirable neural pathway for the brain.<sup>5</sup> Others have used more advanced brain imaging techniques such as functional magnetic resonance imaging (fMRI) to quantify, real-time, the segments of the brain that are activated when attractive and averaged faces are visualized, most commonly, the nucleus accumbens, orbitofrontal and prefrontal cortices.<sup>58,59</sup> Regardless of the biological explanation of attractiveness, it is noteworthy that a significant number of articles reviewed for this proposal employed an averaging technique to create an attractive cohort of images.<sup>20,22,34,44,45,56,60-70</sup>

Not all koinophilia research points towards the common theme that “what is the mathematical average of a population is always most attractive of that population.” Rhodes *et al.* decided to investigate the contribution that perceptual adaptation impacts the perception of facial attractiveness.<sup>71</sup> To test this, they exposed 48 college students to normal and attractive phases during a preadaptation phase, followed by exposure to distorted faces during an adaptation phase. In the subsequent postadaptation rating phase, the highest rated faces trended towards those that were distorted, and these findings were statistically significant.<sup>71</sup>

Perrett *et al.* was one of the first groups to challenge the work of Langlois. They demonstrated that a composite of 15 attractive faces is perceived as more attractive than a 60-face composite derived from the same sample of patients that included the 15 attractive faces.<sup>16</sup> Interestingly, they also demonstrated that when features of the attractive composites were exaggerated, the faces were perceived as even more attractive.<sup>16</sup> Despite

these findings, it is evident that koinophilia explains, at least in part, the psychobiology behind the perception of attractiveness.

## SYMMETRY

Symmetry is one of the most studied concepts with regards to attractiveness research. The oldest *Pubmed* record on the study of facial attractiveness was an odontologic case study on the pathogenesis of facial asymmetry presented at the Proceedings of the Royal Society of Medicine in 1911.<sup>1</sup> For decades, evolutionary biologists have studied the role that physical symmetry in primates has on mate selection.<sup>49</sup> One recurring theory that explains its importance in nature is that it signals genetic quality and is therefore used as a tool for natural selection.<sup>9,49,63</sup> Interestingly, some have posited that symmetry has no influence on the perception of facial attractiveness,<sup>72</sup> while others have argued that it is at least as important as koinophilia in this field of research.<sup>60</sup>

Komori *et al.* wanted to determine the contribution that both koinophilia and symmetry play on facial attractiveness.<sup>60</sup> This group recruited an equal proportion of male and female college students and acquired 2D photographs which were then standardized. They used a special technique to create the mirror-reversed version of each face for comparative analysis. A second cohort was then recruited to rate the different versions of each face. This group used multiple regression analysis and found that both symmetry and koinophilia had positive effects on the ratings of male photos. On the other hand, female faces were only positively impacted by koinophilia and not symmetry.<sup>60</sup>

As previously discussed, Schmid *et al.* explored the importance of symmetry in their study in which they performed a multiple stepwise regression to assess the impact of this

feature, as well as the neoclassical canons and the golden ratios on facial attractiveness.<sup>21</sup> Importantly, their analysis revealed that symmetry had the strongest relationship with attractiveness compared to the other components they tested.<sup>21</sup> In stepwise fashion, they revealed that for both male and female raters, symmetry of the mouth and nose were deemed most important in terms of their impact on attractiveness, followed by the upper tips of the lips.<sup>21</sup>

After performing extensive meta-analyses on the factors that contribute to facial attractiveness, Gillian Rhodes is convinced that symmetry is a significant contributor.<sup>49</sup> Rhodes attributes the results of early studies that found no correlation between attractiveness and symmetry to researchers' methodologies. Initial symmetry research generated symmetrical images by reflecting each hemiface of individual photos about the midsagittal plane, thus creating two chimeras that were prone to actually magnifying imperfections.<sup>49</sup> Conversely, in more modern studies where perfectly symmetrical faces are created by blending mirror-reversed and normal images, faces are perceived as more attractive than the originals.<sup>49</sup>

One such modern study utilized a Procrustes fit procedure where shape data was partitioned into symmetric and asymmetric components that could be reflected within a face to create symmetric and asymmetric versions of that face.<sup>73</sup> These symmetric and asymmetric components were then converted to covariance matrices and subjected to principal component analyses to determine the degree that each component contributed to overall symmetry or asymmetry. Though the sample of raters was small for this study (n = 10), the study was sufficient to detect that there was a significant reduction in attractiveness ratings for faces with fluctuating asymmetries.<sup>73</sup>

Little *et al.* postulated that because both symmetry and sexual dimorphism signal positive heritable traits to potential mates, they should be positively correlated with each other and also trend in the same direction as perceived attractiveness.<sup>61</sup> Images were collected from European university students, the Hadza ethnic group from Tanzania, and free-ranging macaque in Puerto Rico. Symmetry and sexual dimorphism measurements were taken and the images with 15 highest and lowest asymmetry scores were made into composites. Ratings revealed that symmetry was correlated with masculinity for males and femininity for females, and both were positively correlated with attractiveness.

## **SEXUAL DIMORPHISM**

The theory surrounding preferences for sexually dimorphic features in facial attractiveness is almost purely derived from an evolutionary basis. Masculinized features are associated with levels of testosterone, for example.<sup>67</sup> Similarly, estrogen-dependent characteristics are associated with health and fertility in women.<sup>67</sup> Based on these findings, Perrett *et al.* postulated that masculinized and feminized faces would correlate with facial attractiveness in images of males and females, respectively.<sup>67</sup> Feminized faces of females were preferred over average faces, but interestingly, masculinized male faces were negatively correlated with attractiveness while also being associated with the perception of dominance and negative attributions.<sup>67</sup> This is just one example, but demonstrates the seemingly counter-intuitive findings that can arise out of sexual dimorphism research.

Hoss *et al.* sought to determine if there was a direct correlation between masculinized features in males, feminized features in females and sex classification. Due to the conflicting findings in the literature, this group also thought it important to determine

whether this relationship held true independent of facial attractiveness.<sup>74</sup> Interestingly, they found that highly masculine male faces, but not highly feminine female faces were associated improved sex classification. Additionally, attractiveness improved sex classification of images independent of sexual dimorphism.<sup>74</sup>

Rennels, Bronstad and Langlois decided to delve deeper into the inconsistent findings in the literature that suggest adults prefer the images of more feminized over masculinized males.<sup>68</sup> Masculinized and feminized male average faces were presented to male and female raters to determine preference. Using a forced-choice procedure, they found that adults prefer the more feminized male face when the choice was feminized versus masculinized. However, if an average face was added to the selection choice, the masculine face was preferred.<sup>68</sup> These results demonstrate the complexity of this topic, and its weighted contribution to facial attractiveness.

The concept of male dominance is an interesting one that deserves more attention. Perrett *et al.* found that hypermale faces were perceived as dominant, which was associated with negative attributes.<sup>68</sup> Additionally, multiple studies have pointed to a shifting preference of female raters towards more masculinized males as they enter the fertile period of their menstrual cycle.<sup>49</sup> One group looked at women's preference for masculinized males in the context of speed dating. Interestingly, they found that hypermale faces are associated with dominance and were more likely to garner short-term interest and get asked for a second date, but less likely to be perceived as good candidates for long-term interest.<sup>75</sup>

From the evidence, it's abundantly clear that feminized females are perceived as more attractive.<sup>9,49,67</sup> This perhaps explains why the commercial space of cosmetics, a set of



products dedicated to highlighting female features, is so wildly successful. Etcoff *et al.* studied this phenomenon which they termed the ‘extended phenotype,’ and demonstrated that makeup is positively correlated with facial attractiveness.<sup>76</sup>

Importantly, Scott *et al.* examined the effects of exaggerated facial sex characteristics on populations of industrialized and non-industrialized populations.<sup>12</sup> This group studied the preferences of rater cohorts from around the world including more developed areas such as Shanghai, China and Bristol City, United Kingdom, as well as less developed regions like the Kadazan-Dusun from the Sabah region of Malaysia and the Tchimba from the Kunene region of Namibia. The preferences for hypermale and hyperfemale faces that were observed in the urban regions did not hold true for raters who live in less well-developed societies unexposed to Western culture.<sup>12</sup> This interesting finding points to the concept of ‘perceptual narrowing,’ which was previously defined as the “decrease in the discrimination ability between objects to which we are not regularly exposed during critical times of our development.”<sup>22</sup> Exposure to cultural influences, or lack thereof, is a factor that must be considered when creating a model that attempts to identify the aesthetic ideal within a population.

## **OTHER CONSIDERATIONS**

One of the proposed components of facial attractiveness that is extremely difficult to quantify is youthfulness. Compared to the other known factors: koinophilia, sexual dimorphism, and symmetry, the body of literature regarding the impact of youthfulness on facial attractiveness is severely lacking. Anecdotally, there is no question that individuals who seek dermatological and plastic surgery interventions for aesthetic facial

enhancements are usually striving to maintain a youthful look. Botulinum toxin (botox) is an injectable cholinergic antagonist that causes short-term neuromuscular blockade and is primarily used to relax and release the wrinkles in the skin that are associated with aging. Similarly, the facelift procedure is designed to tighten the skin to give a more youthful appearance.

Pocheron *et al.* previously demonstrated that facial contrast, or the relation between the luminance of skin and surrounding facial features, is inversely correlated with age in Caucasian women.<sup>77</sup> Recently, this group went a step further to investigate whether this relationship held true for other ethnicities.<sup>78</sup> As expected, the researchers uncovered that youthfulness is directly correlated with facial contrast. Specifically, this phenomenon was studied in Chinese, Latin American and Black South Africans. Interestingly, when the images were digitally altered for low and high face contrast, the relationship to the perception of youthfulness held constant.<sup>78</sup>

Lambros and Amos have been studying facial aging for over a decade and collecting 3D images of subjects during this period.<sup>70</sup> Their large sample, which is almost exclusively composed of Caucasian females, was averaged to provide tremendous insight into the facial changes associated with aging. Specifically, their research reveals that as Caucasian females age, their lid aperture gets smaller, the lower eyelid rises, the lid-cheek junction migrates, and the columellar base moves posteriorly.<sup>70</sup> In a different study on aging, researchers employed statistical shape analysis to examine a very large sample of photos to reveal that head size is directly correlated with age.<sup>79</sup>

One set of components that cannot be overlooked with regards to their implications in facial attractiveness is that of skin color and texture.<sup>50</sup> Little *et al.* performed an

extensive review on the evolutionary basis of facial attractiveness and surmised that skin texture and color are accurate reflections of skin health, and elements thought to be factored in attractiveness assessment.<sup>50</sup> Some have speculated that carotenoid-based skin color is also associated with health, but more recent research did not observe this relationship.<sup>9</sup>

In one study specifically interested in answering the question about the impact of skin color on facial attractiveness, Vera Cruz examined the preferences of individuals from Mozambique, who are traditionally dark-skinned people. Of the 240 raters in the study, photos of light-skinned individuals were preferred to those who are dark-skinned.<sup>13</sup>

Another group examined the skin color from a cross-cultural race standpoint.<sup>22</sup> White Scottish and Black South African college students were queried to their attractiveness preferences by exposing them to photos of White Scottish and Black South African individuals. Interestingly, White Scottish and Black South Africans raters demonstrated high agreement for White European faces, but Black South Africans rated Black faces much higher than did the White Scottish.<sup>22</sup> Further investigation revealed that the Black South African observers relied on color cues when judging attractiveness, whereas White Scottish observers relied more on shape cues. The authors speculated that the reason for the discrepancy was related to 'perceptual narrowing,' whereby the lack of exposure to Black faces led to a decreased ability to distinguish between more and less attractive Black faces. All of the above-mentioned factors should be considered when trying to conceptualize the gestalt that comprises facial attractiveness.

## **CHAPTER 4: METHODOLOGY TO OBJECTIVELY QUANTIFY FACIAL ATTRACTIVENESS**

### *APPROACH*

Due to the scope of the experimental question, this proposal will employ a multiphase strategy that will begin with hypothesis-generating data and descriptive analysis. Similar to the model papers (Bashour and Schmid *et al.*),<sup>20,21</sup> subjects will be recruited and anthropometric landmarks will be used to generate composite photos as well as to measure deviation scores. Similarly, raters will be recruited and asked to judge the attractiveness of the photos using a standard Likert scale. The key differentiating features of this proposal are as follows: subjects will be recruited specifically based on ethnicities, anthropometric landmarks will be more inclusive to capture more extensive facial proportion data, principal component analysis and machine learning will be employed to determine the weight of proportions and attractiveness components to the overall gestalt, and objective measures including functional magnetic resonance (fMRI) as well as eye tracking tools will be employed to objectively quantify neurological response to the images displayed.

### *HYPOTHESES*

The first null hypothesis is that the composite 3D image of all photos of a particular ethnic group (stratified by age and gender) will not represent the ideal aesthetic for that group. The alternative hypothesis is that the composite 3D image of all photos of a particular ethnic group (stratified by age and gender) will represent the ideal aesthetic of that group.

A second null hypothesis is that the composite 3D image of the photos containing 20% of the most attractive individuals of a particular ethnic group (stratified by age and gender) will not represent the ideal aesthetic for that group. The alternative hypothesis is that the composite 3D image of the photos containing 20% of the most attractive individuals of a particular ethnic group (stratified by age and gender) will represent the ideal aesthetic of that group.

A third null hypothesis is that the Koinophilia Index, sexual dimorphism score, degree of symmetry, and skin tone will not vary amongst the aesthetic ideal composite photos of various ethnic groups. The alternative hypothesis is that the Koinophilia Index, sexual dimorphism score, degree of symmetry, and skin tone will not vary amongst the aesthetic ideal composite photos of various ethnic groups.

A fourth null hypothesis is that proposed components of facial attractiveness examined in this proposal all contribute equally to overall perceived facial attractiveness. The alternative hypothesis is that the proposed components of facial attractiveness examined in this proposal contribute to overall perceived facial attractiveness, but to different degrees or not at all.

The final null hypothesis is that the perception of facial attractiveness is unrelated to the age, gender, and ethnicity of the subject and/or the age, gender, and ethnicity of the observer. The alternative hypothesis is that the perception of facial attractiveness is related to the age, gender, and ethnicity of the subject and/or the age, gender, and ethnicity of the observer.

The goal of this research is to use techniques such as 3D digital photography, anthropometric analysis, averaging, and subjective and objective rating measures to

determine the aesthetic ideal of a particular cohort. Once identified, statistical modeling will be employed to determine the contributions that each component mentioned plays in the gestalt that comprises facial attractiveness. Machine learning will be employed to assist in the complex nature of landmark identification in 3D images, and potentially identify facial proportions that contribute to facial attractiveness not previously described in the literature. Ultimately once complete, this system will be able to identify facial components of individuals and identify areas that contribute to or detract from the perceived attractiveness they desire. Hopefully this will result in a tool that patients and surgeons can use to assist in the planning of surgical and non-surgical facial enhancement, as well as an evidence-based device for the assessment of post-surgical outcomes.

## **Subject Recruitment**

Ideally, the target population is the entire adult population who has never undergone facial surgery, hasn't recently received any cosmetic injectables, or been subject to any significant facial trauma. In reality, the target population is primarily consumers of plastic surgery, which likely varies by ethnicity and geographical location. According to the *2016 Plastic Surgery Statistics Report*, 49% of cosmetic procedures in the United States were performed on individuals age 40 to 54.<sup>80</sup> The vast majority (70%) of individuals were Caucasian, with Hispanics (10%), African-Americans (8%), Asian-American (7%) and other ethnicity (5%) making up the remainder of those receiving cosmetic interventions.<sup>80</sup>

### *PHASE 1*

The initial pilot study, or phase 1 of this proposal, will include the analysis of already-accessible 3D photographs. Dr. Val Lambros is a plastic surgeon in Newport Beach,

California, and is a foremost expert on capturing and analyzing 3D photos. He has been collecting 3D images of subjects he's encountered for over a decade,<sup>70</sup> and currently has a collection of approximately 130 photos of young females (age 20-30) and 120 photos of older females (age  $\geq$  68 years). This convenience sample includes a mostly Caucasian cohort of typical plastic surgery patients in Newport Beach and the surrounding areas of southern California. Initial analysis of these photos will be conducted for a proof-of-concept study to secure funding for the next phase.

## *PHASE 2*

The accessible population for phase 2 of this proposal is limited to individuals in Orange County, California. The primary sampling units are the Aesthetic & Plastic Surgery Institute (APSI) at the University of California, Irvine Medical Center (UCIMC) in Orange, California, UC Irvine Health Pacific Coast Plastic Surgery (PCPS) located in Costa Mesa, California where subjects will be recruited who are seeking plastic and reconstructive surgery consultation. Undergraduate and graduate students, as well as faculty and staff may also be recruited from the main campus of the University of California, Irvine.

The sampling frame is based on demographic data obtained over the past several years. The majority of patients who are treated at the APSI and PCPS are Caucasian females, therefore UC Irvine students, faculty and staff will be recruited to increase the sampling rate. Financial compensation will likely be required to incentivize participation. Currently, the clinics do not stratify Asian race into subgroups, so efforts will be made to collect this information prospectively moving forward.

From our accessible population, we will employ probability sampling whereby proportionate stratified sampling will be used to recruit an equal number of Caucasian and

a second ethnicity. Because phase 2 of this study is for hypothesis-generating data, analysis will be initiated once an adequate sample of two ethnicities is reached. We will not discriminate with regards to who we enroll in the study, however, so that we can begin to build a database of photos of individuals from all ethnicities encountered.

The only inclusion criterion is that the patient be of adult age ( $\geq 18$ ). The exclusion criteria include any previous facial plastic surgery, use of any cosmetic fillers within the past two years, botox injection within the past three months, or history of any significant facial trauma.

## **Photograph Sample Acquisition**

### *PHASE 1*

As previously discussed, Dr. Val Lambros has a collection of approximately 130 photos of primarily Caucasian young females (age 20-30) and 120 photos of older females (age  $\geq 68$  years). The images were collected in a standardized fashion using a handheld Vectra H1 3D imaging system (Canfield Scientific Inc., Parsippany, NJ) under standardized lighting and subject position parameters.

### *PHASE 2*

The efforts described in this proposal are supported by the research division of Canfield Scientific Inc. and discounted equipment will be provided for the acquisition of 3D images. Specifically, two Vectra XT 3D floor-standing imaging systems will be purchased which allow for automated and rapid (3.5 millisecond) acquisition of 3D photos in standardized fashion. Additionally, one Vectra H1 3D handheld imaging system will be



purchased that will allow for the remote capture of subject images in the event that photos need to be captured on UC Irvine's main campus, for instance.

Dr. Lambros serves as a technical advisor on this project and will provide expertise and training on how to capture properly lit and appropriately positioned photos. Special attention will be made to ensure that each photo is taken in a standardized fashion. The subject will be asked to have a neutral facial expression and hair will be tied back as to not obscure the face in any way. The distance and height of the face will be standardized.

### **Image Grouping and Analysis**

Phase 1 of this project will be a proof-of-concept analysis on already collected 3D images. As such, the images are already grouped into two cohorts: 130 photos of Caucasian young females (age 20-30) and 120 photos of older females (age  $\geq$  68 years).

As subject photos are acquired for phase 2 of the study, work will begin with regards to image landmarking and composite photo generation. Images will be grouped by sex and ethnicity. Meaning only Caucasian females will be grouped together, where Caucasian males will be placed in a separate group. Stratification by age will be determined based on the sample collected.

#### *LANDMARK IDENTIFICATION*

Initially, images will be manually demarcated based on the anthropometric landmarks derived from key studies.<sup>20,44,65,73</sup> These particular landmarks were chosen to maximize the ability to measure symmetry, sexual dimorphism score, and degree of youthfulness in all three dimensions (Figures 1 & 2, Table 9). The angles of the face identifiable in the lateral view will also be considered in the creation of a composite score

of facial dimensions (Figure 3). Consideration will be given to employing principal component analysis to determine the overall contributions of individual proportions as they impact the three attractiveness components previously mentioned.

Figure 1. Anthropometric Landmarks for proportions assessment (frontal view).

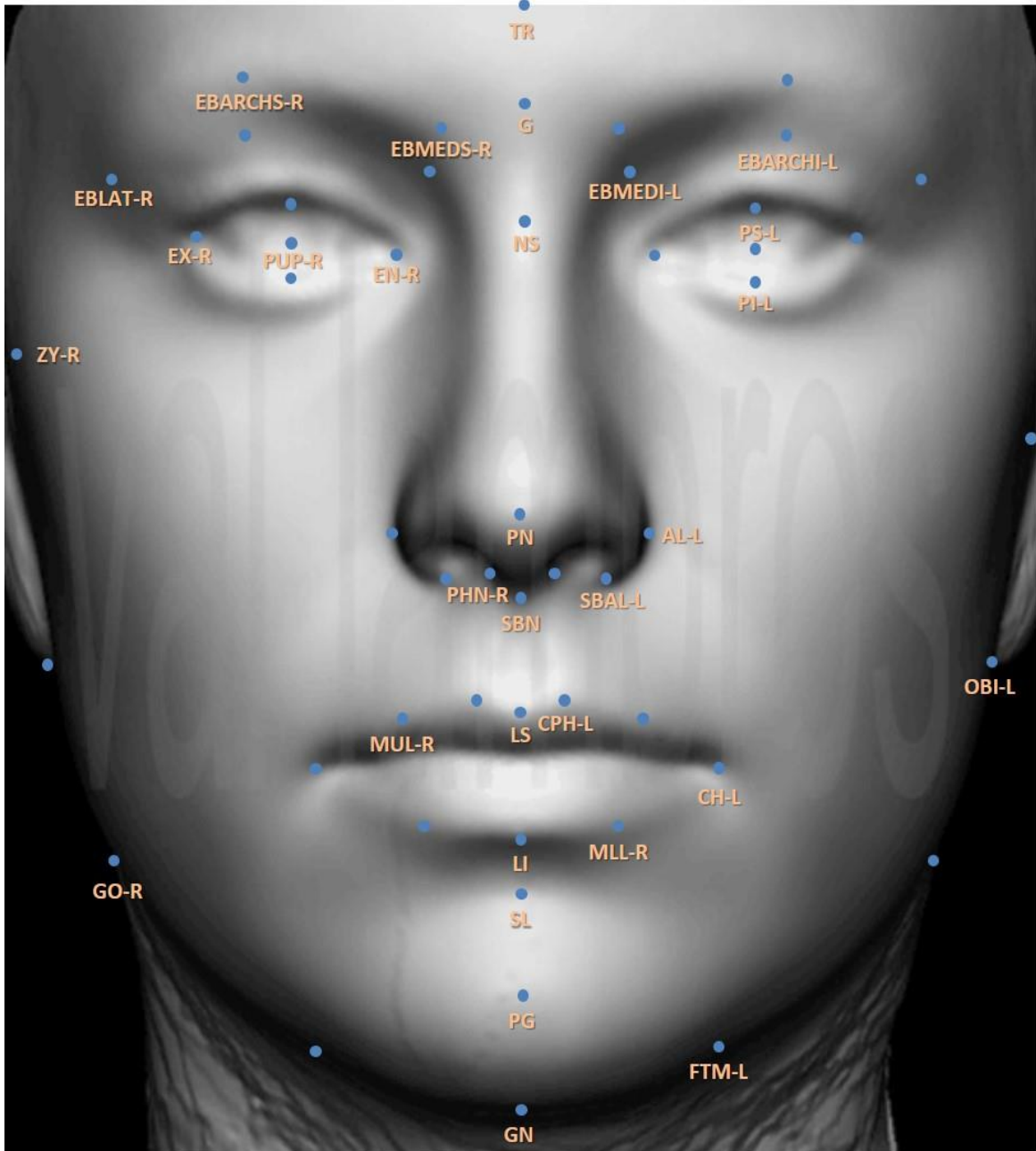


Figure 2. Anthropometric Landmarks for proportions assessment (lateral view).

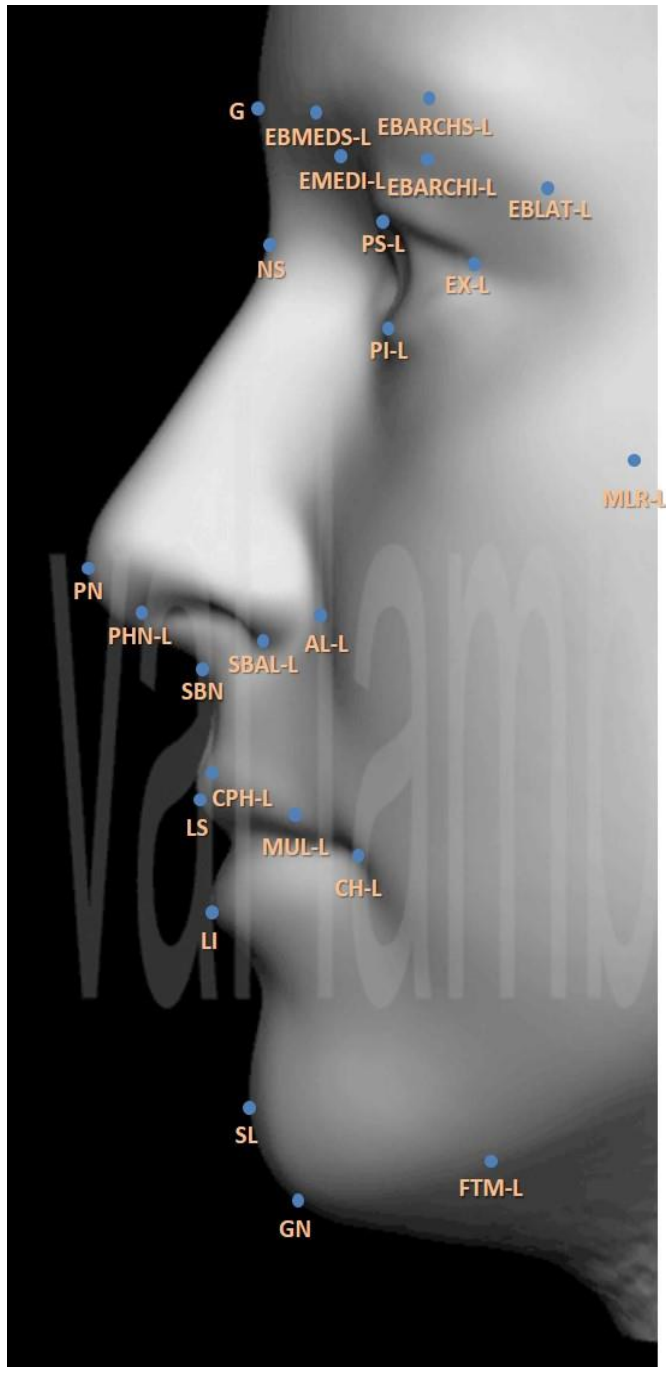
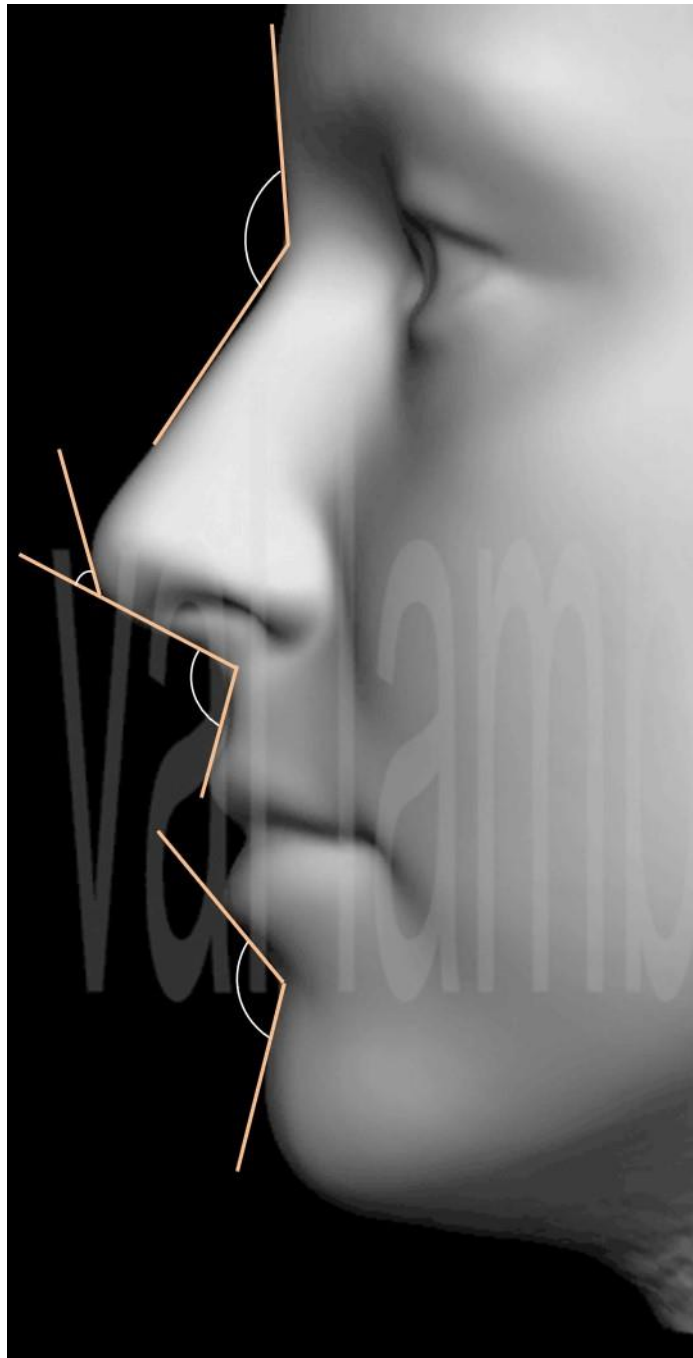


Table 9. Anthropometric landmarks for proportion assessment.<sup>20,44,65,73,81</sup>

<b>Abbreviation</b>	<b>Name</b>	<b>Description</b>
TR	Trichion	The point where the normal hairline and midline of forehead intersect.
G	Glabella	The most prominent midpoint between the eyebrows
EBARCHS	Eyebrow arch superiorly	The position of the eyebrow arch superiorly
EBARCHI	Eyebrow arch inferiorly	The position of the eyebrow arch inferiorly
EBMEDS	Medial eyebrow superiorly	The position of the medial eyebrow superiorly
EBMEDI	Medial eyebrow inferiorly	The position of the medial eyebrow inferiorly
EBLAT	Lateral eyebrow	The position of the lateral eyebrow
PS	Palpebrale superius	The highest point in the midportion of the free margin of each upper eyelid
PI	Palpebrale inferius	The lowest point in the midportion of the free margin of each lower eyelid
PUP	Pupil	The midpoint of the pupil
EN	Endocanthion	The point at the inner commissure of the eye fissure
EX	Exocanthion	The point at the outer commissure of the eye fissure
ZY	Zygion	The most lateral point of each zygomatic arch
NS	Nasion	Midline point between the nasal root and nasofrontal suture, above the line that connects the two inner canthi
PN	Pro-nasale	The most protruded point of the apex nasi
AL	Alare	The most lateral point on each alar contour
PHN	Philtrum-nasale	Lateral point at the intersection of the columella base and the philtrum
SBN	Sub-nasale	The midpoint of the angle at the columella base where the lower border of the nasal septum and the surface of the upper lip meet
SBAL	Sub-alar	Labial insertion of the alar base
OBI	Otobasion inferius	The point on each elevated margin of the philtrum, just above the vermilion line
LS	Labiale superius	The midpoint of the upper vermilion line
LI	Labiale inferius	The midpoint of the lower vermilion line
MUL	Mid-upper lip	The midpoint between crista philtri and cheilion on the vermilion border of the upper lip
CPH	Crista philtrum	The point on each elevated margin of the philtrum, just above the vermilion line
CH	Chelion	The point at each labial commissure
MLL	Mid-lower lip	The midpoint between cheilion and labiale inferius on the vermilion border of the lower lip
SL	Sub-labiale	The most superior midpoint on the labiomental soft tissue contour that defines the border between the lower lip and the chin
PG	Pogonion	The most prominent midpoint of the chin
GO	Gonion	The most lateral point on the mandibular angle
FTM	First-third mandible	The most medial third of the mandible, along the jawline
GN	Gnathion	The lowest median point on the lower border of the mandible

Figure 3. Anthropometric Angles for proportions assessment (lateral view).



### *KOINOPHILIA INDEX*

Similar to the work completed by Bashour *et al.*, image composites will be made with a progressive number of subjects per composite.<sup>20</sup> This means, for phase 1 of the

study that includes 120 photos, composites of 2, 4, 8, 15, 30, 60, 90 and 120 subjects will be created. The composites will be created as previously described,<sup>70</sup> or by using novel software that is currently in beta at Canfield Scientific.

All feature points listed in figures 1 and 2 as well as the angles demonstrated in figure 3 will be considered in the creation of the average composite photos. For each ethnicity and sex, a composite attractiveness index (CAI) will be created from the composite photos generated as previously described.<sup>20</sup> For each individual photo, landmark measurements will be taken and a total deviation score, or koinophilia index (KI) will be created. Based on the early work by Perret *et al.*, it is reasonable to suspect that the averaged face of attractive individuals will be perceived as more attractive than the averaged face of the entire cohort.<sup>16</sup> Therefore,  $KI_{Sum}$  will represent the deviation score from the total cohort, while  $KI_{Ideal}$  will represent the deviation score from the most attractive subpopulation of the total cohort.

Consideration will also be given to performing regression analysis on the contribution that each facial proportion contributes to attractiveness ratings as a function of koinophilia. For instance, once the CAI is determined, deviations from individual face proportions will be measured and used as continuous independent variables (e.g. **alare-alare**, **endocanthion-endocanthion**, etc.). After the individual photos are rated, linear regression analysis can be used to determine the degree that each facial proportion deviates from the individual proportions of the ideal, and how each of these feature points contribute to overall attractiveness in the setting of koinophilia, for example:

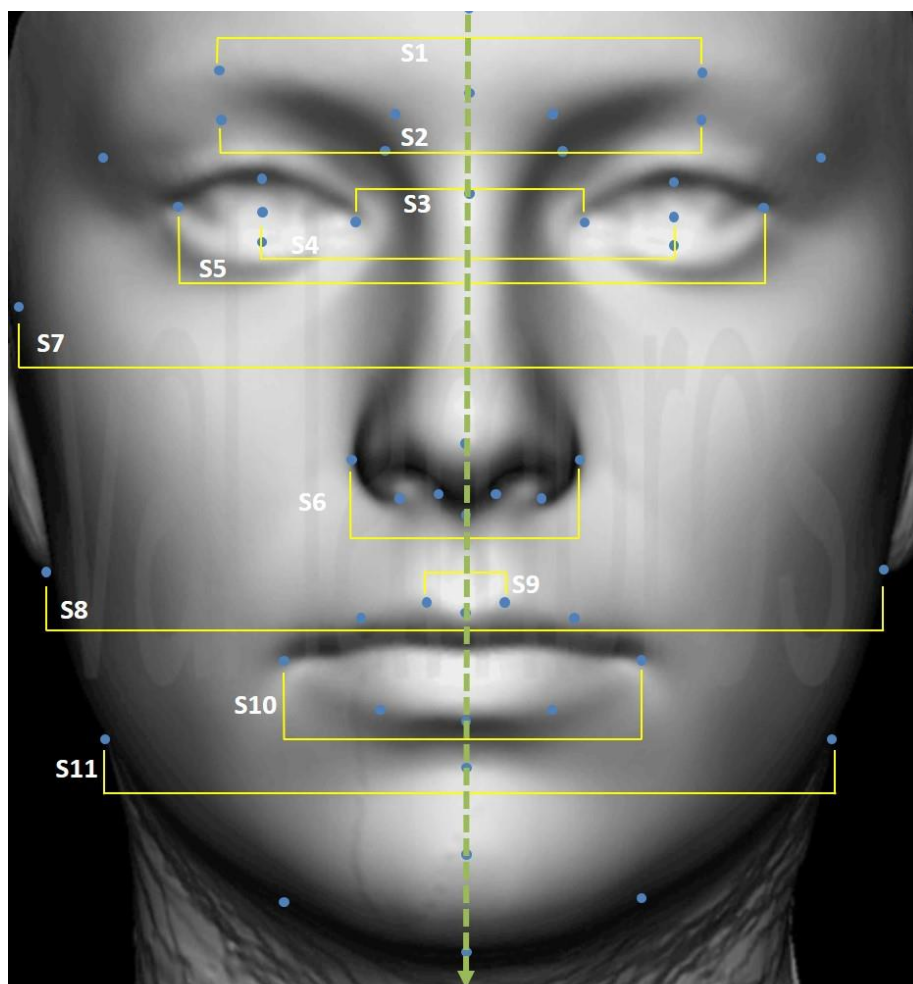
$$KI = \beta_0 + \beta_1AL + \beta_2EN + \dots$$

This technique can be employed in the evaluation of the other components of attractiveness such as symmetry, sexual dimorphism and youthfulness.

### *SYMMETRY*

Symmetry will be measured and scored similar to the methods used by Schmid *et al.*<sup>21</sup> Symmetry will be scored using the following component: eyebrows, eyes, nose, ears, lips and face (figure 4).

Figure 4. Anthropometric proportions for symmetry assessment (frontal view).



Schmid *et al.* determined that the mean difference equation of the symmetry equations evaluated was most strongly associated with attractiveness,<sup>21</sup> therefore the following equation will be used

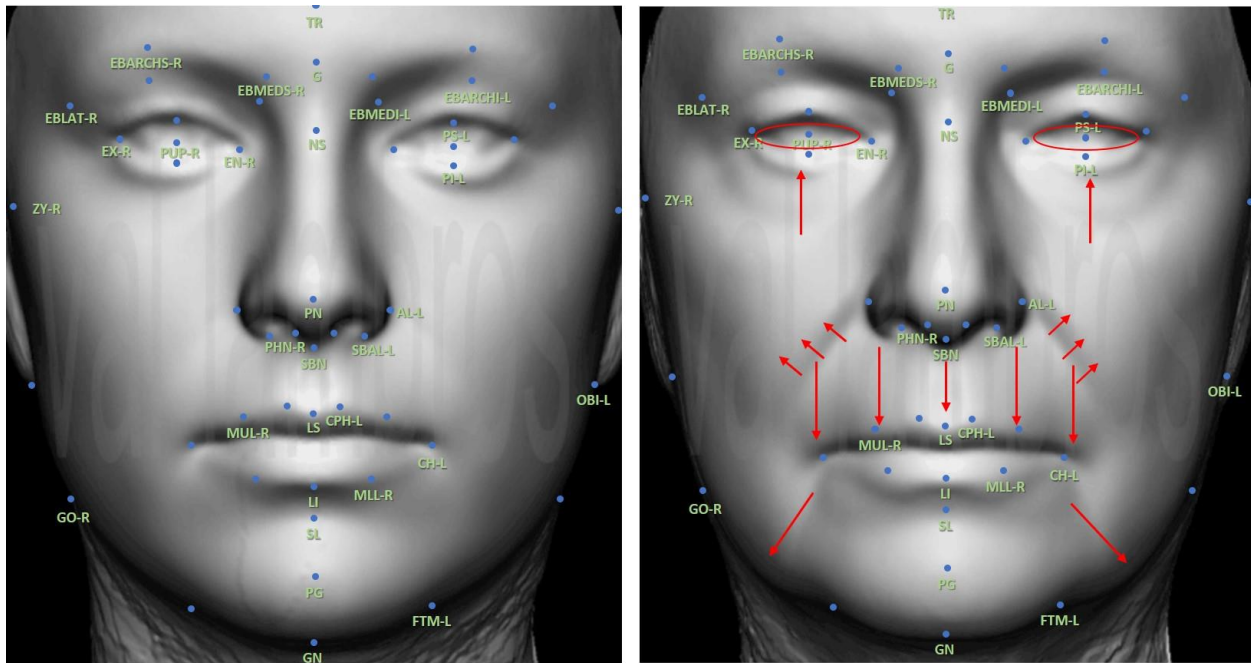
$$\text{Adjusted Difference: } FSM_{AdjDiff}(d) = (d_L - d_R) / [(d_L + d_R)/2]$$

and deviation from this equation will be used to score symmetry.

### YOUTHFULNESS

The changes primarily associated with aging include narrowing of the eyes, raising of the lower eyelid, thinning of the upper and lower lip with migration of the horizontal labial fissure, prominence of the nasolabial fold, posterior movement of the columellar base, elevation of the alar bases, and ptosis of the pro-nasale (figure 5).<sup>70</sup>

Figure 5. Anthropometric proportions associated with aging (frontal view).



Similar to the KI scoring system, the proportions of the young ideal face (20-30 years) will be compared to the proportions of the aged face ( $\geq 68$  years). The deviation of the aged proportions from the young proportions will yield a youthfulness score which will be used to determine the extent of the changes related to aging.

Phase 1 of this study will be limited to the assessment of koinophilia, symmetry and youthfulness. Once proof-of-concept has been established using the three components of

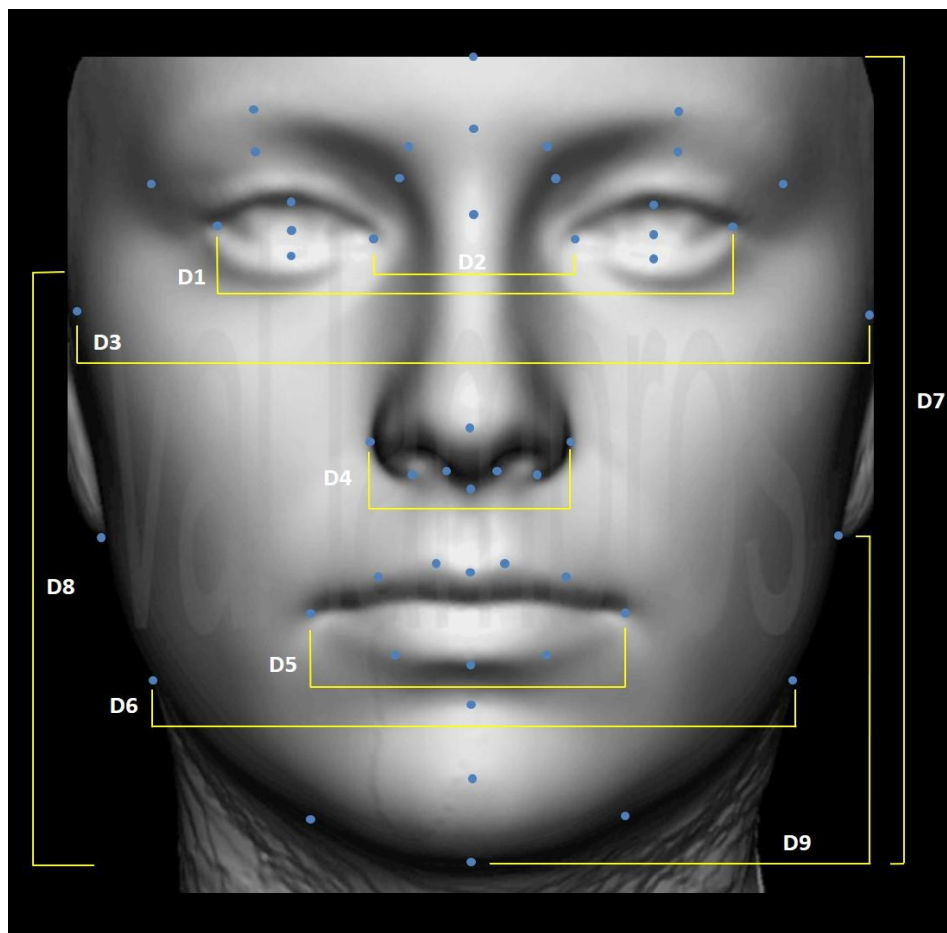


attractiveness, efforts will be made to more precisely define the remaining components and their impact on beauty.

### *SEXUAL DIMORPHISM*

Sexual dimorphism score (SDS) will be derived from feature points first described by Little *et al.*<sup>61</sup> The distance between specific points will be measured and used to calculate four ratios based on the following distances: cheekbone prominence ( $D3/D6$ ), jaw height/lower face height ( $D9/D8$ ), lower face height/face height ( $D8/D7$ ), and face width/lower face height ( $D3/D8$ , Figure 3).

Figure 6. Facial proportions assessment for assessment of sexual dimorphism.



It is predicted that the upper and lower limits of these proportions will depend on the ethnicity examined. Additionally, previous evidence has demonstrated that less urbanized societies prefer more masculinized faces, therefore the observer must be accounted for when examining the role that sexual dimorphism weighs on attractiveness perception.

### *SKIN TONE*

The current body of literature would indicate that within particular ethnic groups, fairer skin tone is positively associated with attractiveness.<sup>82</sup> Anecdotally, it would seem that based on the tanning habits of Americans, there is likely a degree of skin pigmentation that is considered more attractive than a typical hypopigmented state. Prior to the acquisition of 3D photographs, a test photo will be acquired with the subject's face next to X-Rite ColorChecker Classic Chart for later determination of skin pigmentation as previously described.<sup>9</sup>

## **Machine Learning**

Machine learning offers a number of advantages over traditional methods of analysis of facial photos. Dr. Charless Fowlkes is an expert in computer vision, associate professor of computer science at UC Irvine, and an advisor on this project. Once we begin the acquisition of 3D photos, our first task will be to automate the process of landmark demarcation. Simple use of active shape models has proven inefficient in automating this process,<sup>64</sup> but the addition of principal component analysis to an algorithm may obviate the inherent inefficiencies. For example, the software provided by Canfield Scientific has the ability to automatically identify anthropometric landmarks, however, these nodal points usually have to be manual adjusted before analysis can begin. Computers can take the

inputs of active shape models-based landmark identification versus the manual corrections and improve upon this process.

The next and even greater challenge involves employing machine learning for the identification of facial measurements and proportions associated with attractiveness not previously described in the literature. Data mining is the manner in which computers utilize data to learn and perform new tasks. Supervised data mining is essentially goal-directed machine learning in which the computer is fed information based on a set of parameters or previously acquired knowledge. For instance, in the setting of sexual dimorphism, a computer can be instructed that males typically have thicker, more angular jaws and wider noses and use this information to attempt to distinguish female photos from male photos<sup>83</sup> Because there are virtually infinite measurements and proportions that are associated with facial attractiveness within a particular ethnicity, an unsupervised data mining approach may prove optimal to determining the degree of jaw-angularity or nose width that distinguishes an attractive male from an attractive female in a particular ethnic cohort.

The implications of machine learning in this field are broad and beyond the scope of this proposal but will likely provide new insights into different ways in which humans perceive the human face and facial attractiveness. More complex approaches such as deep neural networks can also be employed if the previously described computational methods prove insufficient.<sup>83</sup>

## **Rater Recruitment**

For phase 1 of the study convenience sampling will be employed to recruit adult volunteer raters at UCIMC and UC Irvine main campus. A minimum of 36 participants will be recruited (18 male and 18 female) and asked to rate the attractiveness of photos subjectively (see below). This is similar to the methodology employed by Schmid *et al.*<sup>21</sup>

For phase 2 of the proposal, students, faculty and staff are the primary source for raters recruited. Quota sampling will be employed to ensure that there is an equal cohort of male and female raters. Based on the research by Coetzee *et al.*,<sup>22</sup> perceptual narrowing is minimized by employing raters who are culturally privileged, therefore the only exclusion criteria will be American-born. Inclusion criteria is adult age ( $\geq 18$  years). In total, a minimum 50 raters will be recruited (25 male, 25 female), similar to the survey arm employed by Bashour.<sup>20</sup>

## **Subjective and Objective Rating Systems**

### *SUBJECTIVE RATING SYSTEM*

Similar to previous studies, the raters will be given a questionnaire to measure attractiveness. The raters will be subjected to 3D rotating images of each photo. Because attractiveness is judged more from frontal and lateral profiles, the rotating 3D image will pause on these positions for a short period of time.<sup>73</sup> During four 30-minute sessions, raters will be subjected to a series of photos that will include individual as well as composite photos (e.g. composite of 2, 8, 60, 120 photos). Some photos will be duplicated to assess for individual rater reliability. A ten-point Likert scale will be used to subjectively quantify attractiveness with 1 being associated with very unattractive and 10 indicating the

most attractive. Additionally, cronbach's alpha coefficient will be measured to assess interrater reliability.

### *OBJECTIVE RATING SYSTEM*

For phase 2 of the study, an objective rating system will be employed which includes fMRI and eye-tracking to supplement the subjective rating system. Previous studies have demonstrated that fMRI is a useful tool for collecting real-time data regarding the reward pathways that are activated when individuals view attractive faces, including the nucleus accumbens, ventromedial prefrontal and orbitofrontal cortices.<sup>59,84,85</sup> Eye tracking for pupillary response, granular movements and saccadic eye movements will also be combined to gain additional objective measures as previously described.<sup>6,86</sup> It is anticipated that correlations between eye movements and pupillary size will correlate with brain activity that may uncover other areas of potential research.

## **Statistical Analysis**

The primary dependent variable for this study is facial attractiveness rating. Secondary dependent variable includes adherence to the neoclassical canons and adherence to phi or the 'golden ratios.' The primary independent variables for this study include ethnicity (nominal), Koinophilia Index (e.g. averageness – continuous), facial symmetry (continuous), sexual deviation score (e.g. degree of masculinity/femininity – ordinal or continuous) the morphological changes associated with aging (ordinal or continuous), skin tone (ordinal), sex of the subject (dichotomous), sex of the rater (dichotomous), ethnicity of the subject (ordinal), and ethnicity of the rater (ordinal).

### *Proposed Multiple Linear Regression Equation (Phase 1)*

$$\text{Facial Attractiveness} = \beta_0 + \beta_1\text{KI} + \beta_2\text{Symmetry} + \beta_3\text{Youth} + e_i,$$

where KI = koinophilia index (measure of averageness), Youth = degree of age-related changes to facial measurements and Symmetry = symmetry of the face as measured on a vertical axis

Phase 1 analysis will be exclusively conducted on the images of Caucasian females. The regression equation will allow the user to predict the attractiveness score of an individual based on their degree of symmetry, youthful proportions, and the amount to which their facial proportions deviate from the mathematical average.

*Proposed Multiple Linear Regression Equation (Phase 2)*

$$\text{Facial Attractiveness} = \beta_0 + \beta_1\text{KI} + \beta_2\text{SDS} + \beta_3\text{Youth} + \beta_4\text{Symmetry} + \beta_5\text{Tone} + \beta_6\text{S} + \beta_7\text{R} + \beta_8(\text{S} \times \text{R}) + \beta_9\text{E} + \beta_{10}\text{D} + \beta_{11}(\text{E} \times \text{D}) + \beta_{12}(\text{E} \times \text{SDS}) + e_i,$$

where KI = koinophilia index (measure of averageness), SDS = sexual dimorphism score, Youth = degree of age-related changes to facial measurements, Symmetry = symmetry of the face as measured on a vertical axis, Tone = skin tone of the patient, S = the sex of the subject, R = the sex of the rater, S x R = impact of subject-rater relationship (sex), E = the ethnicity of the subject, D = the ethnicity of the rater, E x D = impact of subject-rater relationship (ethnicity) and E x SDS = impact of subject ethnicity and sexual dimorphism score.

The multiple linear regression equation for phase II is complex, and it is anticipated that additional data will be generated from Phase 1 that modifies our statistical approach. Based on the number of covariates, it is anticipated that a minimum of 120 subjects per group analyzed will be needed to reach statistical significance.

## **Anticipated Results and Next Steps**

Based on the extensive research indicating that phi and the neoclassical canons don't represent the individual or average proportions or measurements of various ethnicities doesn't imply that these ethnicities (non-European, white) do not possess individuals with attractive faces. On the contrary, each ethnicity likely possesses individuals whose attractiveness ratings follow a normal distribution when examining a sample that represents that entire cohort. I predict that the faces of each ethnicity studied based on the proposal herein will fit a normal distribution of attractiveness scores and will also correlate to activation of the reward pathways of the nucleus accumbens, ventromedial prefrontal and orbitofrontal cortices.

### *PHASE 1 (6-8 Months)*

The initial pilot study will require two separate ratings sessions. The first will involve a cohort of raters to judge individual and composite photos for baseline ratings and analysis. The age, race and gender of each rater will be recorded and subgroup analysis for interrater reliability will be tested. The beta values of each component will be analyzed to detect the weight that each factor impacts the perception of facial attractiveness. Additional regression analysis will be performed on each individual component to determine whether factors in that component can be eliminated from consideration of impact in the form of principal component analysis. I predict that averageness will impart the largest contribution to facial attractiveness perception, followed by symmetry, sexual dimorphism, youthfulness, and skin tone. Data would suggest that sexual dimorphism may be more important for female than male subjects.

It is also anticipated that that gender, age, and ethnicity will have significant impacts on the perception of attractiveness. Based on evolutionary theory, it is a distinct possibility that males will value the images of younger females, since fertility peaks earlier in life for women. On the other hand, women may prefer men of slightly more advanced age, because from a cultural standpoint, older men are likely to have more financial stability and be able to provide for a family. Based on the idea of perceptual narrowing, it is also likely that individuals of one ethnicity may rate less familiar ethnicities as lower in attractiveness. At the completion of Phase 1, I anticipate robust data that will allow the determination of hypotheses 1, 2 and the majority of 3, with the exception of skin tone analysis. These first three hypotheses are focused on the beauty gestalt and will be more easily studied with the photos I have at my disposal.

#### *PHASE 2 (18 – 24 Months)*

Phase 2 of this proposal will require funding to recruit subjects, raters, obtain high-speed 3D cameras, employ a clinical research coordinator and integrate the use of objective ratings systems. Regardless the outcome of Phase 1, it is anticipated that the data will provide an incentive to explore this approach in other ethnicities. For instance, if raters perceive the  $KI_{Sum}$  of Caucasian women more attractive than the  $KI_{Ideal}$  of women from the same cohort, it will be incumbent upon this project to answer the question whether this relationship holds true for Korean women, for example.

The greatest anticipated bottleneck is the recruitment of subjects. Subjects will be enrolled who are patients at UCIMC and PCPS. To bolster our sample of ethnic patients, we will also recruit students, faculty and staff from the University of California, Irvine, a cohort of approximately 44,000 individuals. Due to the complexity of the proposed phase 2 linear



regression equation, we will seek to reproduce the measures of the phase 1 study (koinophilia, symmetry and youthfulness) and a second ethnicity of female subjects. In addition to these cofactors, the additional factors previously mentioned will be explored, adding covariates as our sample size and recruitment factors permit. I anticipate that at the completion of Phase 2, we will have enough data to answer many of the questions I have raised in this analysis. If successful, Phase 2 will generate findings that will encourage a much broader and robust study that examines the features of even more ethnicities as well as male subjects.

## **CHAPTER 5: SUMMARY AND CONCLUSIONS**

Society is becoming increasingly fixated on self-image with the invention of smart phones and social media applications like Instagram and Snapchat. Social influencers like Kim Kardashian are contributing to a building environment in which the masses feel compelled to alter their physical features in ways that don't necessarily reflect the aesthetic ideal. It is incumbent on the scientific community, and particular the field of plastic surgery, to develop tools that allow us to objectively quantify the norms of beauty for the various ethnic cohorts that we treat on a daily basis. Research has demonstrated that older tools like phi and the neoclassical canons don't apply to various ethnicities and more a modern tool, the 'phi mask,' has only been validated in the population from which it was developed.

Here, I propose a methodology that will allow for the objective quantification of facial attractiveness that is sensitive to the age, sex, and ethnicity of both the subject, and the observer that subject intends to impress. I hope that this proposal eventually results in a multi-discipline, multi-center, collaborative research study that seeks to objectively

quantify the norms and extremes of facial aesthetics much in the same way the human genome project continues to further define the genetic traits that comprise humanity.

## REFERENCES

1. Rushton W. A possible Cause of Facial Asymmetry. *Proc R Soc Med* 1911;4:105-9.
2. Aly A, Tolazzi A, Soliman S, Cram A. Quantitative analysis of aesthetic results: introducing a new paradigm. *Aesthet Surg J* 2012;32:120-4.
3. Ching S, Thoma A, McCabe RE, Antony MM. Measuring outcomes in aesthetic surgery: a comprehensive review of the literature. *Plast Reconstr Surg* 2003;111:469-80; discussion 81-2.
4. Langlois JH, Roggman LA, Casey RJ, Ritter JM, Rieser-Danner LA, Jenkins VY. Infant Preferences for Attractive Faces: Rudiments of a Stereotype? *Developmental Psychology* 1987;23:363-9.
5. Trujillo LT, Jankowitsch JM, Langlois JH. Beauty is in the ease of the beholding: a neurophysiological test of the averageness theory of facial attractiveness. *Cogn Affect Behav Neurosci* 2014;14:1061-76.
6. Zhang Y, Wang X, Wang J, Zhang L, Xiang Y. Patterns of Eye Movements When Observers Judge Female Facial Attractiveness. *Front Psychol* 2017;8:1909.
7. Folstad I, Karter AJ. Parasites, bright males, and the immunocompetence handicap. *The American Naturalist* 1992;139:603-22.
8. Wedekind C. Detailed information about parasites revealed by sexual ornamentation. *Proc R Soc Lond B* 1992;247:169-74.
9. Foo YZ, Simmons LW, Rhodes G. Predictors of facial attractiveness and health in humans. *Sci Rep* 2017;7:39731.
10. Lemley B. Do You Love This Face? *Discover* 2000 February 1, 2000.
11. Farrera A, Villanueva M, Quinto-Sanchez M, Gonzalez-Jose R. The relationship between facial shape asymmetry and attractiveness in Mexican students. *Am J Hum Biol* 2015;27:387-96.
12. Scott IM, Clark AP, Josephson SC, *et al.* Human preferences for sexually dimorphic faces may be evolutionarily novel. *Proc Natl Acad Sci U S A* 2014;111:14388-93.
13. Vera Cruz G. The impact of face skin tone on perceived facial attractiveness: A study realized with an innovative methodology. *J Soc Psychol* 2017:1-11.
14. Galton F. Composite Portraits, Made by Combining Those of Many Different Persons Into a Single Resultant Figure. *The Journal of the Anthropological Institute of Great Britain and Ireland* 1879;8:132-44.
15. Langlois JH, Roggman LA. Attractive Faces Are Only Average. *Psychological Science* 1990;1:115-21.
16. Perrett DI, May KA, Yoshikawa S. Facial shape and judgements of female attractiveness. *Nature* 1994;368:239-42.
17. Thorne C, Chung KC, Gosain A, Guntner GC, Mehrara BJ. *Grabb and Smith's plastic surgery / editor-in-chief, Charles H. Thorne ; editors, Kevin C. Chung, Arun Gosain, Geoffrey C. Gurtner, Babak Joseph Mehrara, J. Peter Rubin, Scott L. Spear. Seventh edition. ed.*
18. The L. Cosmetic procedures: a cause for concern. *Lancet* 2017;390:2.
19. Veale D, Gledhill LJ, Christodoulou P, Hodsoll J. Body dysmorphic disorder in different settings: A systematic review and estimated weighted prevalence. *Body Image* 2016;18:168-86.
20. Bashour M. An objective system for measuring facial attractiveness. *Plast Reconstr Surg* 2006;118:757-74; discussion 75-6.
21. Schmid K, Marx D, Samal A. Computation of a face attractiveness index based on neoclassical canons, symmetry, and golden ratios. *Pattern Recognition* 2008;41:2710-7.
22. Coetzee V, Greeff JM, Stephen ID, Perrett DI. Cross-cultural agreement in facial attractiveness preferences: the role of ethnicity and gender. *PLoS One* 2014;9:e99629.
23. Bashour M. History and current concepts in the analysis of facial attractiveness. *Plast Reconstr Surg* 2006;118:741-56.
24. Fang F, Clapham PJ, Chung KC. A systematic review of interethnic variability in facial dimensions. *Plast Reconstr Surg* 2011;127:874-81.
25. Al-Sebaei MO. The validity of three neo-classical facial canons in young adults originating from the Arabian Peninsula. *Head Face Med* 2015;11:4.
26. Cvicelova M, Benus R, Lysakova L, Molnarova A, Borovska Z. Occurrence of neoclassical facial canons in Caucasian primary school pupils and university students. *Bratisl Lek Listy* 2007;108:480-5.
27. Husein OF, Sepehr A, Garg R, *et al.* Anthropometric and aesthetic analysis of the Indian American woman's face. *J Plast Reconstr Aesthet Surg* 2010;63:1825-31.
28. Jayaratne YS, Deutsch CK, McGrath CP, Zwahlen RA. Are neoclassical canons valid for southern Chinese faces? *PLoS One* 2012;7:e52593.

29. Karaca Saygili O, Cinar S, Gulcen B, Ozcan E, Kus I. The validity of eight neoclassical facial canons in the Turkish adults. *Folia Morphol (Warsz)* 2016;75:512-7.
30. Kusugal P, Ruttonji Z, Gowda R, Rajpurohit L, Lad P, Ritu. Three-dimensional facial analyses of Indian and Malaysian women. *Contemp Clin Dent* 2015;6:332-6.
31. Porter JP. The average African American male face: an anthropometric analysis. *Arch Facial Plast Surg* 2004;6:78-81.
32. Porter JP, Olson KL. Anthropometric facial analysis of the African American woman. *Arch Facial Plast Surg* 2001;3:191-7.
33. Wang D, Qian G, Zhang M, Farkas LG. Differences in horizontal, neoclassical facial canons in Chinese (Han) and North American Caucasian populations. *Aesthetic Plast Surg* 1997;21:265-9.
34. Zhao Q, Zhou R, Zhang X, *et al.* Morphological quantitative criteria and aesthetic evaluation of eight female Han face types. *Aesthetic Plast Surg* 2013;37:445-53.
35. Livio M. The golden ratio and aesthetics. *Plus Magazine*, 22 (<http://plus.maths.org/>) 2002.
36. High definition body sculpting : art and advanced lipoplasty techniques. New York: Springer; 2014.
37. Pallett PM, Link S, Lee K. New "golden" ratios for facial beauty. *Vision Res* 2010;50:149-54.
38. Making The Mask. Marquardt Aesthetic Imaging, Inc., 2014. at <https://www.beautyanalysis.com/research/our-research/making-mask/>)
39. Holland E. Marquardt's Phi mask: pitfalls of relying on fashion models and the golden ratio to describe a beautiful face. *Aesthetic Plast Surg* 2008;32:200-8.
40. Rhee SC, Koo SH. An objective system for measuring facial attractiveness. *Plast Reconstr Surg* 2007;119:1952-3; author reply 3-4.
41. Facial Beauty and the "New" Golden Ratio (or is it just 1.618 in disguise?). 2013. at <https://www.goldennumber.net/facial-beauty-new-golden-ratio/>.)
42. Kaptein YE, Kaptein JS, Markarian A. Vertical Localization of the Malar Prominence. *Plast Reconstr Surg Glob Open* 2015;3:e411.
43. Anand S, Tripathi S, Chopra A, Khaneja K, Agarwal S. Vertical and horizontal proportions of the face and their correlation to phi among Indians in Moradabad population: A survey. *J Indian Prosthodont Soc* 2015;15:125-30.
44. Jang KS, Bayome M, Park JH, Park KH, Moon HB, Kook YA. A three-dimensional photogrammetric analysis of the facial esthetics of the Miss Korea pageant contestants. *Korean J Orthod* 2017;47:87-99.
45. Machard A, Jomier M, Hottelart D, Vie K. Identification of new morphological differences between Chinese and Caucasian faces and influence of BMI on these characteristics. *Skin Res Technol* 2016;22:137-47.
46. Saad A, Hewett S, Nolte M, Delaunay F, Saad M, Cohen SR. Ethnic Rhinoplasty in Female Patients: The Neoclassical Canons Revisited. *Aesthetic Plast Surg* 2018;42:565-76.
47. Pavlic A, Trinajstic Zrinski M, Katic V, Spalj S. Neoclassical canons of facial beauty: Do we see the deviations? *J Craniomaxillofac Surg* 2017;45:741-7.
48. Schmid KK, Marx DB, Samal A. Analysis of landmark data using multidimensional regression. Lincoln, Neb.: University of Nebraska-Lincoln.; 2007.
49. Rhodes G. The evolutionary psychology of facial beauty. *Annu Rev Psychol* 2006;57:199-226.
50. Little AC, Jones BC, DeBruine LM. Facial attractiveness: evolutionary based research. *Philos Trans R Soc Lond B Biol Sci* 2011;366:1638-59.
51. Langlois JH, Roggman LA, Musselman L. What Is Average and What Is Not Average About Attractive Faces. *Psychological Science* 1994;5:214-20.
52. Average. Merriam-Webster, 2018. (Accessed 5/1/2018, at <https://www.merriam-webster.com/thesaurus/average>.)
53. Average. 2018.
54. Olson IR, Marshuetz C. Facial attractiveness is appraised in a glance. *Emotion* 2005;5:498-502.
55. Rubenstein AJ, Kalakanis L, Langlois JH. Infant preferences for attractive faces: a cognitive explanation. *Dev Psychol* 1999;35:848-55.
56. Sharabi SE, Hatef DA, Hollier LH, Jr. Facial attractiveness: is the whole more than the sum of its parts? *Aesthet Surg J* 2010;30:154-60.
57. Little AC, Hockings KJ, Apicella CL, Sousa C. Mixed-ethnicity face shape and attractiveness in humans. *Perception* 2012;41:1486-96.
58. Hahn AC, Perrett DI. Neural and behavioral responses to attractiveness in adult and infant faces. *Neurosci Biobehav Rev* 2014;46 Pt 4:591-603.

59. Pegors TK, Kable JW, Chatterjee A, Epstein RA. Common and unique representations in pFC for face and place attractiveness. *J Cogn Neurosci* 2015;27:959-73.
60. Komori M, Kawamura S, Ishihara S. Effect of averageness and sexual dimorphism on the judgment of facial attractiveness. *Vision Res* 2009;49:862-9.
61. Little AC, Jones BC, Waite C, *et al.* Symmetry is related to sexual dimorphism in faces: data across culture and species. *PLoS One* 2008;3:e2106.
62. Vingilis-Jaremko L, Maurer D. The influence of symmetry on children's judgments of facial attractiveness. *Perception* 2013;42:302-20.
63. Iyengar A, Kulkarni R, Vidya TNC. Koinophilia and Human Facial Attractiveness. *Resonance* 2015;20:311-9.
64. Chen FM, Zhang D. Evaluation of the Putative Ratio Rules for Facial Beauty Indexing. 2014 International Conference on Medical Biometrics (Icmb 2014) 2014:181-8.
65. Djordjevic J, Zhurov AI, Richmond S, Visigen C. Genetic and Environmental Contributions to Facial Morphological Variation: A 3D Population-Based Twin Study. *PLoS One* 2016;11:e0162250.
66. Rhodes G, Lee K, Palermo R, *et al.* Attractiveness of own-race, other-race, and mixed-race faces. *Perception* 2005;34:319-40.
67. Perrett DI, Lee KJ, Penton-Voak I, *et al.* Effects of sexual dimorphism on facial attractiveness. *Nature* 1998;394:884-7.
68. Rennels JL, Bronstad PM, Langlois JH. Are attractive men's faces masculine or feminine? The importance of type of facial stimuli. *J Exp Psychol Hum Percept Perform* 2008;34:884-93.
69. Torrance JS, Wincenciak J, Hahn AC, DeBruine LM, Jones BC. The relative contributions of facial shape and surface information to perceptions of attractiveness and dominance. *PLoS One* 2014;9:e104415.
70. Lambros V, Amos G. Three-Dimensional Facial Averaging: A Tool for Understanding Facial Aging. *Plast Reconstr Surg* 2016;138:980e-2e.
71. Rhodes G, Jeffery L, Watson TL, Clifford CW, Nakayama K. Fitting the mind to the world: face adaptation and attractiveness aftereffects. *Psychol Sci* 2003;14:558-66.
72. Rubenstein AJ, Langlois JH, Roggman LA. What makes a face attractive and why: The role of averageness in defining facial beauty. *Facial attractiveness: Evolutionary, cognitive, and social perspectives*. Westport, CT, US: Ablex Publishing; 2002:1-33.
73. Hatch CD, Wehby GL, Nidey NL, Moreno Uribe LM. Effects of Objective 3-Dimensional Measures of Facial Shape and Symmetry on Perceptions of Facial Attractiveness. *J Oral Maxillofac Surg* 2017;75:1958-70.
74. Hoss RA, Ramsey JL, Griffin AM, Langlois JH. The role of facial attractiveness and facial masculinity/femininity in sex classification of faces. *Perception* 2005;34:1459-74.
75. Valentine KA, Li NP, Penke L, Perrett DI. Judging a man by the width of his face: the role of facial ratios and dominance in mate choice at speed-dating events. *Psychol Sci* 2014;25:806-11.
76. Etcoff NL, Stock S, Haley LE, Vickery SA, House DM. Cosmetics as a feature of the extended human phenotype: modulation of the perception of biologically important facial signals. *PLoS One* 2011;6:e25656.
77. Porcheron A, Mauger E, Russell R. Aspects of facial contrast decrease with age and are cues for age perception. *PLoS One* 2013;8:e57985.
78. Porcheron A, Mauger E, Soppelsa F, *et al.* Facial Contrast Is a Cross-Cultural Cue for Perceiving Age. *Front Psychol* 2017;8:1208.
79. Evison M, Dryden I, Fieller N, *et al.* Key Parameters of Face Shape Variation in 3D in a Large Sample. *J Forensic Sci* 2010;55:159-62.
80. Surgeons ASoP. Plastic Surgery Statistics Report 2016.
81. Swennen GRJ, Schutysse F, Hausamen J-E. Three-dimensional cephalometry : a color atlas and manual. 1st. ed. New York, NY: Springer; 2005.
82. Swami V, Furnham A, Joshi K. The influence of skin tone, hair length, and hair colour on ratings of women's physical attractiveness, health and fertility. *Scand J Psychol* 2008;49:429-37.
83. Amilia S, Sulistiyo MD, Dayawati RN. Face Image-Based Gender Recognition Using Complex-Valued Neural Network. 2015 3rd International Conference on Information and Communication Technology (ICoICT) 2015:201-6.
84. Ito A, Fujii T, Abe N, *et al.* Gender differences in ventromedial prefrontal cortex activity associated with valuation of faces. *Neuroscience* 2016;328:194-200.
85. Kranz F, Ishai A. Face perception is modulated by sexual preference. *Curr Biol* 2006;16:63-8.
86. Valuch C, Pfluger LS, Wallner B, Laeng B, Ansoerg U. Using eye tracking to test for individual differences in attention to attractive faces. *Front Psychol* 2015;6:42.