

A COMPARISON OF CRANIOMETRIC POPULATION AFFINITY ESTIMATIONS
TO SELF-REPORTED IDENTITY

by

Heather J. Nesbitt, B.A.

A thesis submitted to the Graduate Council of
Texas State University in fulfillment
of the requirements for the degree of
Master of Arts
With a Major in Anthropology
August 2022

Committee Members:

M. Katherine Spradley, Chair

Ashley H. McKeown

Kyra E. Stull

COPYRIGHT

by

Heather J. Nesbitt

2022

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

This work is protected by the Copyright Laws of the United States (Public Law 94-553, section 107). Consistent with fair use as defined in the Copyright Laws, brief quotations from this material are allowed with proper acknowledgment. Use of this material for financial gain without the author's express written permission is not allowed.

Duplication Permission

As the copyright holder of this work I, Heather J. Nesbitt, authorize duplication of this work, in whole or in part, for educational or scholarly purposes only.

DEDICATION

For my Nana, who always knew I could.
I only wish you could watch me walk across that graduation stage.
I love you.

ACKNOWLEDGMENTS

This research was conducted with the help of many individuals. To my family, thank you for your unwavering support in both my academic and personal endeavors. To my friends, thank you for the laughs in between schoolwork and constant encouragement. To my partner, thank you for following me and my dreams to Texas. And to my committee, thank you for all your guidance throughout this project. I would be nowhere without my mentors. It truly takes a village.

Additionally, this material is based upon work supported by the Grady Early Research Endowment in Forensic Anthropology. Thank you for funding my time at Texas State.

Finally, I would like to thank and acknowledge the donors of the Texas State Donated Skeletal Collection. Without these selfless individuals, none of this research would be possible. Thank you.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.....	v
LIST OF TABLES	viii
LIST OF FIGURES.....	ix
LIST OF ABBREVIATIONS.....	x
CHAPTER	
I. INTRODUCTION.....	1
Background.....	1
II. MATERIALS AND METHODS.....	10
Data Collection.....	10
Demographics of the Texas State Donated Skeletal Collection.....	11
Demographics of the Forensic Data Bank and FORDISC.....	16
Data Analysis.....	18
FORDISC Protocol.....	19
Limitations.....	22
III. RESULTS.....	23
Summary of Results.....	23
FORDISC Classifications by Populations.....	24
White Individuals.....	24
Black/African American Individuals.....	25
Hispanic Individuals.....	27
Indigenous American Individuals.....	28
Asian Individuals.....	28
Multiracial/Mixed Ancestry Individuals.....	29
Classification Rates by Sex.....	30
Misclassifications.....	31
IV. DISCUSSION.....	32

Future Research.....	38
V. CONCLUSION.....	41
APPENDIX SECTION.....	43
LITERATURE CITED.....	46

LIST OF TABLES

Table	Page
1. Table 1. Standard cranial measurements used for analysis and their abbreviations.....	11
2. Table 2. TXSTDSC Donor Information (sex and race).....	13
3. Table 3. Breakdown of Donors by Self-Described Race.....	16
4. Table 4. Forensic Data Bank groups included in FORDISC 3.1 by sex and total.....	17

LIST OF FIGURES

Figure	Page
1. Figure 1. TXSTDSC Donor Information (sex and race).....	13
2. Figure 2. Breakdown of Donors by Self-Described Race.....	16
3. Figure 3. Posterior Probability Distribution for All Populations.....	24
4. Figure 4. Posterior Probability Distribution for White Individuals.....	25
5. Figure 5. Posterior Probability Distribution for Black/African American Individuals.....	26
6. Figure 6. Posterior Probability Distribution for Hispanic Individuals.....	27
7. Figure 7. Posterior Probability Distribution for Multiracial/Mixed Ancestry Individuals.....	30

LIST OF ABBREVIATIONS

Abbreviation	Description
TXSTDSC	Texas State Donated Skeletal Collection
BIPOC	Black, Indigenous, People of Color
mm	Millimeters
ZMB	Zygomaxillary breadth
ASB	Biasterionic breadth
MOW	Mid-orbital width
MDH	Mastoid height
DFA	Discriminant function analysis
FDB	Forensic Data Bank
FACTS	Forensic Anthropology Center at Texas State
p.p.	Posterior probability

I. INTRODUCTION

In forensic anthropology, ancestry is estimated as part of the biological profile, which also includes estimations of sex, age, and stature, as well as the documentation of identifying characteristics (i.e., pathologies or trauma) of an unidentified individual. While each component is unique, some components can be dependent on others. For example, sex and age can be population-specific (Spradley 2016). Ancestry estimations do not imply that biological races exist based on metric or nonmetric differences. Rather, the estimation assesses how similar or dissimilar an individual is to a specific group and its known frequencies of metric or non-metric features. Because the estimated demographics aim to facilitate the positive identification of unidentified human remains, and there is dependence among the parameters, correct forensic anthropological population affinity estimations are essential.

The purpose of this research is to compare FORDISC 3.1 (Jantz and Ousley 2005) classifications with self-reported identity. I examined 77 crania from the Texas State Donated Skeletal Collection (TXSTDSC) using the craniometric standards in the Data Collection Procedures Manual 2.0 (Langley et al. 2016) and ran analyses in FORDISC 3.1.

Background

Race is paradoxical as it is socially real but not biologically. However, it affects nearly every aspect of daily life, especially within US society (Jenkins 2019). Race exists in the “sociopolitical context,” whereas ancestry is “derived from genomic analyses,” (Popejoy 2020). Even still, race has palpable influences on biology because of social

factors. It is fluid in the sense that individuals may change how they identify throughout their life, especially in admixed populations (Hahn et al. 1996), unlike the skeletal form which is relatively static.

In forensic anthropology, population affinity is an estimation of membership indicating similarity to well-defined reference groups. Current methods depend on the correlation of skeletal morphology to an individual's social race or geographic origin to provide accurate estimations (Dunn et al. 2020). This is a more representative definition than ancestry and changes the overall attitude toward ancestry and the estimation process. Moreover, this definition addresses that biological races do not exist, but that skeletal morphology is influenced by social factors and geographic location (Spradley and Jantz 2022). Shared genetics, population history, and social constructs inform the skeletal morphology and subsequently population affinity (Dunn et al. 2020). We analyze the phenotypic expression of the interactions of social and biological elements emphasizing the need for a biocultural approach when estimating population affinity.

Craniometric measurements and nonmetric traits are commonly used to estimate population affinity as compared to postcranial elements (Spradley 2016). Both methodological approaches to population affinity estimates can produce high accuracy rates depending on various factors (e.g., number of measurements, traits selected) (Bulbeck 2011; Hughes et al. 2011). However, population affinity can be difficult to estimate because of the complex relationship between social constructs and skeletal morphology and because of the lack of representative sample groups and knowledge of population histories. Additionally, population affinity contains cultural and historical factors that affect self-identity, as well as the analysis of an individual's remains (Cunha

and Ubelaker 2019). Because our goal is to aid the identification of deceased individuals, assessing forensic anthropologists' ability to estimate population affinity accurately is of great importance. This research aims to address these concepts.

Human biological variation is present at the genetic and phenotypic levels. Genetically, variation occurs as mutations at the chromosomal levels as well as through gene flow and genetic drift. These evolutionary processes are crucial to furthering diversity within gene pools. Phenotypes are the outward or physical expression of these genes. In addition to genetic influences, the environment in which an individual lives before puberty influences growth and development and the expression of traits (Hochberg 2011). Even still, there is more variation within a group than between, meaning that humans are more alike than different. We belong to one, single species. The cranium is plastic in its morphological expression as evident in differences visible between first-generation U.S.-born individuals and their counterparts born in their native countries (Boas 1912, Relethford 2004). Secular changes do occur, however, genetic constraints still play a role in morphology (Martínez-Abadías et al. 2009).

Population structure does affect the overall morphology. What this means is that anthropologists can use metric and nonmetric data to assess group membership. The practice of positive assortative mating is still largely the cultural norm in the U.S. "The percentage of married-couple households that are interracial or interethnic grew across the United States from 7.4 to 10.2 percent from 2000 to 2012-2016," (Rico et al. 2021). While social structures and taboos have lessened regarding between-group coupling, the majority of people choose companions who look similar to themselves as evident in the

U.S. census data (Rico et al. 2021). A 2021 Gallup poll found that 94% of Americans support interracial marriage (McCarthy 2021). However, the notion that humans have admixed past the point of affinity estimation's use has not yet been realized. Within the United States, the systems that prohibited between-group intermingling have legally been overturned. Yet, an examination of the racial composition of recent censuses (2010 and 2020) indicates that positive assortative mating practices are still the norm. Therefore, cranial morphology still captures population affinity. Understanding contemporary American mating practices allow for the use of craniometrics to estimate an unidentified individual's group affiliation.

Currently, there is a lack of studies that assess the accuracy of population affinity estimation (Thomas et al. 2017; Hahn et al. 1996). Furthermore, because prior research limits evaluations to how correct anthropologists are at estimating the affinity when compared to DNA analysis, results are presented in a three-group model that does not accurately depict a person's identity. Thomas et al. (2017) found that forensic anthropology affinity estimations are correct 90.9% using a three-group classification scheme (i.e., the unknown was classified into the African, European, or Asian group). The study uses the terms European, African, and Asian when comparing positively identified persons with DNA. These broad groups are not compared to the individual's self-identity reported social race from the next of kin, so while these tests produce general ideas related to the affinity of the individual, the models fail to account for how people identify in their daily lives.

The U.S. Census Bureau requires a minimum of five racial categories to be provided on surveys: White, Black or African American, American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander (“Race” N.D.). Even still, five categories are not representative of real life. “People who identify their origin as Hispanic, Latino, or Spanish may be of any race,” (“Race” N.D.), which further complicates the notion of affinity when compared to identity. Would someone of an Afro-Latinx background claim “Black” or “Hispanic” on their government documents? Would the same individual be estimated as being “Black” or “Hispanic” by a forensic anthropologist? The term “Hispanic” is an American term used to describe individuals from Spanish-speaking countries (e.g., Cuba, Costa Rica, and Mexico). It does not encompass the diversity of these nations. Furthermore, Hispanic individuals make up the second largest population in the United States yet skeletal collections lack robust samples (Spradley 2013). Similarly, the term Asian is also an oversimplification of the various countries and ethnicities of the continent. Genetically as well as cranial and dental morphology varies for individuals born in the North (i.e., China) and the Southeast (i.e., Cambodia). While these terms are commonly used, they gloss over the diversity that is present. Census data predicts that the number of individuals born of “recent mixed ancestry,” will increase significantly by 2050 which will influence racial and ethnic boundaries (Perez et al. 2009). However, the majority of the population follows positive assortative mating patterns (Robinson et al. 2017). Accounting for social rules, people mate with people who share phenotypic similarities. Because of this, it is possible to estimate affinity due to the interconnectedness of population history and cranial morphology (Spradley 2021). These bureaucratic racial categories demonstrate the

complexity of affinity and social identity because while an individual may check a box, there may be a more specific, representative group that they use in daily life. Three-group models and census categories do not account for the diversity that exists within humans.

FORDISC 3.1 (Jantz and Ousley 2005) currently has 13 reference sample groups. However, the scope of each differs. For example, White is a race denoting light or fair skin complexion; it is not an ethnicity or a nationality. On the other hand, Black or African American is both a race and ethnicity as there is a shared cultural experience as well as phenotypic characteristics. The term Hispanic signifies a shared language, Spanish but also has a geographic context (Latin America). This term condenses the variation and diversity of Central and South America, so is it helpful? Especially considering the proposed research questions surrounding self-identity? FORDISC also has a Guatemalan reference sample, which further complicates the terminology of affinity estimation; when does someone stop being Guatemalan and become Hispanic? FORDISC also has 4 nationalities included (Guatemalan, Japanese, Chinese, and Vietnamese). These are all different levels of population. So, how does terminology affect anthropological assessments? Should we standardize terminology and if so, how?

Accuracy rates of affinity estimation hold merit only if it expedites the identification process; if anthropologists use categories that are not relevant to society, it may impede said process. Assessing the accuracy rates of affinity estimation using the three-group model does not accurately reflect contemporary society and how people self-identify. More representative samples and methods are needed to account for the growing diversity. Consequently, cross-referencing an individual's self-identified race with the population affinity estimation can provide an insight into the applicability of population

affinity estimation outside of the forensic context related to social identity (i.e., a person's sense of who they are based on their membership(s) in relevant groups such as race or sexual orientation), (Tajfel and Turner 2010). Because anthropologists aim to help identify deceased individuals, understanding how people identified themselves in life is of note.

Hahn and colleagues (1996) assess the reliability of population affinity estimation in various data forms (e.g., self-reported identity and funeral director death certificates). They argue that understanding population affinity is critical to health status and compare population affinity assignments to demonstrate discrepancies between self and proxy reported estimates. Hahn et al. concluded that individuals with admixed heritage, such as Native American and Hispanic individuals in New Mexico, tended to change what race they selected on government documents throughout life (1996). This phenomenon highlights that while forensic anthropologists may accurately estimate the population affinity of an individual's remains, we do not consider the self-identification and how that may affect reports and positive identifications. A biocultural approach is necessary to population affinity estimation because its very nature is a mixture of social, historical, and biological concepts. Understanding population genetics and history is essential in affinity estimation.

Because race is a social construct, my study inherently has implications for contemporary society. Ancestry is not based on concrete biological differences but is a conglomeration of psychological, cultural, and physical characteristics. The explicit concept of self-identity is missing from forensic anthropological literature. While it is inherent to population affinity estimation, anthropologists focus on the concordance of

self-identity to skeletal morphology; not what it means to the individual. However, the way in which an individual sees themselves inherently affects several aspects of life. Comparing self-reported racial identity with anthropologically assessed reports may give insight into what affects how ancestral identity and social race are perceived by the public and forensic anthropologists. It is not enough to only estimate affinity. We as a field must consider that an individual may not identify with our estimated population affinity and how this may implicate identifications. To end that, my research questions are:

1. How do forensic anthropological population affinity estimates compare with donor-specific identity? Identity is arguably more complex in the TXSTDSC and its context (i.e., Texas) because of the larger portion of Hispanic individuals as well as Central and South American migrants.
2. Within the United States, how do people self-identify?
3. Are there trends with who is misclassified based on population, sex, or any other factor?

Recently, members of the field have called into question the validity and ethicality of affinity estimations. DiGangi and Bethard argue for the abolishment of the practice because it is inherently linked with typological and racist ideologies. The authors also state that search efforts decrease when an unidentified individual is estimated to be a member of the BIPOC community (DiGangi and Bethard 2021). However, the purpose of the biological profile is to provide law enforcement agencies with the resources necessary to narrow down possible identifications. So, while it is argued to potentially cause a delay in identification, many forensic anthropologists view it as an important factor of the profile. The purpose of this research is not to argue for or against either camp, but rather

to provide quantitative results regarding accuracy so that practitioners can be well-informed about the concordance between the population affinity estimates and the self-identity of the decedents.

The literature review demonstrates that although there is ample information available on forensic population affinity estimation methodologies, there is little research that addresses the direct impact self-identity has on the accuracy rates of estimations. This thesis will add to the body of knowledge available on affinity estimation. It explores supplemental information for populations often misestimated. This research has implications for the analysis of unidentified human remains because it compares self-identification and may provide information that impacts the field of biological anthropology.

II. MATERIALS AND METHODS

Data Collection

To address the question of how forensic anthropological population affinity estimates compare with donor-specific identity, craniometric data was collected on individuals of the Texas State Donated Skeletal Collection (TXSTDSC). Because the cranial complex is the most informative region of the skeleton related to population affinity and several validation studies exist, this is an appropriate method (Jantz and Ousley 2005). Additionally, craniometrics is arguably the most objective, validated, and widely used methodological approach to assess populational differences when paired with biological distance statistics. Spreading and sliding calipers were used following the standards outlined in the Data Collection Procedures for Forensic Skeletal Material 2.0 handbook (Langley et al. 2016). Caliper based measurements were used because they cost less than MicroScribe digitizers and are therefore more accessible to practicing anthropologists. Values were recorded in millimeters (mm).

All standard measurements are included except for zygomaxillary breadth (ZMB), biasterionic breadth (ASB), mid-orbital width (MOW), and mastoid height (MDH) for a total of 23 measurements (Figure 1, Table 1). Three of the excluded measurements (ZMB, ASB, and MOW) are recent additions to the Forensic Data Bank (FDB) collection and drastically reduce sample sizes, especially in BIPOC groups, and mastoid height has a high interobserver error rate (Langley et al. 2016); for both of these reasons the measurements were excluded. For any bilateral measurement (e.g., orbital breadth), the left side was included in the analysis unless trauma prevented such; in those instances, measurements from the right side were included (Langley et al. 2016).

Table 1. Standard cranial measurements used for analysis and their abbreviations.

Measurement	Abbreviation
Maximum Cranial Length	GOL
Nasal Height	NLH
Maximum Cranial Breadth	XCB
Nasal Breadth	NLB
Bizygomatic Breadth	ZYB
Orbital Breadth	OBB
Basion-Bregma Height	BBH
Orbital Height	OBH
Cranial Base Length	BNL
Biorbital Breadth	EKB
Basion-Prosthion Length	BPL
Interorbital Breadth	DKB
Max. Alveolar Breadth	MAB
Frontal Chord	FRC
Max. Alveolar Length	MAL
Parietal Chord	PAC
Biauricular Breadth	AUB
Occipital Chord	OCC
Upper Facial Height	UFHT
Foramen Magnum Length	FOL
Minimum Frontal Breadth	WFB
Foramen Magnum Breadth	FOB
Upper Facial Breadth	UFBR
Mastoid Height	MDH
Biasterionic Breadth	ASB
Midorbital Width	MOW
Zygomaxillary Breadth	ZMB

Source: FORDISC Help

(https://www.statemachine.net/software/Fordisc/support/Help/Fordisc3_Help.pdf)

Demographics of the Texas State Donated Skeletal Collection

This research was conducted using remains from donors from the Texas State Donated Skeletal Collection. Therefore, it is important to understand the collection's demographic composition. The TXSTDSC is a whole-body donation system at the

Forensic Anthropology Center at Texas State (FACTS). Individuals can pre-register or family members and next of kin can donate remains (Gocha et al. 2021). Donors must meet specific requirements. The maximum weight is 500lbs and any person with an active infectious disease, such as hepatitis B and C, is not accepted (Gocha et al. 2021). Once an individual is accepted, they are required to answer a three-page questionnaire that includes comprehensive demographic information (e.g., biological sex, adult socioeconomic status, etc.) and lifestyle/habitual activities, such as frequency of alcohol and tobacco use. As of November 2021, the TXSTDSC has received 710 donors. The demographic breakdown of the collection is as follows:

“639 donors (90%) identified as White (non-Hispanic), 32 donors (4.5%) identified as Hispanic, 22 donors (3%) identified as Black or African American, two donors identified as Native American, two donors identified as Asian (one Indian and one Laotian), and one donor identified as Lebanese. The remaining 12 donors identified as multiracial or of mixed ancestry,” (Gocha et al. 2021).

Biological males are overrepresented at 58% with biological females representing the remaining 42%. Regarding age, the collection is predominantly composed of older adults with the mean age-at-death being 66 years old (Gocha et al. 2021). Using individuals with known affinity assesses how well anthropological methods account for diversity in estimations.

For the purpose of this research, I requested a list of individuals to be equally represented across the demographics of the collection. My list included all BIPOC donors that were relatively equal in sex distribution, included younger individuals if possible, and only a subset of White individuals. Only adult individuals with complete cranial remains were examined. Those that exhibit severe trauma that can impact the cranial landmarks were excluded. Because prosthion is a diagnostic landmark in several key measurements, individuals who exhibited edentulism or any trauma were not included.

These exclusions allow me to use the same suite of traits for each donor because my question is how FORDISC compares to self-identity, not its mathematical accuracy. The sample includes a random sample of White individuals (due to their overrepresentation) and every BIPOC donor that met the criteria. The final sample size was $n = 77$, with 48 males and 29 females (Tables 1 and 2). To ensure a blind analysis, I had no access to the donor forms that contain the demographic information prior to data collection. Only after the completion of my FORDISC analyses, did I learn of the demographic breakdown of my sample regarding sex and population groups.

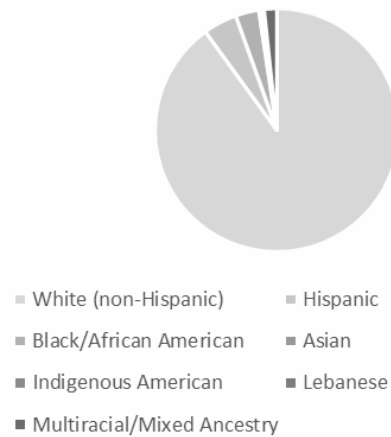


Figure 1. TXSTDSC Donor Information (sex and race).

Table 2. TXSTDSC Donor Information (sex and race).

Donor Number	Sex	Race (From Donor Form)
2015.053	Female	American Indian or Alaskan Native
2017.049	Male	American Indian or Alaskan Native
2014.025	Male	American Indian or Alaskan Native, Hispanic, White
2018.055	Male	American Indian or Alaskan Native, White
2019.008	Male	American Indian or Alaskan Native, White
2012.034	Male	Asian (Indian)

2010.014	Female	Black
2012.010	Female	Black or African American
2016.022	Female	Black or African American
2017.065	Female	Black or African American
2019.010	Female	Black or African American
2011.016	Male	Black or African American
2012.037	Male	Black or African American
2013.063	Male	Black or African American
2014.008	Male	Black or African American
2014.013	Male	Black or African American
2014.052	Male	Black or African American
2017.055	Male	Black or African American
2014.053	Male	Chinese, Hispanic, White
2012.040	Female	Hispanic
2013.038	Female	Hispanic
2014.001	Female	Hispanic
2014.028	Female	Hispanic
2017.045	Female	Hispanic
2017.061	Female	Hispanic
2018.053	Female	Hispanic
2010.008	Male	Hispanic
2010.015	Male	Hispanic
2012.011	Male	Hispanic
2012.014	Male	Hispanic
2014.054	Male	Hispanic
2015.048	Male	Hispanic
2015.062	Male	Hispanic
2017.022	Male	Hispanic
2017.067	Male	Hispanic, White
2016.043	Male	Laotian
2013.012	Female	Lebanese
2011.012	Female	White
2012.032	Female	White
2012.036	Female	White
2012.049	Female	White
2013.017	Female	White
2014.011	Female	White

2014.021	Female	White
2015.005	Female	White
2015.012	Female	White
2015.026	Female	White
2015.060	Female	White
2016.008	Female	White
2016.031	Female	White
2016.046	Female	White
2018.004	Female	White
2009.003	Male	White
2010.007	Male	White
2010.010	Male	White
2011.001	Male	White
2011.014	Male	White
2011.015	Male	White
2012.013	Male	White
2012.016	Male	White
2012.020	Male	White
2012.025	Male	White
2012.038	Male	White
2013.045	Male	White
2014.015	Male	White
2014.048	Male	White
2014.065	Male	White
2015.008	Male	White
2015.013	Male	White
2015.029	Male	White
2016.021	Male	White
2016.030	Male	White
2016.056	Male	White
2018.003	Male	White
2018.027	Male	White
2018.039	Male	White
2019.034	Male	White

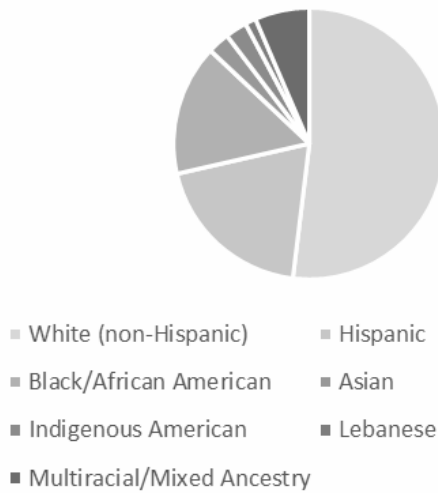


Figure 2. Breakdown of Donors by Self-Described Race.

Table 3. Breakdown of Donors by Self-Described Race.

Population/Race	Number of Donor
Indigenous American	5
Asian (Indian)	1
Black/African American	12
Hispanic	15
Laotian	1
Lebanese	1
White	40
Multiracial	2
Total	77

Demographics of the Forensic Data Bank and FORDISC

This research will be analyzed using FORDISC 3.1, which compares the unknown to the reference samples found in the Forensic Data Bank (FDB). Therefore, it is important to understand the demographic composition of the FDB. There are 13 population groups in FORDISC: American Black (male and female), American Indian

(male and female), American White (male and female), Hispanic (male and female), Japanese (male and female), Chinese (male only), Guatemalan (male only), and Vietnamese (male only). Except for the American Indian group, all the individuals comprising the reference samples are born after 1930 (Jantz and Ousley 2005). Some of the names of these groups do not necessarily reflect how an individual would self-identify and are antiquated. For the remainder of this research, Indigenous American will be used instead of American Indian, and American Black will be written as Black/African American as seen in Table 1. Table 4 shows the breakdown of the Forensic Data Bank by population and sex (Jantz and Ousley 2005).

Table 4. Forensic Data Bank groups included in FORDISC 3.1 by sex and total.

Forensic Data Bank Groups			
Population	Male	Female	Total
American Black	224	137	361
American Indian	59	32	91
American White	737	454	1191
Chinese	80	0	80
Guatemalan	83	0	83
Hispanic	281	74	355
Japanese	84	58	142
Vietnamese	51	0	51

Source: Jantz and Ousley (2005)

The following information about the sources of each sample is found in the FORDISC 3.1 Help file (Jantz and Ousley 2005). The Black/African American group is composed of African Americans from different states in the southeast and mid-Atlantic region, including the Terry Collection. The Indigenous American sample consists of forensic casework from the Southwest. The American White group comes from across

the country and is primarily Euro-Americans with some European-born individuals. The Chinese male group is composed of Hong Kong University cadavers from the 1970s. The Hispanic reference group comes from New Mexico and the Pima County Medical Examiner's Office, which deals with a large number of undocumented migrant deaths (primarily from Mexico). None of the samples are from the Caribbean. The Japanese sample includes individuals born in the 20th century. The final group, the Vietnamese sample, comes from a "Killing Fields" massacre site in Vietnam near the Cambodian border. It is of note that the Asian samples are country-specific whereas Hispanic is a broad term that denotes a continent. The difference between these groups may affect classification rates.

Of the 77 donors included in this study, 8 do not have a reference sample in FORDISC or are of multiracial background. The three individuals who do not have a reference sample are Laotian, Lebanese, and Indian and 5 multiracial individuals of differing backgrounds. I expect these individuals to have low classification rates if they classify at all.

Data Analysis

The data was input into FORDISC 3.1 (Jantz and Ousley 2005), a software that employs discriminant function analysis to classify unknowns into population affinity and sex groups using standardized measurements (Ousley and Jantz 2012) and compared to the self-reported identity on donor sheets. The software version used in this research is FORDISC 3.1.320.

FORDISC uses craniometric data to statistically compare an unidentified cranium

to known reference groups through discriminant function analysis (DFA) to provide posterior probabilities for population affinity estimations (Ousley and Jantz 2012). It maximizes between-group differences and minimizes within-group variation. A posterior probability is the probability or likelihood of membership for the unknown individual in each group based on the relative distances to the groups, and they all sum to 1. These probabilities are based on the assumption that the unknown individual belongs to one of the reference groups. An output over 0.8 is considered a high posterior probability. Typicality probabilities are an indication of group membership. FORDISC calculates three typicalities: F-distribution, Chi-square distribution, and ranked Mahalanobis distances (Jantz and Ousley 2005). These demonstrate how typical an unknown is for a particular group, based on the average variability of all the groups in the analysis. DFAs classify unknown individuals into a particular group (such as a sex or ancestry group) based on similarity to the known reference. Should an individual not belong to any of these groups, the program may still provide a classification (Ousley and Jantz 2012). However, if an individual is too dissimilar to all of the included reference group samples, the unknown individual will not classify (Ousley and Jantz 2012). The typicality probability is provided to indicate when a practitioner should not classify an unknown. FORDISC is a tool that users can use to assess population affinity. Because the input data can be manipulated, an analysis protocol was developed to ensure the data is not skewed and to mitigate any potential bias.

FORDISC Protocol

1. For all samples, the initial analysis is run with all populations (male and female) because the focus is not sex estimation and the program has more

male reference groups. If there is an apparent cluster of one sex (e.g., all male populations classify with 0.0 posterior probabilities), those with low values will be removed from the remaining analyses.

2. Any measurement that is one or more standard deviations from the mean was remeasured. If they are correct, they remain included in the analysis. If not, the measurement was changed.
3. Any measurement that is more than 3 standard deviations from the mean is excluded from the remainder of the analyses regardless of if they are correct.
4. Regarding posterior probabilities, > 0.8 is considered the threshold for confident estimations.
5. Low posterior probabilities (e.g., 0.01), were removed.
6. The threshold for typicalities is anything greater than or equal to 0.05 as this represents the individual is typical of 95% (of the group).
7. If an individual's classification is split between 2 or more groups relatively evenly, the results are recorded as such.
8. If an estimation is split between Hispanic and Guatemalan, they are condensed into a single group (Hispanic) as this is the official term for U.S. documents.
9. Forward-Wilks' L Stepwise regression was used when donors were not typical of any group. This allows for a more conservative estimation.

Any scenarios not mentioned in the above protocol were analyzed on a case-by-case basis. Following FORDISC analysis, the percentages of each sample were recorded.

The two classifications with the highest posterior probabilities were then put into a table for each individual. Group membership, or a strong classification, was considered concrete for posterior probabilities of 0.8 and above. For those with split classifications, each population was recorded for comparison as this may indicate the individual was of two ancestral groups (i.e., multiracial or mixed ancestry) or that their group is not included in the FORDISC reference samples, such as the Laotian individual.

The FORDISC outputs were then compared to the donor form. For the comparison of self-identity to the forensic estimates, I quantified the amount of correct positive identifications with self-reported population affinity. I organized the data by population affinity to provide percentages of correct estimates and corresponding self-identity to the rates for each population. In categorizing the data, I demonstrate how correct forensic anthropologists are and provide trends based on the groups.

The complexities of this topic were further explored by analyzing the correct and incorrect classifications by self-reported identity. For estimations that are different from the individual's self-identity, each incorrect assessment was recorded. Then, the frequency of each group per population was calculated. By evaluating which individuals and which groups are more likely to be misclassified, anthropologists will understand which populations are lacking from the reference samples and further the dialogue on inclusive estimation techniques. I also explore any patterns of factors causing misclassification. Specific measurements that have a significant effect will be studied more in-depth. Considering those who do not fall within the 90% confidence interval or whose posterior probability is under 0.8 were reexamined. I determined what populations are generally assessed correctly and those which have high misclassification rates.

Because my research addresses how correct forensic anthropologists are at estimating population affinity when compared to self-identification, the chosen data are appropriate in answering it since my analysis is based on assessing individuals with known identities.

Limitations

In all research, there are limitations that must be addressed. My sample size is small in relation to the potential identities in the U.S. and the current population. However, it is appropriate for the proposed questions. Additionally, because the TXSTDSC has a 100-mile pickup radius, the number of individuals who would like to donate their remains versus those who do so is biased favoring people who live in Texas, which is not representative of every group in the United States. Furthermore, FORDISC 3.1 only contains thirteen reference samples with White males being significantly overrepresented. The Indigenous American or Alaskan Native group contains historic samples, which can vary from contemporary populations. “Hispanic” refers to an ethnicity that is only used in the U.S.; it does not capture the diversity of Central, Latin, and South America. The program is also only appropriate for its included populations. Regarding self-described or next of kin reported identity, the records provided to me by the TXSTDSC do not indicate which donors self-reported. Thus, there is an inherent bias because an individual may have considered themselves something different than what their family reported. Additionally, while the TXSTDSC questionnaire provides a space to write in one’s race/ethnicity, the information provided to me only includes the U.S. Census categories. Ultimately, these limitations have implications for this project, but the results still answer the proposed research questions.

III. RESULTS

Summary of Results

Of the 77 individuals, 50 (64.9%) were classified into the population reported on the donor forms (Table 2). 24 of the 50 (48%) met or exceeded the posterior probability threshold of 0.8. All of the donors met the typicality threshold (0.05) for at least one group; none were too dissimilar to classify. Because my focus was solely on population affinity, groups with sex-specific samples were condensed into a single population (e.g., Hispanic males and females are combined into Hispanic), however, a breakdown by sex is provided.

American White individuals had a correct classification rate of 87.5% with 35 of the 40 individuals classifying as White in FORDISC. They were the group with the highest correct estimates. The group with the second highest correct classification (83.3%) is Black/African American. 10 of the 12 Black/African American individuals classified correctly. However, the posterior probabilities were not strong.

Hispanic individuals have a classification rate of 35.7% (5 of the 14). However, of the 9 donors that misclassified, two classified into the Indigenous American group which shares a similar population history. Two Indigenous American donors were included, and both were misclassified into White (0%). Similarly, Asian donors represented a small portion of the sample. The Laotian individual classified into Chinese and the Indian donor into Japanese. So, while FORDISC did not have the exact reference samples, the classification rate for the Asian individuals is either 0% or 100%, depending on interpretation. Additionally, the Lebanese donor classified as White and so according to

the U.S. Census categories, the estimation is correct. However, the individual did not self-identify as White and so it is incorrect.

The final group was people of multiracial/mixed ancestry. Five donors considered themselves as multiracial. There are no clear trends in misclassifications for these individuals. Three of the five did classify into one of the parental groups (60%). The following sections detail the results from the analysis for each group. The subsequent sections contain classification rates for each population and discuss the results in the above table.

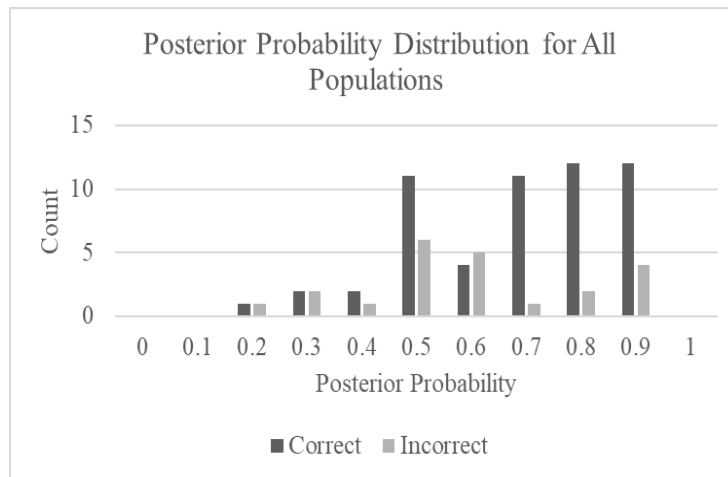


Figure 3. Posterior Probability Distribution for All Populations.

FORDISC Classifications by Populations

White Individuals

This study included forty-one (males = 25 and females = 14) White individuals (including one Lebanese following U.S. census categories) according to the TXSTDSC donor forms, with twenty-five males and fourteen females. However, because the

Lebanese donor did not self-identify as white, they were removed from the White sample. So, the total number of White individuals is 40 (males = 25 and females = 13). Of the 40, 35 individuals correctly classified (87.5%) as White with a mean posterior probability of 0.737, this mean value indicates strong confidence in the statistical analysis's ability to discern them from other groups. When the sexes are separated, the classification rate increases for males and decreases for females. Both groups perform relatively well within the confines of the FORDISC framework. This makes sense when you consider that White individuals have historically been the majority population in the United States and thus, were prioritized by researchers. Overrepresentation is not inherently bad; it allows for more variation to be captured, which ultimately betters estimation methodologies. So, the classification rate for White donors was expected.

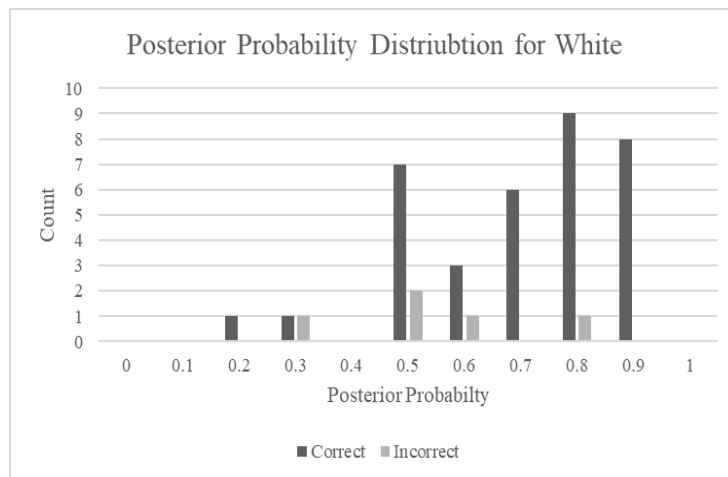


Figure 4. Posterior Probability Distribution for White Individuals

Black/African American Individuals

Data from 12 Black or African American individuals were recorded for this analysis. Using FORDISC, ten of the twelve (83.3%) were correctly classified making

this population the second highest of correct classifications. However, the posterior probabilities did not meet the 0.80 threshold to be considered a strong classification. The mean posterior probability is 0.6589. This indicates that FORDISC is able to correctly classify Black/African American individuals, but the classifications themselves are not strong/confident. This may be due in part to the analyses being run against all populations, not sex-specific. Additionally, the FDB only has 361 individuals, which does not encompass the diversity within this population. The two individuals that misclassified varied in their estimated group membership. Donor 2010.014 classified as a Vietnamese Male with a posterior probability of 0.285 followed by Black/African American Female with a posterior probability of 0.214. Donor 2013.063 classified as a White Male with a posterior probability of 0.614 followed by Indigenous American Male with a posterior probability of 0.337. Both individuals met the typicality threshold of 0.05.

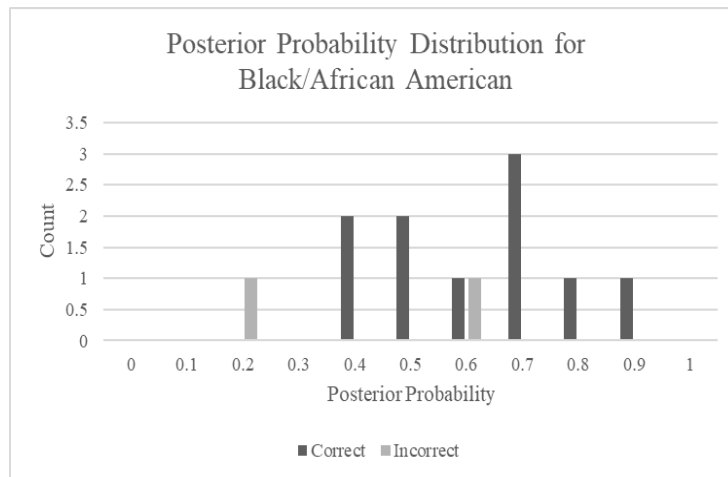


Figure 5. Posterior Probability Distribution for Black/African American Individuals

Hispanic Individuals

Hispanic individuals make up 4.5% of the TXSTDSC (Gocha et al. 2022). Fifteen self-ascribed Hispanic individuals were included in this research. Hispanic individuals correctly classified 35.7% (5 of the 14) of the time, but the posterior probabilities were relatively low. The group the Hispanic individuals misclassified the most is Black/African American (5 donors), followed by Indigenous American (2 donors) and White (2 donors). The final misclassified individual classified as Japanese. Females classified correctly more frequently than males. Dudzik and Jantz (2016) examined the misclassification rates for Hispanic groups and found that Japanese crania share morphological similarities. As such, they often misclassify as one another (Dudzik and Jantz 2016). For this reason, as well as the Texas demographic context of these cases, analyses of donors whose classifications split between the two populations had the Japanese groups excluded.

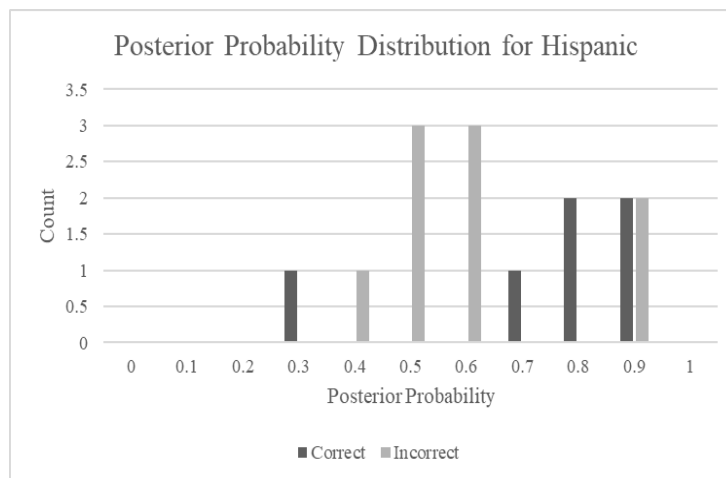


Figure 6. Posterior Probability Distribution for Hispanic Individuals

Indigenous American Individuals

Two donors self-identified as Indigenous American/Alaska Native. Neither were correctly classified in FORDISC; both classified as White (0%). Additionally, two donors considered themselves to be Native American and White. The highest posterior probability for both was White followed by a North Asian group (Chinese for donor 2018.055 and Japanese for 2019.008). This is consistent with the peopling of the Americas when considering the level of admixture. However, this would not aid the identification process. The fifth and final donor who identified as Native American also reported having Hispanic and White heritage. Donor 2014.025 classified into Japanese and then White.

Asian Individuals

As of November 2021, only two TXSTDSC donors are of Asian descent, one Indian and one Laotian. Both were included in the analyses. Donor 2012.034, while Indian, classified as Japanese with a posterior probability of 0.568. The second highest classification was Chinese (0.128). The classification for donor 2016.043 was split between Chinese (p.p. 0.579) and Japanese (p.p. 0.392). The mean posterior probability is 0.5735, which does not meet the confidence threshold. However, when the posterior probabilities are added, Donor 2012.034's total posterior probability for Asian groups increases to 0.696 and Donor 2016.043's combined posterior probabilities total 0.971. Both individuals classified relatively strongly as Asian but neither population has a reference sample in FORDISC 3.1. The classification rate is either 0% or 100% depending on the interpretation.

Multiracial/Mixed Ancestry Individuals

Twelve donors of the TXSTDSC identified as multiracial or of mixed ancestry. Of these, five were included in this study. Two donors (2018.055 and 2019.008) self-identified as American Indian or Alaskan Native and White. One donor (2014.025) self-identified as American Indian or Alaskan Native, Hispanic, and White. Donor 2014.053 self-identified as Chinese, Hispanic, and White on their donor form and donor 2017.067 self-identified as Hispanic and White.

There is no pattern to how these donors classified. Two donors self-reported as Native American/Alaskan Native and White. Both classified as White followed by Asian. One individual described themselves as Native American/Alaskan Native, Hispanic, and White and classified as Japanese with a posterior probability of 0.523 followed by White Male 0.248. However, due to the low posterior probabilities, a practicing anthropologist should not report these results as estimations as they do not meet the threshold for confident estimations. Another donor self-reported being Chinese, Hispanic, and White. When analyzed in FORDISC, they classified as Black/African American with a strong posterior probability (0.727). The final multiracial donor classified as Hispanic female (p.p. 0.378) and Hispanic male (0.307). However, when run against only male reference groups, the posterior probability and typicality scores increased to 0.922 for Hispanic Male, followed by Guatemalan male (p.p. 0.058) and White male (p.p. 0.020). Using the sex-specific reference groups, the classification meets the threshold. This individual self-identified as both Hispanic and White, but the FORDISC classification favored Hispanic. So, of the 5 multiracial individuals, 3 (60%) classified as one parental group while the

other 2 did not. This highlights the need for collections to include individuals of mixed backgrounds.

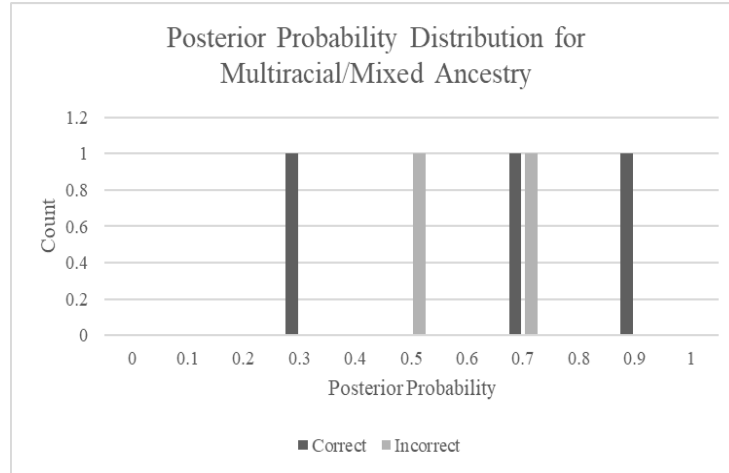


Figure 7. Posterior Probability Distribution for Multiracial/Mixed Ancestry Individuals

Classification Rates by Sex

Of the 51 with correct classifications, 31 were male and 20 were female. Males classified correctly 64.5% of the time (n = 48) and females classified correctly 69% (n = 20) of the time. Because males are overrepresented in the collection and FORDISC samples, I hypothesized they would classify better than females. However, the opposite is true. This could be because when an unknown individual is run against the male reference groups, there are more classification options. Nonetheless, there is no significant difference in the rate of classification by sex. These results will be discussed in the following chapter in relation to the research questions. Suggestions on future research are also provided.

Misclassifications

33.8% of the donors included in this research were misclassified. These results are the inverse of the correct classifications. The group with the highest rate of misclassifications is Indigenous American (100%). The group with the second highest misclassification rate is Hispanic (64.3%), followed by the multiracial/mixed ancestry individuals with a classification rate of 40%. This value is based on three of the five classifying into one of their parental groups, but the misclassification rate is 100% when considering that no individual was classified into all of their appropriate parental groups. The misclassification rate for White individuals is 12.5% and is 16.7% for Black/African American individuals. For the two Asian groups, the misclassification rate is 0% if you interpret their results to suggest they are of Asian descent and not hinder identification efforts. If not, the rate is 100%.

When considering why some groups classified so well, the same is true for those that did not. BIPOC samples, including multiracial individuals, are less represented in skeletal collections. Therefore, the variation within a population is less likely to be captured, affecting classification rates.

IV. DISCUSSION

This thesis set out to answer three questions: 1) How do forensic anthropological population affinity estimates compare with donor-specific identity? 2) Within the United States, how do people self-identify? and 3) Are there trends with who is misclassified based on population, sex, or any other factor?

This research is a preliminary study on the accuracy of craniometrics and FORDISC 3.1 when used to estimate population affinity and demonstrates that current anthropological methods perform relatively well (64.9% correct) when used to classify unidentified skeletal remains. Regarding the first question, the performance of population affinity estimates varies by group. Second, it is too early to determine what if and how craniometric features affect self-identity for U.S. individuals. Because of the limited sample of multiracial individuals and their low classification rates, I found no consistent pattern indicating how and/or why an individual self-identified a certain way. Classification trends will be discussed in the following sections.

Results were particularly positive for individuals with a single ancestral background that also had a robust reference sample (e.g., White donors). This number (64.9%) considered the two Asian donors (Indian and Laotian) classifying into Asian reference groups as correct because their group is not present in FORDISC. On the contrary, the overall classification rate included the Lebanese individual classifying as White as incorrect; while this individual is considered White on the U.S. census, they did not report that on their donor form. If all three are considered incorrect classifications, the overall rate decreases to 62.3%. However, because the Asian classifications would not

hinder an identification, and because Lebanese is considered White for government purposes, they were considered correct. Thus, the overall rate of correct classifications is 64.9% (50 of the 77 individuals).

Craniometric assessments have a high classification rate when compared to self-reported identities, which impacts the field because it provides concrete, quantitative data for the question of ‘what we are really estimating?’. In the post-Daubert era where findings are expected to be statistically reliable, this corroborates that affinity estimation is useful when certain limitations are acknowledged. Nonetheless, groups that lack comprehensive reference samples are more likely to be misclassified. Additionally, issues with self-reported racial identity are not accurately reflected in the samples. For example, the Lebanese donor correctly classified using US census categories (White, non-Hispanic), but that individual did not claim that identity in daily life; they wrote in “Lebanese” on the donor form. FORDISC 3.1 has zero populations that represent Arab ethnic backgrounds in addition to others. So, in a forensic context, this may hinder the identification process.

Similarly, Native American populations have endured forced assimilation and many individuals possess admixed heritage. European colonization diversified the gene pool among other things. However, government documents, such as a driver’s license, do not allow for a person to select more than one racial category. As such, an individual may be of both Indigenous and European descent but choose to select the identity more obvious for society or the one with greater cultural significance (i.e., based on skin color). So, FORDISC misclassifications may be linked to these limitations as well as the fact that many of the Native American reference samples are historic. Cranial morphology

experiences secular change so there is a discordance when comparing modern individuals to temporally older individuals even if they have the same affinity. However, cultural practices may hinder the available samples. For example, several Native American tribes (Navajo and Sioux) participate in tree burials while others (e.g., Kiowa) bury their deceased beneath the ground. With the history of forced assimilation, the Trail of Tears, and other measures to disrupt Indigenous lifestyles, Native American individuals may be hesitant to donate their remains to predominantly White institutions. Further, the population history of Latin America could attribute to misclassifications. Individuals who consider themselves Hispanic may have a significant portion of Indigenous genetic ancestry, blurring the separation between groups. Gene flow and admixture may skew these classifications.

Additionally, people of multiracial or “mixed ancestry” do not classify as well as individuals of single ancestral populations. This is because discriminant function analysis has to classify a sample into a single group. There is no way to know for certain the number of people in existing collections who are of mixed ancestry because of a plethora of reasons including but not limited to social restrictions, personal preference, or lack of knowledge/ignorance. Being White was the only acceptable race for a large portion of U.S. history therefore individuals often claimed one racial identity and not the other. So, while we assume reference samples contain people from isolated populations, there is no way to be certain, which directly influences our analyses. That being said, all of the multiracial individuals included in this study have reference samples for their ancestral groups in FORDISC. I expected their classifications to mimic their self-reported group

membership; it was interesting to see that there was no apparent pattern in their classifications.

Regarding the low classification rate of Hispanic individuals (35.7%), there are several potential factors. These analyses did not include priors (i.e., sex), however, several methodological approaches are population-specific (Spradley 2016). Because of this, re-running the data against the sex of each of donor, may increase the classification rate. Another factor may be related to the Hispanic sample's composition in the FDB. The Hispanic individuals included in FORDISC's samples were supplied by the Pima County Medical Examiner's Office as well as New Mexico with the majority of individuals being Mexican migrants. The context of the TXSTDSC Hispanic donors is different; some are immigrants, and others are citizens by birthright. Additionally, in Texas, undocumented migrants are from several Central American countries (e.g., Guatemala, El Salvador, and Mexico). So, the cranial morphology may be different because these are different populations. The term "Hispanic" is broad and does not capture the different nationalities and ethnicities – or the diversity within these groups – making it problematic when considering samples.

To continue, this study demonstrates that there is no concrete distinction between correct and incorrect classifications. FORDISC 3.1 outputs are supposed to inform an anthropologist's interpretations. For example, Donor 2016.043 self-described as a Laotian Male. FORDISC analysis produced a posterior probability of 0.579 for Chinese Male, 0.392 for Japanese Male, and 0.028 for Vietnamese Male. These posterior probabilities add up to 0.971, an extremely significant value if the populations are condensed into Asian. The Indian individual (donor 2012.034) does not have a reference

sample group in FORDISC. Again, there may be cultural factors affecting the donation rates. For example, an estimated 77% of India practices Hinduism, a religion where all but saints and children are cremated (Majumdar 2018). So, while anthropological collections lack substantial samples of Indian individuals, a large majority of the population may believe in different death practices, limiting the number of interested donors. Because FORDISC does not contain samples from every single population, a forensic anthropologist may interpret these results to mean that the individual is likely Asian. Case context would also be used to estimate the population affinity. So, while this estimation does not capture that the individual was Laotian – or even Southeast Asian (i.e., similar to the Vietnamese sample), the results would likely *not* have negatively impacted attempts made by law enforcement to match potential identifications. That being said, FORDISC does not have a Laotian reference sample, so it was impossible to correctly match the estimate to the donor’s self-reported affinity.

As hypothesized, White individuals had the highest classification rate. White donors make up approximately 50% of the FORDISC sample (Table 1) and 90% of the TXSTDSC. The overrepresentation is evident throughout both datasets. Furthermore, anthropometric studies prioritized White populations followed by Black populations for most of the discipline’s history, with some pseudoscientists searching for racial superiority. But, another reason for the overrepresentation is that White Americans comprised the largest proportion of the population for decades. Prioritization of the data collection of any group allows for more of the variation to be captured which allows for better estimation methodological approaches and ultimately increases classification rates. More data means more variety, which allows for a more significant range of variation and

therefore there is more knowledge when facing an unknown. So, there is more data and literature available. Simply put, the available Asian and Hispanic samples are smaller than that of American White, and Black groups. Although, the only group that had a high percentage of correct classifications and high posterior probabilities is the White donors. So, there is room for improvement for all other groups. Large, diverse reference samples ensure confidence when faced with future unknowns or data.

To continue, anthropologists should consider the context of unidentified human remains when estimating population affinity. For example, Texas and the U.S. have a large Hispanic population (62.1 million people as of 2020) (Krogstad et al. 2022) and a smaller Japanese population (1.498 million people as of 2019) (Budiman 2021). Therefore, removing the Japanese reference groups from FORDISC analyses reduces misclassification rates for groups that share similar population history and structure (Dudzik and Jantz 2016). Because population affinity estimates are a tool to aid in the identification process of missing or unidentified individuals, understanding the demography of an anthropologist's region is imperative. When employing FORDISC, understanding what populations to include in your analyses increases the likelihood of classification.

Regarding sex, there is no significant difference in classification rates. Females were correctly classified into their donor-specified population 69% of the time while males were 64.5%. Because males are overrepresented in the FDB, TXSTDSC, and other skeletal collections, I hypothesized that males would have a higher classification rate and in general, classify better. However, the data indicates there is no apparent pattern between the sexes; both classify relatively well.

The FDB and other institutions need to prioritize supplementing existing collections to include Middle Eastern/Arab populations, more-specific Hispanic and Asian groups, and females for existing reference groups. However, FORDISC relies on data submissions from other anthropologists. So, while there is a big push to diversify collections and samples, data sharing is needed. Anthropologists need to actively submit new reference samples to programs like FORDISC if change is to happen. A limitation to capturing global diversity in collections is cultural norms and mistrust of former imperial nations. For example, the primary religion in India is Hinduism and a common death practice is cremation. Therefore, it is difficult to acquire skeletal samples from that region. Researchers need to rely on in-country members of the discipline to procure and collect samples, or at the very least cranial data to propel research. It is necessary to make space for scientists who are community members of these groups. For these reasons, forensic anthropology must adapt in its methods for data collection. Investing in new and different technologies, such as computed tomography (CT) scans, is the future. Cranial information can be collected on living individuals who otherwise may not have donated their remains. There are resources available that will aid in the field's quest to transform skeletal collections to be more representative of society's demography.

Future Research

This research was conducted using blind analysis. While this type of analysis was appropriate for the project, priors are not a bad thing. Demographic information and other priors are informative and can aid the identification process. Anthropologists typically estimate sex before assessing population affinity so by running these analyses using the known sex of each donor, the individuals and populations who have poor classification

rates may perform better. Because there is a dependence among the parameters and certain estimation methods are population and/or sex-specific, reanalyzing the data with donor demographic information may provide higher classification rates.

While this work demonstrates the utility of craniometric population affinity estimations, we as researchers and practicing forensic anthropologists should always consider the social and cultural implications of said assessments. We are approaching a paradigm shift in the way we classify and report affinity. We have a duty and a responsibility to educate law-enforcement personnel, as well as the public, on what our terminologies and findings actually report. It is not enough to simply change the name from ancestry to affinity without ensuring that the public grasps what we are referring to. Outreach is necessary as well as making articles and data freely available. Data sharing will also help diversify the field. Additionally, publishing brief communications in plain speech could help dispel any myths regarding racial differences. Continuing to include affinity in skeletal analyses and reports holds merit. However, anthropologists must be cautious as our work has often been used to propel racist and dangerous ideas, especially within the United States. In the technological golden age, researchers need to prioritize accessible and freely available publications in order to counteract misinformation the general public may have. Self-reported identity and craniometrically assessed population affinity estimates correlate, but that is not to say that there are no negative consequences for reporting said results (DiGangi and Bethard 2021).

This research emphasizes the need for more diverse, larger samples and estimation practices. Comparing widely sampled populations (e.g., American White) with traditionally underrepresented groups (e.g., Latin American, Asian, and Indigenous

populations) should generate additional research focused on supplementing existing standards to account for more diverse samples. Consequently, anthropologists should expand their estimation techniques to include these previously excluded groups to increase positive identifications. Anthropologists must prioritize collecting data from underrepresented groups. However, a biocultural approach is necessary as there may be underlying social factors preventing groups from donating their remains (e.g., religion). Therefore, outreach focused on educating the public on the benefits of donating to science is essential so long as ethical standards on informed consent are met. Further, anthropologists continue to develop methods for the application of CT scans and other technology for skeletal data collection. This is another potential avenue to increase samples that reflect the variation in cranial morphology.

Forensic anthropologists are tasked with creating a biological profile for unidentified skeletal remains which includes population affinity. Census data show that there is a steady increase in minority populations within the United States (“Race” N.D.). The discovery of unidentified remains is inevitable. The demography of this transition will likely be reflected in forensic casework. Therefore, skeletal collections must diversify and grow in conjunction in order to maintain utility. All of these considerations would help to make population affinity estimation more applicable to American society and forensic casework.

V. CONCLUSION

Anthropology studies how human variation changes over space and time from several perspectives (behavioral, biological, linguistic, etc.). Variation is evident in all individuals and populations both genetically and phenotypically. This emphasizes the need for a holistic understanding of population affinity as it relates to skeletal biology. Within the forensic context, correct estimations increase the likelihood of positive identifications. Therefore, calculating accuracy rates of affinity estimates provide insight into how well the standards account for variation and limit biases such as interobserver error.

This study examines how well craniometric population affinity estimations compare to donor-specified identity. For the sampled populations, the methods produce high classification rates depending on a variety of factors such as population and sex. The results indicate groups with robust representation in anthropology skeletal collections (TXSTDSC, FDB) classify with high posterior probabilities while others do not. As expected, White populations have high classification rates paired with high posterior probabilities. Black/African American donors classified correctly 83.3% of the time, but the posterior probabilities were not high. In contrast, Native American, Hispanic, and Asian groups are less likely to be correctly classified due in part to small reference groups and sample sizes. Finally, there is no apparent pattern between the sexes; both classify relatively well.

Ultimately, anthropological ancestry (population affinity) estimation has a sordid history (e.g., eugenics, white supremacy, and other pseudosciences), but by standardizing

estimation techniques, bias is limited and allows anthropologists to evaluate how we understand and classify biological human variation. This research has the potential to change public understanding of population affinity and demonstrate how it can be a tool for facilitating positive identifications rather than a divisive factor as seen in the current political climate. Additionally, this research demonstrates the usefulness of the biological profile and highlights the fluidity of identity and how it can be biologically embodied. Further research is needed to discern what specific factors affect one's self-identity and how this impacts identification rates. Skeletal collections must adapt as demography changes. Representative and diverse samples are overdue, and it is the responsibility of forensic anthropologists to push for these developments while also respecting the cultural beliefs of the groups.

To conclude, there is a relatively high level of concordance between craniometric affinity estimations and donor identity. Increasing the sample size may highlight new patterns and corroborate the findings.

APPENDIX SECTION

APPENDIX A. Comparison of FORDISC Classifications to the donor-reported Race.

Donor Number	Race (From Donor Form)	FORDISC Classification	Posterior Probability
2015.053	American Indian or Alaskan Native	White Female	0.946
2017.049	American Indian or Alaskan Native	White Male	0.908
2014.025	American Indian or Alaskan Native, Hispanic, White	Japanese Male	0.523
2018.055	American Indian or Alaskan Native, White	White Male	0.743
2019.008	American Indian or Alaskan Native, White	White Male	0.983
2012.034	Asian (Indian)	Japanese Male	0.568
2010.014	Black	Vietnamese Male	0.285
2011.016	Black or African American	Black/African American Male	0.435
2012.010	Black or African American	Black/African American Female	0.502
2012.037	Black or African American	Black/African American Male	0.721
2013.063	Black or African American	White Male	0.614
2014.008	Black or African American	Black/African American Male	0.962
2014.013	Black or African American	Black/African American Male	0.529
2014.052	Black or African American	Black/African American Male	0.824
2016.022	Black or African American	Black/African American Female	0.480
2017.055	Black or African American	Black/African American Male	0.645
2017.065	Black or African American	Black/African American Male	0.746
2019.010	Black or African American	Black/African American Male	0.745
2014.053	Chinese, Hispanic, White	Black/African American Male	0.727
2010.008	Hispanic	Hispanic Male	0.771
2010.015	Hispanic	Japanese Male	0.398

2012.011	Hispanic	Black/African American Male	0.615
2012.014	Hispanic	Black/African American Male	0.509
2012.040	Hispanic	Indigenous American Female	0.554
2013.038	Hispanic	Black/African American Female	0.659
2014.001	Hispanic	Hispanic Female	0.916
2014.028	Hispanic	Hispanic Female	0.817
2014.054	Hispanic	Black/African American Male	0.533
2015.048	Hispanic	Indigenous American Male	0.651
2015.062	Hispanic	Black/African American Male	0.447
2017.022	Hispanic	White Male	0.992
2017.045	Hispanic	White Female	0.952
2017.061	Hispanic	Hispanic Female	0.879
2018.053	Hispanic	Hispanic Female	0.927
2017.067	Hispanic, White	Hispanic Female	0.378
2016.043	Laotian	Chinese Male	0.579
2013.012	Lebanese	White Female	0.850
2009.003	White	White Female	0.892
2010.007	White	White Male	0.969
2010.010	White	Black/African American Male	0.698
2011.001	White	White Male	0.326
2011.012	White	White Female	0.742
2011.014	White	White Male	0.536
2011.015	White	Vietnamese Male	0.866
2012.013	White	White Male	0.780
2012.016	White	White Female	0.628
2012.020	White	White Male	0.938
2012.025	White	Chinese Male	0.339
2012.032	White	White Male	0.504
2012.036	White	White Female	0.711
2012.038	White	White Male	0.765
2012.049	White	White Female	0.743
2013.017	White	White Female	0.848
2013.045	White	White Male	0.980
2014.011	White	White Female	0.871
2014.015	White	White Male	0.514

2014.021	White	White Female	0.579
2014.048	White	White Male	0.930
2014.065	White	White Male	0.904
2015.005	White	White Female	0.816
2015.008	White	Indigenous American Male	0.545
2015.012	White	White Female	0.828
2015.013	White	White Female	0.243
2015.026	White	White Female	0.983
2015.029	White	White Male	0.865
2015.060	White	White Female	0.584
2016.008	White	White Female	0.840
2016.021	White	White Male	0.621
2016.030	White	White Male	0.529
2016.031	White	White Female	0.899
2016.046	White	White Male	0.554
2016.056	White	Hispanic Male	0.583
2018.003	White	White Male	0.706
2018.004	White	White Female	0.677
2018.027	White	White Male	0.947
2018.039	White	White Male	0.856
2019.034	White	White Male	0.905

LITERATURE CITED

- Boas, Franz. "Changes in the Bodily Form of Descendants of Immigrants." *American Anthropologist* 14, no. 3 (1912): 530–62. <https://doi.org/10.1525/aa.1912.14.3.02a00080>.
- Budiman, Abby. "Japanese in the U.S. Fact Sheet." Pew Research Center's Social & Demographic Trends Project. Pew Research Center, May 24, 2022. <https://www.pewresearch.org/social-trends/fact-sheet/asian-americans-japanese-in-the-u-s/>.
- Bulbeck, D. 2011. "Principles Underlying the Determination of Population Affinity with Craniometric Data." *Mankind Quarterly* 52 (1): 35–89.
- Cunha, Eugénia, and Douglas H. Ubelaker. "Evaluation of Ancestry from Human Skeletal Remains: A Concise Review." *Forensic Sciences Research* 5, no. 2 (2019): 89–97. <https://doi.org/10.1080/20961790.2019.1697060>.
- DiGangi, Elizabeth A., and Jonathan D. Bethard. "Uncloaking a Lost Cause: Decolonizing Ancestry Estimation in the United States." *American Journal of Physical Anthropology*, January 18, 2021. <https://doi.org/10.1002/ajpa.24212>.
- Dunn, Rhian R., Micayla C. Spiros, Kelly R. Kamnikar, Amber M. Plemons, and Joseph T. Hefner. 2020. "Ancestry Estimation in Forensic Anthropology: A Review." *Wiley Interdisciplinary Reviews: Forensic Science* 2 (4): 1–26. <https://doi.org/10.1002/wfs2.1369>.
- Dudzik, Beatrix, and Richard L. Jantz. "Misclassifications of Hispanics Using Fordisc 3.1: Comparing Cranial Morphology in Asian and Hispanic Populations." *Journal of Forensic Sciences* 61, no. 5 (2016): 1311–18. <https://doi.org/10.1111/1556-4029.13123>.
- Gocha, Timothy P., Sophia R. Mavroudas, and Daniel J. Wescott. "The Texas State Donated Skeletal Collection at the Forensic Anthropology Center at Texas State." *Forensic Sciences* 2, no. 1 (2021): 7–19. <https://doi.org/10.3390/forensicsci2010002>.
- Hahn, Robert A., Benedict I. Truman, and Nancy D. Barker. "Identifying Ancestry." *Epidemiology* 7, no. 1 (1996): 75–80. <https://doi.org/10.1097/00001648-199601000-00013>.
- Hochberg, Ze'ev. "Developmental Plasticity in Child Growth and Maturation." *Frontiers in Endocrinology* 2 (2011). <https://doi.org/10.3389/fendo.2011.00041>.

- Hughes, Cris E., Chelsey A. Juarez, Taylor L. Hughes, Alison Galloway, Gillian Fowler, and Shirley Chacon. “A Simulation for Exploring the Effects of the ‘Trait List’ Method’s Subjectivity on Consistency and Accuracy of Ancestry Estimations*.” *Journal of Forensic Sciences* 56, no. 5 (2011): 1094–1106. <https://doi.org/10.1111/j.1556-4029.2011.01875.x>.
- Jantz, Richard L., and Stephen D. Ousley SD. *FORDISC 3: Computerized Forensic Discriminant Functions*. Version 3.1. The University of Tennessee, Knoxville, 2005.
- Jenkins, Katharine. “II—Two Routes to Radical Racial Pluralism.” *Aristotelian Society Supplementary Volume* 93, no. 1 (2019): 49–68. <https://doi.org/10.1093/arisup/akz008>.
- Krogstad, Jens Manuel, and Luis Noe-Bustamante. “Key Facts about U.S. Latinos for National Hispanic Heritage Month.” Pew Research Center. Pew Research Center, February 16, 2022. <https://www.pewresearch.org/fact-tank/2021/09/09/key-facts-about-u-s-latinos-for-national-hispanic-heritage-month/>.
- Langley, Natalie R., Lee M. Jantz, Stephen D. Ousley, Richard L. Jantz, and George Milner. Data collection procedures for forensic skeletal material 2.0. The University of Tennessee Knoxville, (2016): https://fac.utk.edu/wp-content/uploads/2016/03/DCP20_webversion.pdf.
- Majumdar, Samirah. “5 Facts about Religion in India.” Pew Research Center. Pew Research Center, May 31, 2022. <https://www.pewresearch.org/fact-tank/2018/06/29/5-facts-about-religion-in-india/>.
- Martínez-Abadías, Neus, Mireia Esparza, Torstein Sjøvold, Rolando González-José, Mauro Santos, and Miquel Hernández. “Heritability of Human Cranial Dimensions: Comparing the Evolvability of Different Cranial Regions.” *Journal of Anatomy* 214, no. 1 (2009): 19–35. <https://doi.org/10.1111/j.1469-7580.2008.01015.x>.
- McCarthy, Justin. “U.S. Approval of Interracial Marriage at New High of 94%.” Gallup.com. Gallup, February 21, 2022. <https://news.gallup.com/poll/354638/approval-interracial-marriage-new-high.aspx>.
- Ousley, Stephen D., and Richard L. Jantz. “Fordisc 3 and Statistical Methods for Estimating Sex and Ancestry. A Companion to Forensic Anthropology.” Essay. In *A Companion to Forensic Anthropology*, edited by Dennis C. Dirkmaat, 311–29. Malden, MA: Wiley-Blackwell, 2012.
- Perez, Anthony Daniel, and Charles Hirschman. “The Changing Racial and Ethnic Composition of the US Population: Emerging American Identities.” *Population and Development Review* 35, no. 1 (2009): 1–51. <https://doi.org/10.1111/j.1728-4457.2009.00260.x>.

- Relethford, John H. “Boas and beyond: Migration and Craniometric Variation.” *American Journal of Human Biology* 16, no. 4 (2004): 379–86. <https://doi.org/10.1002/ajhb.20045>.
- Rico, Brittany, Lydia Anderson, and Rose M. Kreider. Growth in interracial and interethnic married-couple households. U.S. Census Bureau, October 8, 2021. <https://www.census.gov/library/stories/2018/07/interracial-marriages.html>.
- Robinson, Matthew R., Aaron Kleinman, Mariaelisa Graff, Anna A. Vinkhuyzen, David Couper, Michael B. Miller, Wouter J. Peyrot, et al. “Genetic Evidence of Assortative Mating in Humans.” *Nature Human Behaviour* 1, no. 1 (2017). <https://doi.org/10.1038/s41562-016-0016>.
- Spradley, M. Katherine, and Richard Jantz. “What Are We Really Estimating in Forensic Anthropological Practice, Population Affinity or Ancestry?” *Forensic Anthropology*, 2022. <https://doi.org/10.5744/fa.2021.0017>.
- Spradley, M Katherine. “Project IDENTIFICATION: Developing Accurate Identification Criteria for Hispanics,” National Criminal Justice Reference System (2013): 69.
- Spradley, M. Kate. “Metric Methods for the Biological Profile in Forensic Anthropology: Sex, Ancestry, and Stature.” *Academic Forensic Pathology* 6, no. 3 (2016): 391–99. <https://doi.org/10.23907/2016.040>.
- Spradley, M. Katherine. “Use of Craniometric Data to Facilitate Migrant Identifications at the United States/Mexico Border.” *American Journal of Physical Anthropology* 175, no. 2 (2021): 486–96. <https://doi.org/10.1002/ajpa.24241>.
- States News Service*. 2020. “5 Questions: Alice Popejoy on Race, Ethnicity and Ancestry in Science,” August 21. <https://search-ebscohost-com.libproxy.txstate.edu/login.aspx?direct=true&db=edsgbc&AN=edsgcl.633201208&site=eds-live&scope=site>.
- Tajfel, Henri, and John C. Turner. “An Integrative Theory of Intergroup Conflict.” In *Rediscovering Social Identity.*, edited by Tom Postmes and Nyla R. Branscombe, 173–90. *Key Readings in Social Psychology*. New York, NY: Psychology Press, 2010.
- Thomas, Richard M., Connie L. Parks, and Adam H. Richard. “Accuracy Rates of Ancestry Estimation by Forensic Anthropologists Using Identified Forensic Cases.” *Journal of Forensic Sciences* 62, no. 4 (2017): 971–74. <https://doi.org/10.1111/1556-4029.13361>.
- U.S. Census Bureau Quick Facts: United States. U.S. Census Bureau, n.d. <https://www.census.gov/quickfacts/fact/table/US/RHI625219>.