Ethnic identity and genetic ancestry in New Mexicans of Spanish-speaking descent

Meghan Healy
University of New Mexico

Follow this and additional works at: https://digitalrepository.unm.edu/anth_etds

Part of the Anthropology Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the Electronic Theses and Dissertations at UNM Digital Repository. It has been accepted for inclusion in Anthropology ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.
Meghan Elizabeth Healy
Candidate

Anthropology
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Keith Hunley, Chairperson

Heather Edgar

Phillip Gonzales

Yann Klimentidis
ETHNIC IDENTITY AND GENETIC ANCESTRY IN NEW MEXICANS OF SPANISH-SPEAKING DESCENT

by

MEGHAN ELIZABETH HEALY

B.S., Biological Sciences, University of Notre Dame, 2001
M.S., Anthropology, University of New Mexico, 2006

DISSERTATION
Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy Anthropology The University of New Mexico Albuquerque, New Mexico

July, 2019
DEDICATION

To Oscar, world’s best grad school comrade, confidante, and consigliere

Plus je connais les hommes, plus j’aime mon chien
ACKNOWLEDGMENTS

I am immensely grateful to my dissertation committee, particularly my advisor, Keith Hunley, who has been incredibly patient and understanding over my long graduate career. His excitement for learning about human evolution is contagious, and his commitment to and talent for both research and teaching epitomizes what I think academia is supposed to be all about. It has been a privilege to be his student. Heather Edgar, besides being the co-PI on our Heritage New Mexico project with Keith Hunley, provided me with valuable guidance and much-appreciated encouragement throughout the years. I also admire her work-life balance: she’s a rare example of a successful biological anthropologist who manages to maintain an active social life and cool hobbies to boot. Felipe Gonzales helped me turn my final paper from a fledgling idea into a full dissertation chapter, and in doing so, helped me regain the confidence I needed to finish up and graduate. Yann Klimentidis brought valuable experience and insights from his own Ph.D research in New Mexico to my committee, and I greatly appreciate his feedback and corrections on my whole manuscript as well as his friendship.

This research could not have happened without our study participants, who generously gave us their time, shared their stories with us and showed so much enthusiasm for our work. I wish we could all meet up and discuss the research findings together. My upbringing in a quirky town in New Mexico played an important role in shaping my interest in anthropology and in this project in particular: thanks to the rich history and character-filled people of Socorro.

I am so grateful to have worked with Carmen Mosley on our project. I cannot imagine a better teammate for tasks ranging from label-making to blood sampling, from
prying into family histories to troubleshooting databases. Brandon Pettit was a fantastic assistant and gives me hope that Millennials might turn out all right even if they can’t read cursive or address a letter. I am also grateful to the other members of our research team and to collaborators and consultants on the project and my dissertation: Shannon Bermea, Kate Rusk, Marianne Berwick and her lab, Sylvia Rodriguez, Jeff Long, Laura Gómez, Jack Baker, Estevan Rael-Galvez, and Casandra Salgado.

Some of my best friends have also been some of my best colleagues: I am thankful to Jennifer Spence, Bonnie Young, and Jessica Gross for being brilliant, fun, and funny, and for being there with me through the beginning, middle and end of grad school, respectively. I can’t list all of my incredible friends and colleagues here; suffice it to say that I have been fortunate to know them, laugh and cry with them, and learn from them.

Through some unknowable combination of genes and environment, my parents gave me a boundless interest in the world around me and a critical mind for looking at it. I owe everything to them, of course. My whole extended family is wonderful and wonderfully complicated, and I feel lucky to have had their love and support.

Finally, I could never have done this without Drew Levin. He is always the calm in the storm for me, and he is amazingly willing to go along with my ideas even though previous ones have ended in getting lost in the mountains and biking along Greek highways in the dark. He somehow kept believing that I could do this even during times when I thought it was impossible. While it has been challenging to complete my Ph.D with small children along for the ride, their smiles and hugs always more than compensate for sleep and productivity. My thanks and my love go to Norah and Drake for making life more interesting and delightful every single day.
ETHNIC IDENTITY AND GENETIC ANCESTRY IN NEW MEXICANS OF SPANISH-SPEAKING DESCENT

by

Meghan Elizabeth Healy

B.S., Biological Sciences, University of Notre Dame, 2001
M.S., Anthropology, University of New Mexico, 2006
Ph.D., Anthropology, University of New Mexico, 2019

ABSTRACT

This dissertation focuses on a regional population, New Mexicans of Spanish-speaking descent (NMS), to explore the nature of identity-related substructure in admixed populations and its implications for research and policymaking. We looked at the relationship between ethnic/ethnoracial identity and genomic ancestry in NMS in two studies. In the first, we collected genomic ancestry data using 270 autosomal microsatellites in 98 New Mexicans who self-identified as Hispanic or Latino and provided more detailed information on their ethnoracial identities. We tested for genetic substructure in this sample along with 13 other admixed samples from the Americas. The New Mexican sample showed evidence of genetic substructure linked to self-identification in two main groups: those who had recent ancestors from Mexico, who showed higher Native American ancestry, and those whose families had lived in New
Mexico for many generations and emphasized their “Spanish” heritage, who showed higher European ancestry. Analyses of the additional admixed samples demonstrated genetic substructure to be nearly universal in admixed populations in the Americas and suggested that it may often be linked to social identity. We then interviewed 507 NMS and obtained data on ethnic identity, age, birthplace and historical ties to different regions. We estimated genomic ancestry in this sample using 291,917 single nucleotide polymorphisms (SNPs). We again found genomic ancestry differences between groups of NMS who used different ethnic identity nomenclature, and found corresponding differences in the birthplaces of participants and their recent ancestors and the time-depth of family ties to New Mexico between groups. Our data on ethnic identity were collected using both open-response and fixed-choice techniques. We found that the data obtained from both forms of questioning together provided richer information than either dataset alone, providing insights into important features such as strength of commitment to identity nomenclature. Broadly, this research emphasizes the importance of recognizing and accounting for identity-based substructure in admixed populations, which reflect historical patterns of migration, colonization, and cultural change. Failing to do so can confound research on the genetic basis of disease and on causes of variation in health outcomes and social inequality.
# TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................... xi

LIST OF TABLES .............................................................................................................. xiii

CHAPTER 1: INTRODUCTION .......................................................................................... 1

Purpose ........................................................................................................................... 1

Background .................................................................................................................... 2

Research Design ............................................................................................................ 11

Guide to Dissertation ..................................................................................................... 13

References ..................................................................................................................... 14

CHAPTER 2: SOCIAL-GROUP IDENTITY AND POPULATION SUBSTRUCTURE IN ADMIXED POPULATIONS IN NEW MEXICO AND LATIN AMERICA .................................................................................................................. 18

Abstract ......................................................................................................................... 19

Introduction ..................................................................................................................... 19

Materials and Methods ............................................................................................... 20

Results ............................................................................................................................ 23

Discussion ....................................................................................................................... 30

Conclusions .................................................................................................................... 32

Supporting Information ............................................................................................... 32

Acknowledgments ......................................................................................................... 32

Author Contributions .................................................................................................. 32

References ..................................................................................................................... 33
LIST OF FIGURES

Chapter 2

Fig 1. Scatter plot of PC1 and PC2 color-coded by geographic region.................................25
Fig 2. Average FIS.............................................................................................................26
Fig 3. Bar plot of the mean individual-level ancestry for each sample averaged
across 10 runs of structure at K = 10 ............................................................................27
Fig 4. Histograms of individual-level European ancestry averaged across 10 runs
of structure at K = 10 ...................................................................................................28
Fig 5. Histograms of NM-HL and model-based European ancestry. .................................29

Chapter 3

Fig. 1 Proportion of foreign-born NMS by county, Proportion of people who
speak Spanish at home by county ...............................................................................40
Figure 2. Age distributions for NMS groups .....................................................................46
Figure 3. Proportion of individuals with ancestors from Spain and/or Mexico,
Proportion of individuals reporting family ties to colonial New Mexico through
each side of the family ..................................................................................................48
Figure 4. Genomic ancestry and PCA ............................................................................50
Figure 5. European, Native American and sub-Saharan African ancestry ....................52
Figure 6. European, Native American, and sub-Saharan African ancestry by
family ties to geographic locations ..............................................................................53
Figure 7. Birthplaces of participants and their parents and grandparents ......................55

Chapter 4

Figure 1. Comparison of frequencies of identity terms used by NMS as
open-ended responses and those chosen as fixed-choice responses ...........................80

Supplementary Figures

Chapter 2

Appendix D S1 Fig. African ancestry.................................................................143

Appendix D S2 Fig. Native American ancestry..............................................144
LIST OF TABLES

Chapter 2
Table 1. HL sample sizes and locations.................................................................22
Table 2. NMS ethno-racial and ethnic subgroup identity......................................23
Table 3. Ethnicity-race, participants and parents..................................................24

Chapter 3
Table 1. Second-choice ethnic identity terms for each first-choice term..................44
Table 2. Comparative genomic ancestry estimates for admixed populations in Mexico and the US (proportion (standard deviation or range))..........................51

Chapter 4
Table 1. Comparison of terms and term frequencies in responses to open-ended and fixed-choice identity questions .................................................................77
Table 2. Overall & individual term changes between open-ended & list-based identity questions ..............................................................................................................81
Table 3. Contingency table for open-ended vs. list-based identity terms...................85

Supplementary Tables

Chapter 2
Appendix D S1 Table. Ethnic subgroup participant and parents............................145
Appendix D S2 Table. Credible region analyses....................................................146

Chapter 3
Appendix A Table S1. Comparisons of selected demographic data from Bernalillo County and New Mexico ........................................................110
Appendix A Table S2. Genomic ancestry proportions ............................................110
Appendix A Table S3. Genomic ancestry, Wilcoxon rank sum test p-values ..............111

Appendix A Table S4. Birth location .....................................................................112
CHAPTER 1: INTRODUCTION

Purpose

The historical record of New Mexico suggests that the modern population of the state includes subgroups of people of Spanish-speaking descent with different migration periods and different ancestors. Admixture studies have already shown substantial differences in genomic ancestry estimates between populations of Spanish-speaking descent in different regions of the Americas (Bryc et al., 2010; Lisabeth et al., 2011). Even in a single state with a relatively small population like New Mexico, it may not be appropriate to assume genetic homogeneity among those of Spanish-speaking descent. This is important because populations such as New Mexicans of Spanish-speaking descent can be ideal for research in areas such as admixture mapping, but this and other types of genetic research depend on the assumption that the population being studied is unstructured.

In New Mexico, there is a widely-held cultural perception today that the people of Spanish-speaking descent in the state are divided into at least two distinct subgroups. One of these groups consists of the descendants of early Spanish colonists, while another main group comprises more recent immigrants from Mexico (Nieto-Phillips, 2008).

In contrast with cultural perceptions like this of multiple, distinct groups of people of Spanish-speaking descent, researchers and policy-makers often use “Hispanic/Latino” as identifiers at regional, national and even international levels and assume it to have ethnic, genotypic and phenotypic correlates that set this group apart from other racial and/or ethnic groups (Rodríguez, 2013). It is possible that New Mexicans of Spanish-speaking descent (NMS), or even people of Spanish-speaking descent in broader regions,
may be genetically indistinguishable from a single, randomly-mating population. If this is the case, then it should be straightforward to perform population association studies aiming to identify genetic components of multifactorial disease within this group. On the other hand, the pattern of genetic variation in NMS and other populations of Spanish-speaking descent may be structured according to socioculturally-defined subgroups. If so, this structure must be understood and accounted for in research to prevent false associations between genes and disease and confounding factors in research looking at health and social inequality.

The patterning of variation in NMS is also interesting from a purely sociocultural perspective. Humans group themselves and each other at many different levels and for different purposes. Ethnic identification is an important component of self-identity for many people. In the genomic age, people are increasingly using their own DNA to explore aspects of their ancestry and biology. A broader understanding of genetic structure in NMS has the potential to help individuals understand the complexity of the relationship between ethnic identity and genetics, and to help NMS and others better interpret their own genetic results obtained from popular genetic testing companies. As anthropologists, we are interested in exploring the complexity of the relationship between biology and culture. The importance of this research lies in appropriately interpreting and contextualizing our results in a way that informs fellow researchers, the NMS population, and the broader public about the nature of identity-related substructure in a regional population and its implications for research and policymaking.

Background

Origins of New Mexicans of Spanish-speaking Descent
The Americas were originally populated more than 12,500 years ago, and until 1492, the people of the Americas remained isolated from other human populations (Sans, 2000). In the 16th-19th centuries, European colonialism brought long-separated human populations into contact in many different places around the globe. Christopher Columbus’s landing on Hispaniola in 1492 initiated the process of European colonization of the Americas. Large numbers of Africans, especially West Africans, were also brought by Europeans to the Americas as slaves, with at least 120,000 Africans arriving in Mexico alone between 1521 and 1650 (Davidson, 1966). Latin American populations today have ancestors primarily from different combinations of these three “parental” populations: Native Americans, Europeans and West Africans.

Historical documents provide some information about the nature of the interactions between these populations in the Americas. For example, we know that many more Spanish men than women made the journey to the New World in the colonial era, creating a dearth of potential wives for the Spanish colonists (Marshall, 1939; Sans, 2000). To push population expansion, the Spanish government supported marriage between Spanish men and non-Spanish women. However, careful attention was paid to the ancestry of the offspring, especially later in the colonial period when greater Spanish ancestry became more important for accessing social status and privileges (D. Hayes-Bautista, 1980; Seed, 1982).

The land that is now the state of New Mexico has a long and complicated history of human settlement that began at least 12,000 years ago when Native Americans arrived. Evidence of the Clovis and Folsom cultures in New Mexico provides some of the earliest information about big game hunters in the Americas. A decline in large mammals led to
the gradual rise of more sedentary populations, eventually culminating in the rise of the Ancestral Puebloan (formerly often referred to as “Anasazi”, but see Adler et al., 1996) and Mogollon cultures. The Ancestral Puebloans became the dominant culture in the Southwest by about A.D. 1050, building multistoried adobe apartment houses and cliff dwellings known as Pueblos. Drought conditions in the late 13th century led to movement and changes for the Puebloans, with population centers remaining in the Rio Grande Valley and to the west in the Zuni-Acoma area (Roberts & Roberts, 2006). In the 14th century, nomadic peoples in the Athabascan language group, originating in Canada, arrived in New Mexico (Seymour, 2012). These were the ancestors of the Navajo and Apache peoples.

In 1598, Spanish explorers arrived from Mexico and became the first Europeans to settle in New Mexico (Gonzalez, 1969). These first Spanish colonists took lodging in the native pueblos and adopted material goods and farming techniques from the Pueblo Indians (Chavez, 2002); acculturation was pervasive from the very beginning. With the Pueblo Revolt in 1680, many Spanish families left New Mexico permanently, but some original families returned along with new settlers from Mexico in 1693 (Bustamante, 1991; Roberts & Roberts, 2006). While historical records indicate extensive intermixture between the Spanish and the indigenous peoples of the area, many New Mexican families have long considered themselves to be exclusively of Spanish descent (Bonilla et al., 2004; Ruiz & Sánchez Korrol, 2006). Others have claimed both Spanish and Native American ancestors, and still others, termed Genízaros, may have spoken Spanish and had Spanish surnames, but been of full Native American parentage (Devor, 1980; Magnaghi, 1990). Immigration from Mexico has continued until the present day, but it
has experienced additional dips and surges with events and political trends such as the Treaty of Guadalupe Hidalgo, the fight for statehood (Gómez, 2007), and U.S. immigration reform. Americans of northern European descent and other immigrants became an important component of the state’s population beginning in the mid-19th century. When talking about their heritage, NMS currently use a variety of terms that connect them to places, cultures, languages, and ancestors, as discussed below.

**New Mexicans of Spanish-speaking descent today**

In recent government and mass media use in the United States, “Hispanic” is regularly used to describe an ethnic group united by Spanish language and cultural heritage. It theoretically includes all Spanish-speaking populations of Latin America, while “Latino/a” also can be used to include those from Latin America who speak other Romance languages, e.g., Brazilians. Some researchers give definitions for these and other signifiers: individuals born in Mexico are “Mexican” regardless of their current residence, a “Mexican American” is a United States citizen born in the United States of Mexican descent, and a “Chicano/a” is a person of Mexican descent who was born in the United States and “who possesses a political consciousness of himself or herself as a member of a historically and structurally oppressed group” (Rinderle, 2005). In a common university-level anthropology textbook, all Americans of Mexican descent are termed “Chicanos” (Kottak, 2012). While the terms are clearly defined in these and other contexts, the definitions do not always agree with each other, and the ways that people conceptualize and use these terms in real life vary widely (Comas-Díaz, 2001; D. E. Hayes-Bautista & Chapa, 1987). In New Mexico today, people identify their connections to their Spanish-speaking ancestors in different ways, using terms including Hispanic,
Latino/a, Chicano/a, Mexican/Mexican-American, Spanish/Spanish-American/Spaniard and Nuevomexicano/a. A failure to recognize the variation behind signifiers such as these may cause cultural homogenization and lead to contradicting research findings (Rinderle, 2005). Confusion and controversy concerning the applicability of these identifiers to questions about race, ethnicity and origins may even lead to the removal of those designations from the 2020 census, to be replaced simply by asking which “categories” people identify with (Cohn, 2015).

New Mexico currently stands out as the state with the largest proportion of “Hispanic/Latino” citizens, with 46.3% of its people identifying themselves in this group in 2010 (U.S. Census Bureau, 2010). In the United States overall, 16.3% identified as “Hispanic/Latino”. New Mexicans identifying as “Hispanic/Latino” are often grouped together for government and research purposes, but this broader grouping ignores the variation suggested by the identification terms listed above. Many counties in the center of the state have even higher proportions of people who identify as “Hispanic/Latino”. In the other U.S. states that border Mexico, the highest proportions of “Hispanic/Latino” people are found along the border with Mexico, whereas in New Mexico, border areas have lower proportions of “Hispanic/Latino” citizens than the San Luis Valley in the north-central part of the state. The San Luis Valley was relatively isolated from outside influences through most of the 18th and 19th centuries, and the people of this area have been found to be even more likely than other New Mexicans to emphasize their Spanish roots (Chávez, 1984; Quintana, 1974). Many in this region trace their families’ histories back to individuals who received land grants from the Spanish crown in the colonial period.
Admixture, Ancestry and Ancestors

Admixture is the process of interbreeding between formerly isolated groups. Isolation is important to this process, because during the period of isolation, genetic differences accumulate due to genetic processes including mutation, genetic drift, natural selection, and gene flow with other groups. The human species spread out around the globe from Africa around 50,000-60,000 years ago with a process of serial founder effects (Deshpande et al., 2009). In many regions, populations remained fairly isolated from other human groups until recent long-range migrations like European colonialism (Hellenthal et al., 2014; Li et al., 2008). Research on admixture provides insights into the interactions between long-separated human population groups and the biological and socioeconomic consequences of these interactions. For example, technology-driven power differentials often have significant consequences for the dynamics of human groups coming together in space and time, and we can investigate some of these dynamics at the genetic level, from archaic groups and modern humans in the Old World to sex-biased admixture in the Americas (e.g., Bryc et al., 2010; Higham et al., 2014; Kim & Lohmueller, 2015; Vernot & Akey, 2015; Wang et al., 2008).

Genetic ancestry, also referred to as genomic ancestry or biogeographic ancestry, is an estimate for an individual of the percentages of autosomal DNA that originated from ancestors in different places, or in the case of uniparentally-derived markers, where an individual’s maternal or paternal lineage came from. Most current research in this area focuses on estimating the percentage of an individual’s DNA that came from ancestors who lived in different broad regions (Africa, the Americas, Asia, Europe and Oceania) about 500 years ago, prior to the beginning of large-scale colonization by Europeans (see
Royal et al., 2010; Shriver & Kittles, 2004). For NMS, the historical record indicates that these ancestors primarily include the indigenous people of the Americas, the Spanish and other European colonists who came to Mexico and other parts of the New World in the early colonial period, Africans who initially were brought to the New World as slaves, and Europeans who arrived later from other parts of the U.S.

For every generation going back in time, the number of genealogical ancestors for an individual, i.e., the number of people who are parents to the following generation, theoretically doubles. This creates an exponential growth rate for the number of ancestors as the generation number before the present increases. Using a generation time of 25 years, 500 years corresponds to 20 generations. Going back to 20 generations ago, an individual has 1,048,576 genealogical ancestors, at a time when the entire world population was likely less than half of this number (Durand, 1974). Due to the prevalence of intermarriage between somewhat close relatives through human history, the number of actual ancestors for an individual is nowhere near the number of nodes on the genealogical tree (Coop, 2013). Still, the number of ancestors an individual has from 20 generations ago is certainly on the scale of thousands at the very least.

For the most recent generations, individuals have a substantial portion of autosomal DNA from each ancestor, e.g., 50% from each parent and just about 25% from each grandparent. Due to genetic recombination during meiosis, the amount of genetic material inherited from specific ancestors from a given generation varies stochastically. Moving back in time only seven generations, a small proportion of an individual’s 128 genealogical ancestors are unlikely to have provided any DNA to the descendant in question, while others might have provided as much as 2.2% (Coop, 2013). The
proportion of ancestors who are not represented in an individual’s genome increases when proceeding backwards through time, leading to a decrease in the number of genetic ancestors relative to the number of genealogical ancestors. For example, at 11 generations in the past, there are 2048 nodes on an individual’s genealogical tree, but on average, only 29% of those nodes (about 594 ancestors) passed on DNA that ended up in the individual descendant of interest.

Humans throughout history have taken meaning from finding connections with specific ancestors from hundreds of years before themselves. The New Testament, for example, uses a genealogy of Jesus to connect him through his paternal line via Joseph back to earlier Biblical figures including Abraham and Adam. Several recent or current television shows have taken people, particularly celebrities, and looked at their genetic and genealogical histories, often making a reference to specific ancestors. For example, actress Valerie Bertinelli was linked to Edward the First, King of England from 1239 to 1307 (“Who Do You Think You Are? Season 5, Episode 4,” 2014). The program neglected to mention that at 700 years ago, she would have had approximately 268 million genealogical ancestors overall, and the probability of her having any genetic material from Edward the First would be close to zero (the probability of inheriting zero DNA from ancestor 700 years ago is equal to \(\exp\left(-\frac{(22+33*(k-1))/2}{(k-1)}\right)\), or 0.999993).

Many NMS who participated in our research have extensive genealogical records and can trace specific lineages back to Spaniards who were among the first Europeans to arrive in New Mexico with the expedition of Don Juan de Oñate in 1598, or others who arrived during the resettlement of New Mexico in the next century after the Pueblo Revolt. Many subjects expressed to us an expectation that their genealogical information
would correspond to their genetic ancestry results, i.e., that a genealogical link to a Spanish colonial ancestor indicated that they would have primarily European (Spanish) DNA. The information they had was often more complete on the male side of the family tree, so their records were usually for paternal ancestors, often the ones from whom subjects inherited their surnames. A modern individual has less than a 2% chance of inheriting any autosomal DNA from ancestors 16 generations ago, such as these founding conquistadors. Even if represented, these ancestral conquistadors and family settlers (pobladores) would have passed on a tiny amount of DNA (less than 1%) to the descendant. The overwhelming majority of our participants’ genetic material, therefore, came in tiny portions from multitudes of ancestors, most of whom are not likely to be represented on their family trees.

The exception is the non-pseudoautosomal region of the Y chromosome: since it is inherited without recombination through the male lineage, direct male descendants indeed may have inherited the bulk of their Y chromosomes directly from these documented ancestors. Since the Y chromosome contains relatively few genes compared to other chromosomes, however, a shared Y chromosome is not likely to indicate many shared phenotypes aside from that of being male. Mitochondrial DNA is also inherited as a single unit, without recombination, but since it is primarily transmitted maternally, the relevant ancestors are unlikely to be represented in deep family trees.

In summary, our perceptions about our continental genomic ancestry are often based on genealogical knowledge. However, any single genealogical ancestor from more than a few generations back is unlikely to have contributed significant amounts of DNA to a person living today. This can lead to mismatches between known family history and
measured genomic ancestry. It is important that researchers working with genomic ancestry convey this complexity to research participants and populations.

**Research Design**

**Initial Phase**

The University of New Mexico funded an initial exploration of genetic ancestry in New Mexico Hispanics (NMH). In that study, genetic data were collected for 270 autosomal microsatellites in 98 NMH identified through the Cancer Genetics Network (PI Marianne Berwick). Using published genotypes from seven African, eight European, and twenty-nine Native American populations (Cann et al., 2002), individual genetic ancestry proportions were estimated for the NMH sample and for 13 other admixed populations from throughout the Americas.

**Pilot Research**

Pilot research was conducted in 2009 to examine variation in ethnic identity nomenclature among New Mexicans with Spanish-speaking ancestors. Semi-structured one-on-one interviews were conducted with 25 participants. The responses were used to inform the development of the questionnaire for the main phase of the dissertation research.

**Main Phase**

The National Science Foundation awarded funding for a broader project looking at genetic, sociocultural and phenotypic variation in New Mexico Hispanics (co-PIs Keith Hunley and Heather Edgar). The research protocol (HRPO 10-310) was approved by The University of New Mexico Office of the Institutional Review Board, and all participants provided written informed consent to participate in the study. Members of the research
team (Meghan Healy or Carmen Mosley) conducted face-to-face, one-on-one structured interviews with a total of 507 participants who self-identified as New Mexicans of Spanish-speaking descent (NMS).

Participation was open to individuals who identified as NMS and who were 18 or older. To encourage recruitment of participants identifying with a broad variety of ethnic identity terms, the term “New Mexican of Spanish-speaking descent” was used in all recruitment materials. Participants were recruited through advertisements at the University of New Mexico, community centers located throughout Bernalillo County, and flyers mailed to addresses in areas of Bernalillo County with high percentages of people who identified as “Hispanic, Latino, or Spanish” on the 2000 United States Census. As shown in Table S1 in Appendix A, Bernalillo County is representative of New Mexico as a whole in terms of age, income, education, English and Spanish language proficiency, and ethnic identity.

The full research questionnaire is shown in Appendix B. Responses were entered in real time into a Microsoft Access database. Interview duration ranged from one to three hours, with an approximate mean duration of 1.75 hours.

During the interviews, mouthwash samples were collected for obtaining DNA. For DNA collection, each participant swished 10 mL of Original Mint Scope Mouthwash for one minute, then spit the mouthwash into a plastic cup, which was then transferred to a labeled 15 mL tube. A modified Puregene protocol was used for extracting purified DNA from the mouthwash samples. Sample concentration was assessed using a NanoDrop (Thermo Fisher Scientific), and samples with insufficient DNA (n=23, 4.5%) were recollected and extracted during follow-up sampling. Sufficient DNA was obtained
in all 23 follow-up samples, for an overall DNA sample collection rate of 100%. Aliquots of purified DNA were genotyped using an Illumina HumanCytoSNP-12 DNA Analysis BeadChip Kit of 291,917 SNPs at the University of Michigan DNA Sequencing Core. All SNP call rates exceeded 99%. In addition to participant samples, HGDP-CEPH (Cann et al., 2002) samples from Africa (Yoruba, Mandenka, Bantu, and San), Europe (French, Adygei, Orcadian, Russian, Sardinian, and Tuscan) and the Americas (Mexican Pima, Maya, Colombian, Karitiana, and Surui) were genotyped for the same markers at the University of Michigan DNA Sequencing Core. Genetic ancestry was estimated using the model-based ADMIXTURE program (Alexander, Novembre, & Lange, 2009) with the HGDP-CEPH groups serving as parental populations.

**Guide to Dissertation**

This is a hybrid dissertation, in which published papers and manuscripts to be submitted for publication take the place of the standard body of the dissertation. Chapters 2 and 3 are published papers in peer-reviewed journals, while Chapter 4 is in preparation for submission to a peer-reviewed journal. All three papers have multiple authors with Meghan Healy as first author. Appendix C provides citations and detailed author contributions for Chapters 2 and 3. Chapter 2 tests for genetic substructure within a sample of New Mexicans who self-identified as Hispanic or Latino and in 13 other admixed populations in the Americas. It also examines the relationship between genetic substructure and ethnoracial identity in the New Mexico sample population. This chapter was published in PLoS One (Healy et al., 2017). Chapter 3 examines the relationship between genomic ancestry, ethnic identity and regional history in a larger sample of
NMS. It was published in the journal Biodemography and Social Biology (Healy et al., 2018). Chapter 4 compares the informativeness of ethnic identity nomenclature data obtained using open-response and fixed-choice methods in NMS and discusses the histories of important ethnic identity terms in New Mexico. We intend to submit this chapter to the journal Sociology of Race and Ethnicity or a similar journal.

References

Who do you think you are? Season 5, Episode 4. (2014). USA: TLC.


CHAPTER 2: SOCIAL-GROUP IDENTITY AND POPULATION SUBSTRUCTURE IN ADMIXED POPULATIONS IN NEW MEXICO AND LATIN AMERICA

Meghan E. Healy¹, Deirdre Hill², Marianne Berwick², Heather Edgar¹, Jessica Gross¹, Keith Hunley¹

1 Department of Anthropology, University of New Mexico, Albuquerque, NM, United States of America,

2 Department of Internal Medicine, University of New Mexico, Albuquerque, NM, United States of America

Citation:

Social-group identity and population substructure in admixed populations in New Mexico and Latin America

Meghan E. Healy¹, Deirdre Hill², Marianne Berwick¹, Heather Edgar¹, Jessica Gross¹, Keith Hunley¹*,

1 Department of Anthropology, University of New Mexico, Albuquerque, NM, United States of America,
2 Department of Internal Medicine, University of New Mexico, Albuquerque, NM, United States of America

* k hunley@unm.edu

Abstract

We examined the relationship between continental-level genetic ancestry and racial and ethnic identity in an admixed population in New Mexico with the goal of increasing our understanding of how racial and ethnic identity influence genetic substructure in admixed populations. Our sample consists of 98 New Mexicans who self-identified as Hispanic or Latino (NM-HL) and who further categorized themselves by race and ethnic subgroup membership. The genetic data consist of 270 newly-published autosomal microsatellites from the NM-HL sample and previously published data from 57 globally distributed populations, including 13 admixed samples from Central and South America. For these data, we 1) summarized the major axes of genetic variation using principal component analyses, 2) performed tests of Hardy Weinberg equilibrium, 3) compared empirical genetic ancestry distributions to those predicted under a model of admixture that lacked substructure, 4) tested the hypotheses that individuals in each sample had 100%, 0%, and the sample-mean percentage of African, European, and Native American ancestry. We found that most NM-HL identify themselves and their parents as belonging to one of two groups, conforming to a region-specific narrative that distinguishes recent immigrants from Mexico and their families who have resided in New Mexico for generations and who emphasize their Spanish heritage. The “Spanish” group had significantly lower Native American ancestry and higher European ancestry than the “Mexican” group. Positive Fst values, PCA plots, and heterogeneous ancestry distributions suggest that most Central and South America admixed samples also contain substructure, and that this substructure may be related to variation in social identity. Genetic substructure appears to be common in admixed populations in the Americas and may confound attempts to identify disease-causing genes and to understand the social causes of variation in health outcomes and social inequality.

Introduction

Admixed populations form when individuals from previously separated populations mate with one another. Research on admixed human populations can assist in detecting alleles
underlying susceptibility to common diseases [1–3], and it can provide important insights into the causes of social inequality and health disparities in societies that are stratified by race or ethnicity [4,5]. Population substructure is potentially an important confound in this research. Substructure occurs when individuals within a larger population mate assortatively due to the presence of geographic barriers, or to structured racial, ethnic, linguistic, or phenotypic variation. Failure to recognize and account for substructure can lead to spurious conclusions about genetic and social causes of disease [4], and it can lead to suboptimal distributions of resources intended to ameliorate the pernicious consequences of racial and ethnic discrimination.

In this paper, we examine substructure associated with racial and ethnic identity in admixed populations in the Americas, concentrating on the “Hispanic or Latino” (HL) population of New Mexico. We chose to focus on New Mexico for three reasons. First, it has the highest proportion of any US state of individuals that identify as HL on the US Census (47% in 2010) [6]. Second, historical and demographic records provide detailed information about the formation of heterogeneity in social identity within the HL population of the state over the past 400 years [7,8]. Third, previous genetic studies have documented the existence of genetic substructure in HL populations in the region [9,10]. These factors make the HL population of New Mexico ideal for studying the relationship between population substructure and social-group identity in the US, and for drawing inferences about this relationship in other admixed populations in the Americas.

Background: Social-group identity in New Mexico

When the first Spanish colonists arrived in the area that would become New Mexico in 1598, they encountered Native Americans whose ancestors had been present in the Americas for more than 13,000 years [11]. The earliest Spanish census and church records enumerated individuals of mixed ancestry as well as exclusive Spanish or Native American ancestry [12,13], indicating that mating between Spaniards and Native Americans occurred from earliest contact [12–15]. These records show that, during the colonial period (1598–1821), Spaniards used a hierarchical caste system to describe and socially rank admixed individuals according to the fraction of Spanish and Native American heritage. During the US territorial period (1850–1912), the population of New Mexico grew as people migrated from other areas of the US. In the face of political marginalization by the new arrivals, the resident population re-embraced its Spanish, i.e., European, roots. Later, in response to increasing migration from Mexico, many New Mexicans began to embrace and romanticize the Spanish heritage of the long-resident population [8,16]. The rate of migration from Mexico increased in 20th century, and today more than half of New Mexicans who identify themselves as HL on the US census also identify as Mexican or Mexican American [17]. As a result of this region-specific history, the notion of comparatively recent Mexican vs. deep Spanish heritage is a salient feature of social identity in New Mexico today [8].

Here, we use four complementary statistical methods to document the existence of genetic substructure in fourteen HL populations with widespread distribution throughout the Americas, and we use the HL population of New Mexico to examine the role played by racial and ethnic identity in formation of this substructure.

Materials and methods

Ethics statement

The University of New Mexico Institutional Review Board approved the protocols used to conduct this research (UNM IRB 98–369). Written informed consent was obtained from all participants prior to data collection.
Terminology
In 1997, the Office of Management and Budget began to use the phrase "Hispanic or Latino" to refer to US residents of Spanish origin or descent, defined as "of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin regardless of race" [18]. The two terms in the phrase were meant to capture regional variation in the US, and in principle, the term "Latino" includes individuals of Portuguese origin, e.g., from Brazil. Based on these conventions, in the current study, we use the term "NM-HL" to refer to the New Mexican HL sample. Additionally, while noting that the two terms are not widely used outside of the US [19], we use "LA-HL" to refer to people of mixed African, European, and Native American ancestry in Latin America.

Self-identified race and ethnicity
Our primary NM-HL sample consists of 98 unrelated adult individuals recruited from the Cancer Genetics Network (CGN) [20], which consists of cancer patients and their first-degree relatives. During initial recruitment for the CGN, telephone interviewers used two questions to elicit information about the ethnic and racial identity of participants and their parents. The first question was "How would you describe your ethnic or racial background?" Participants were asked to choose all that applied from the list:

- White
- Hispanic, Spanish or Latino
- Black or African American
- Native American, Alaskan, Eskimo
- Asian or Pacific Islander
- Other

Because the question contains both OMB race and ethnic categories, we refer to this self-classification as "ethno-racial." We use the acronym "HSL" to refer to the choice "Hispanic, Spanish or Latino."

The second question was "Would you describe yourself as...?" Participants were asked to choose all that applied from the list:

- Mexican, Chicano, New Mexican
- Puerto Rican
- Cuban
- South American or Central American
- Other specified Spanish origin (including European)
- Spanish, Hispanic, or Latino, not otherwise specified

Since these terms refer to subcategories of HL [18], we refer to this self-classification as "ethnic subgroup." We use the acronym "MCN" to refer to the choice "Mexican, Chicano, New Mexican." We address the potential limitations of these terms for NM-HL in the Discussion.
Genetic data
We mailed DNA mouthwash kits to 159 individuals located throughout the state who minimally responded “Hispanic, Spanish or Latino” to the ethno-racial question. DNA extracts from 109 respondents were genotyped for 270 autosomal microsatellite loci from Screening Set 16 from the Marshfield Clinic’s Mammalian Genotyping Service [21]. Eleven individuals failed to genotype, for a final sample of 98 individuals (32 males and 66 females; 83 cancer patients and 15 unrelated controls). We matched these loci to published data from 1,082 individuals in seven African, eight European, 29 Native American, and 13 admixed LA-HL samples from Central and South America [22-24].

The published LA-HL data (Table 1) were used to assess the extent of genetic substructure in HL populations across the Americas. The data were collected for population genetic analyses or disease association studies [23]; the studies did not provide information about the racial or ethnic identity of the individuals. Data for twelve of the samples were collected in single locations in Mexico (Mexico City), Guatemala (Oriente), Costa Rica (Central Valley, abbreviated CVCR), Colombia (four samples from Peque, Medellin, Cundinamarca, Pasto), Chile (two samples from Paposo and Quetalmahue), Argentina (three samples from Salta, Tucuman, and Catamarca). Data for the Brazil sample (Rio Grande do Sul, abbreviated RGS) were collected in two cities (Bagé and Alegrete).

Statistical methods
We used four complementary approaches to examine substructure in the HL samples. First, to explore gross structure among the populations, we summarized the major axes of genetic variation for the HL samples and NM-HL ethno-racial and ethnic subgroups using PCA. The analyses were conducted using the ade4 package [35] in R [36].

The second approach is based on the fact that structured populations have a reduction of heterozygosity relative to that predicted under HWE [37]. To look for this reduction, we calculated Fis for each sample as the average across loci of the quantity 1 - (Observed Heterozygosity/Expected Heterozygosity). Positive values indicate a reduction of heterozygosity relative to that predicted under HWE. The analyses were conducted using GenePop [28].

The third approach is based on research by Verdu and Rosenberg [29] that demonstrated that, under a model in which two parental source populations contributed to an admixed

Table 1. HL sample sizes and locations.

<table>
<thead>
<tr>
<th>Sample</th>
<th>n</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico City</td>
<td>19</td>
<td>Mexico</td>
</tr>
<tr>
<td>Oriente</td>
<td>20</td>
<td>Guatemala</td>
</tr>
<tr>
<td>CVCR</td>
<td>20</td>
<td>Costa Rica</td>
</tr>
<tr>
<td>Pasto</td>
<td>19</td>
<td>Colombia</td>
</tr>
<tr>
<td>Peque</td>
<td>20</td>
<td>Colombia</td>
</tr>
<tr>
<td>Medellin</td>
<td>20</td>
<td>Colombia</td>
</tr>
<tr>
<td>Cundinamarca</td>
<td>19</td>
<td>Colombia</td>
</tr>
<tr>
<td>RGS</td>
<td>20</td>
<td>Brazil</td>
</tr>
<tr>
<td>Quetalmahue</td>
<td>20</td>
<td>Chile</td>
</tr>
<tr>
<td>Paposo</td>
<td>20</td>
<td>Chile</td>
</tr>
<tr>
<td>Catamarca</td>
<td>14</td>
<td>Argentina</td>
</tr>
<tr>
<td>Salta</td>
<td>19</td>
<td>Argentina</td>
</tr>
<tr>
<td>Tucuman</td>
<td>19</td>
<td>Argentina</td>
</tr>
<tr>
<td>New Mexico</td>
<td>88</td>
<td>USA</td>
</tr>
</tbody>
</table>

https://doi.org/10.1371/journal.pone.0185503.t001
population one time in the past, and mating was subsequently random with respect to genetic ancestry, variation in ancestry decreased every generation until all individuals had identical ancestry proportions. To test this "one-time," random-mating model for the HL samples, we first estimated African, European, and Native American ancestry proportions for all individuals using the model-based clustering approach implemented in STRUCTURE [30]. We ran STRUCTURE at values of K from 2 to 15 using the published African, European, and Native American samples as parental sources. For each value of K, we ran STRUCTURE 10 times using a burn-in phase of 25,000 steps and 15,000 MCMC repetitions. Preempting our results, the average African ancestry in the NM-HL samples was low at 1.3%, and the average across the LA-HL samples was only 2.3%. To simplify the analyses of the one-time model, we only considered admixture between Europeans and Native Americans.

We then generated model-based genetic ancestry distributions in which, in the first generation, Europeans mated with one another in proportion 0.542, and Europeans mated with Native Americans in proportion 0.458; these values produced the mean European and Native American ancestry proportions in the NM-HL sample. Each subsequent generation, individuals in the admixed population mated randomly with respect to genetic ancestry. We ran the model 15 times for admixture onset dates of between 1–15 generations before the present. The 15-generation value corresponds roughly to the date 1598 AD, when Spaniards first arrived in New Mexico. These analyses were based on the "single event admixture" model described by Verdu and Rosenberg, in which the variance in ancestry at generation g is given by $s^2(1-\theta)/2^g$, where $s$ is the proportionate contribution from the European parental source population at the time of initial admixture. The model-based ancestry distributions were generated using an original R script written by one of the authors (IMG).

Based on these results, for our fourth approach, we used the 90% credible regions estimated in STRUCTURE to test the hypotheses that individuals in each HL sample had 100%, 0%, and that sample's mean percentage of African, European, and Native American ancestry. We also tested for mean differences in African, European, and Native American ancestry between the NM-HL ethnic subgroups using signed Wilcoxon rank-sum tests.

### Results

In Tables 2 and 3, we show various combinations of choices for the two identity questions. For the ethno-racial question, all participants chose "Hispanic, Spanish, Latino", and 20 participants also chose "White" (Table 2). None of the participants chose "Black or African American," "Native American, Alutian, Eskimo," or "Asian or Pacific Islander."

For the ethnic subgroup question, we created a single "Spanish" subgroup by combining individuals who responded "Other specified Spanish origin (including European)" or

<table>
<thead>
<tr>
<th>Grouping</th>
<th>Categories</th>
<th>Self</th>
<th>Mother</th>
<th>Father</th>
<th>All three</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethno-racial group</td>
<td>HSL</td>
<td>58</td>
<td>50</td>
<td>50</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>HSL &amp; White</td>
<td>20</td>
<td>11</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>HSL &amp; no race</td>
<td>78</td>
<td>79</td>
<td>84</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Non-HSL &amp; White</td>
<td>0</td>
<td>8</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Non-HSL &amp; no race</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Ethnic subgroup</td>
<td>MCN</td>
<td>58</td>
<td>48</td>
<td>56</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>Spanish</td>
<td>34</td>
<td>34</td>
<td>30</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Both MCN and Spanish</td>
<td>4</td>
<td>6</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Neither MCN nor Spanish</td>
<td>2</td>
<td>10</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

https://doi.org/10.1371/journal.pone.0185603.002

PLOS ONE | https://doi.org/10.1371/journal.pone.0185603 | October 4, 2017 | 5 / 17
“Spanish, Hispanic, or Latino, not otherwise specified.” We note that 27 of the 34 individuals in this subgroup chose the first of these two options. Fifty-eight individuals chose MCN (“Mexican, Chicano, New Mexican”) exclusively. Four individuals self-identified as belonging to both the Spanish and MCN subgroup, and two did not identify with either. None of the other ethnic subgroups were chosen by participants. Most participants placed one or both parents either exclusively in the Spanish subgroup or exclusively in the MCN subgroup.

Table 3 compares ethno-racial identity between the participant and each parent, and between the two parents. Most participants placed both parents into the same ethno-racial group that they chose for themselves. A notable exception is that 14 participants identified one parent as non-HSL and White. Of these participants, six identified themselves as MCN, six identified as Spanish, one identified as both, and one identified as neither. No participants identified both parents as non-HSL and White.

For the ethnic subgroup question, the vast majority of participants (> 93%) placed both parents into the same ethnic subgroup that they chose for themselves (SI Table).

Based on these results, we divided the NM-HL sample into five subgroups for the analyses of the genetic data. The first two subgroups contain participants who self-identified exclusively as either Spanish or MCN. The second two subgroups are subsets of the first two; they consist of individuals that identified themselves and both parents as belonging exclusively to one of the subgroups. We refer to these two groups as "All Spanish" and "All MCN." The fifth subgroup consists of participants who identified one parent as non-HSL and White.

### PCA

Fig. 1A is a scatter plot of values for the first two PCs for our full sample of worldwide populations. Both factors separate individuals from Africa, Europe, and the Americas, and both are strongly correlated with European ancestry (R²PC1 = 0.88; R²PC2 = 0.87) and Native American ancestry (R²PC1 = 0.90; R²PC2 = 0.85) in the Native American and LA-HL samples. HSL individuals are dispersed between the European and Native American clusters. Several HSL individuals are also located relatively closely to the African cluster, most notably a single individual from the Brazilian RGS sample.

#### Table 3. Ethnicity-race, participants and parents.

<table>
<thead>
<tr>
<th></th>
<th>Mother</th>
<th></th>
<th>Father</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Self</td>
<td>HSL &amp; White</td>
<td>HSL &amp; no race</td>
<td>Non-HSL &amp; White</td>
<td>Non-HSL &amp; no race</td>
</tr>
<tr>
<td>HSL &amp; White</td>
<td>5</td>
<td>4</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>HSL &amp; no race</td>
<td>3</td>
<td>73</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Father</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Self</td>
<td>HSL &amp; White</td>
<td>HSL &amp; no race</td>
</tr>
<tr>
<td>HSL &amp; White</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>HSL &amp; no race</td>
<td>1</td>
<td>79</td>
</tr>
</tbody>
</table>

https://doi.org/10.1371/journal.pone.0185593
Fig 1B highlights the NM-HL individuals. Although there is overlap in PC values between individuals in different subgroups, individuals from the MCN subgroup (blue outline) are more widely dispersed, and they cluster closer to Native Americans than other subgroups. Four MCN individuals stand apart from the others in their relative proximity to the African cluster. Individuals with a non-HSI White parent cluster nearer to Europeans (two fall within the European cluster), as do, to a lesser extent, Spanish individuals. The mean PC1 values differ between the three subgroups at the 0.05 level. The patterns are similar for the all-MCN and all-Spanish subgroups (not shown).

Fig 1C and 1D highlight two LA-HL samples with high dispersion (Mexico City and Brazilian RGS) and two samples with low dispersion (Chilean Quetalmahue and Costa Rican CVCR) for both PC factors. In addition to the high degree of dispersion, individuals in Mexico City fall into distinctive clusters; six individuals cluster with Europeans, and, at the other extreme, a single individual clusters with Native Americans. This pattern of dispersion is potentially consistent with the existence of substructure within Mexico City. In contrast, individuals in the Quetalmahue and CVCR samples are relatively tightly clustered around their sample means.
Fig 2. Average $F_{ST}$. Positive values indicate a reduction of heterozygosity relative to that predicted under HWE.

https://doi.org/10.1371/journal.pone.0185503.g002

Hardy Weinberg equilibrium

Fig 2 shows the average $F_{ST}$ values for the HL samples and the NM-HL ethnic subgroups. All samples except the Quetalmahue have positive $F_{ST}$ values, suggesting that substructure may be common in HL populations throughout the Americas. In New Mexico, $F_{ST}$ is significantly greater than zero at the 0.05 level in the NM-HL population as a whole and in the MCN subgroup. The latter result is consistent with the existence of additional substructure within the MCN subgroup. $F_{ST}$ is comparatively low in the Spanish and All-Spanish subgroups.

Genetic ancestry and one-time admixture

In the genetic ancestry analyses, the log probability of data, $P(X|K)$, plateaued at $K = 10$. Two of the 10 clusters were specific to African and European individuals, and the remaining eight clusters captured substructure among Native American populations, most notable among relatively isolated populations in Brazil and Paraguay [23,31]. Fig 3 shows a bar plot of the mean individual-level African, European, and combined Native American ancestry for the HL samples. The blue bars show that African ancestry is broadly low, ranging from 0.4% in the Chilean Papos to 6.6% in the RGS. The mean African ancestry in NM-HL is 1.3%.

All samples have comparatively high European and Native American ancestry. European ancestry ranges from a low of 25.0% in the Argentinian Salta to 73.9% in the Brazilian BGS. Conversely, Native American ancestry is lowest in the RGS (19.5%) and highest in the Salta (73.7%). Compared to the 13 LA-HL samples, the NM-HL sample ranks 2nd highest in European ancestry (mean = 72.9%, range = 41.6–97.0%) and the 4th lowest in Native American
ancestry (mean = 25.8%, range = 2.5–57.0%). These ranges are similar to those reported from other studies of HL populations in the US Southwest [9,10].

The histograms of European ancestry in Fig 4 provide a more detailed picture of the degree of variation in ancestry within and among the samples (see S1 Fig and S2 Fig for the African and Native American ancestry distributions). In a handful of locations, the distributions are relatively narrow, e.g., the CVCR and Quelalmahue. In most locations, however, the distributions are broad and heterogeneous, most notably in the Mexico City (range = 2.6%–98%). This broad distribution in Mexico City is consistent with findings from a recent study of genomic diversity in other urban centers in Mexico [3].

For the NM-HL ethnic subgroups, the ancestry distributions are consistent with the PCA results in showing that the Spanish subgroup has a narrower range of European ancestry estimates (56%–94%) than the MCN subgroup (42%–94%) and a significantly higher mean European ancestry (76% vs. 71%, p < 0.05). The range of ancestry is narrower still for the subgroup with one non-HSL White parent (66%–97%), and the mean European ancestry (87%) is significantly greater than it is for the other two subgroups (p < 0.001). These results show that the genetic differences identified in the PCA analyses are the result of differences in continental ancestry among the NM-HL subgroups.

Fig 5 compares the European ancestry distribution for the NM-HL to distributions generated under one-time models of the admixture process with onset times at 3, 6, 9, 12, and 15 generations before the present. The plots show that the observed level of variation in ancestry
in the NM-HL sample exceeds that produced under one-time admixture events occurring any time prior to three generations before the present. These results support findings from the PCA, HWE, and social-group identity analyses. We reject the hypothesis of one-time admixture in NM-HL. Furthermore, if admixture began between the 15th-19th centuries in the 13 LA-HL samples, we reject a one-time model in them as well, with the possible exceptions of the CVCRC and Quetalmahue.

Credible region analyses

The sample sizes are relatively small for many of the HL groups (average = 25), and the 90% credible regions for individual estimates (from STRUCTURE) are fairly broad, averaging 6% (between low and high) across individuals for African ancestry, 23% for European ancestry, and 18% for Native American ancestry. Because these ranges are so broad, to confirm our conclusions about one-time admixture and to further explore the existence of substructure, we used the 90% credible regions to test the hypotheses that individuals in each HL sample had 0%, 100%, and the sample mean African, European, and Native American ancestry percentage. The results are shown in S2 Table. In the NM-HL sample, the 90% credible region excludes 0%
African ancestry for only two individuals, with 14.6% and 12.2% African ancestry respectively, both of whom identify as MCN. The credible region also excludes 100% African ancestry for all 98 individuals. For European ancestry, the 90% credible region excludes 0% for all 98 individuals, and for all but six individuals, it excludes 100%. Of the six individuals, one is in the Spanish (and All Spanish) subgroup, and the remainder have a non-HSL White parent.

The 90% credible region for 17 NM-HL individuals fell above the sample mean (indicating relatively high European ancestry); of these individuals, three identified themselves and their parents as Spanish, four identified themselves and their parents as MCN, and nine had a non-HSL White parent. The credible region for 18 individuals fell below the sample mean (indicating relatively low European ancestry). Of these individuals, three identified themselves and their parents as Spanish, 11 identified themselves and their parents as MCN, and none had a non-HSL White parent. The results are similar (but complementary) for Native American ancestry. These results provide further support for the existence of substructure related to Spanish vs. MCN ethnic subgroup identity in NM-HL. They also confirm that admixture between non-HSL White and NM-HL individuals is common in the state, which is clearly inconsistent with one-time admixture.
Turning to the LA-HL samples, the 90% credible region excluded 100% African ancestry for all individuals in all HL samples, and it included 0% for all individuals in five of the samples. These results confirm that African ancestry is relatively low in these HL populations. European ancestry was relatively narrowly distributed in the CVCRR and Quelalmahu. In these locations, the 90% credible region for ≥ 90% of individuals contains the mean European ancestry of the samples, and, except for one individual in the CVCRR, the 90% credible regions exclude both 0% and 100% European ancestry for all individuals. The numbers are similar for Native American ancestry in these samples. The results are potentially consistent with one-time admixture in these two samples.

The credible region analyses confirm that European and Native American ancestry distributions are much more heterogeneous in the remaining HL samples. In Mexico City (n = 19), European ancestry spans 2.6–98%, and the 90% credible region excludes the sample mean for the majority of individuals [52 Table]. The credible region contains 0% European ancestry for 2 individuals and 100% European ancestry for four individuals. In the Argentinian Salta, the 90% credible region contains 0% European ancestry for six of the 19 individuals. In the Brazilian RGS, already highlighted for their high African ancestry, the 90% credible region contains 100% European ancestry for five of the 20 individuals. These results imply that substructure related to genetic ancestry is ubiquitous in admixed populations throughout the Americas.

Discussion

Our results indicate that substructure related to genetic ancestry is ubiquitous in admixed populations throughout the Americas. Such substructure is potentially an important confound in studies that exploit admixed populations to identify the genetic component of disease, e.g., through genome-wide association and admixture linkage disequilibrium [32]. Substructure is also important in epidemiology and social science studies that seek to identify and eliminate racial and ethnic disparities in health and social welfare. It is important to recognize that group-level identity and its social and health-related correlates may vary widely among peoples that are commonly placed into a single category by the OMB, e.g., HL and African Americans.

HL samples from Latin America

Though the studies of the LA-HL samples did not provide information about the racial or ethnic identity of the participants, we can use historical and demographic information to make inferences about the causes of substructure and the potential correspondence between social group identity and genetic ancestry in these samples. We concentrate on two samples, one showing high dispersion in genetic ancestry (Brazilian RGS) and one showing low dispersion (Costa Rican CVCRR).

In Brazil, the Instituto Brasileiro de Geografia e Estatística (IBGE) recognizes five categories of self-described social group identity that are primarily based on phenotypic traits, including skin color: White, Brown, Black, Yellow, and Indigenous [33]. In 2008, the first three categories comprised 99.1% of the Brazilian population. In the Rio Grande do Sul, which comprises 5.7% of the population of Brazil, that number is 99.5% (White 80.8%, Brown 13.8%, Black 4.9%). Pena et al. [34] demonstrated that the three groups differed with respect to African, European, and Native American ancestry proportions estimated from 40 DNA polymorphisms. Mean European ancestry in individuals that identified as White, Brown, and Black, respectively, was 86%, 44% and 43%. Many individuals that identify as White had almost 100% European ancestry. Respective mean African ancestry values for the three groups were 5%, 44% and 46%. For the Brown and Black groups, African ancestry spanned 0–100%. This pattern of correspondence between social identity and genetic ancestry is consistent with studies
showing that mating is assortative with respect to "color"-based social identity categories in Brazil [35].

In our analyses, European ancestry in the RGS ranges broadly from 14–97%, and the 90% credible region encompasses 100% for five individuals. Native American ancestry ranges from 3–40%, and the 90% credible region spans 0% for eight individuals. African ancestry was particularly heterogeneous; the 90% credible region spans 0% for 17 of the 20 individuals. In these individuals, the mean African ancestry is 1.8%. For the other three individuals, mean African ancestry is 33.6%, and it ranges from 17.2–58.7%. These results are consistent with those from Pena et al. [34], and they strongly suggest that the RGS sample is structured with respect to genetic ancestry.

In the sample from the Central Valley of Coast Rica (CVC), the low dispersion in European and Native American ancestry, combined with the fact that the 90% credible region contains the sample-mean ancestry for all but two individuals (see Fig 2B, S2 Table), is consistent with one-time admixture occurring multiple generations in the past. This scenario is consistent with historical information documenting admixture between males from Spain and a small group of Native Americans females beginning in the 16th century, followed by relative isolation [36]. This history is also consistent with findings from recent genetic studies of low levels of mtDNA diversity (comprised of 83% Native American haplogroups) compared to other Latin American populations [37].

HL sample from New Mexico

We have more direct evidence for ethno-racial substructure in NM-HL. Most NM-HL individuals identify themselves and their parents as belonging to one of two groups; these groups broadly conform to a region-specific narrative that distinguishes relatively recent immigrants from Mexico from individuals whose families have resided in New Mexico for generations and who often emphasize their Spanish heritage. These groups differ from each other with respect to continental ancestry. Epidemiological and social science studies in the region should take this substructure into account, and they should work with individuals and scholars in local communities to identify the social groups that are relevant to New Mexicans. Such research is especially important given the potential lack of congruence between region-specific vs. national-level, e.g., OMB, racial and ethnic categories [38].

Fourteen percent of NM-HL individuals reported having a non-Hispanic White parent. Equal numbers of these individuals identified as Spanish and MCN (6 each). If such admixture has occurred for several generations, it may have eroded even larger ancestry differences that may have once existed between the subgroups. In fact, European ancestry in MCN is higher than it is in other Mexican and Mexican American populations, including those located near the US border [39–42].

These findings and interpretations must be tempered by the fact that the questions we used to elicit information about identity were not perfectly-suited to NM-HL. The first two terms in the Mexican-Chicano/New Mexican category conflate nationality and socio-political identity. Furthermore, the term "New Mexican" may have been associated by some participants with the region-specific term "Nevomexicano," which is typically associated with Spanish heritage [8]. Additionally, of individuals that identified as Spanish, 79% chose "Other specified Spanish origin (including European)" compared to 21% for "Spanish, Hispanic, or Latino, not otherwise specified." The infrequent choice of the latter category may reflect that fact that it combines the potentially region-specific term "Spanish" with broader, more inclusive terms "Hispanic" and "Latino." These identity-related questions highlight the broader problem of capturing how people conceive group-level identity in different regions of the world, especially
when attempting to use generalized terminology that may be inappropriate or meaningless in a particular region.

In this vein, in 2009, the Institute of Medicine (now the National Academy of Medicine) recommended that health disparities researchers include lists of locally relevant fine-grained ethnic subcategories when collecting data [43]. To assist researchers, they compiled a list that separates the terms "Mexican" and "Chicano" into separate categories and includes region-specific terms such as "Nuevo Mexicano." Regional studies, including ours, would benefit from the inclusion of such lists.

We also note that for the NM-HL sample, our European ancestry estimates are higher, and our Native American ancestry estimates lower, than those from other studies of the region [5, 10]. The variation in estimates might reflect differences in sampling methods, e.g., exclusion of individuals with non-HL White parents. It might also reflect differences in the types of markers used in the studies (autosomal microsatellites vs. SNPs). Autosomal microsatellite loci are prone to allelic dropout, which could lead to errors in allele frequency estimation and tests of HWE (e.g., inflating $P_{61}$). We eliminated loci and individuals with high amounts of missing data, but it is possible that dropout affected some of our $F_{ST}$ results.

**Conclusions**

We found that racial and ethnic identity among NM-HL conforms to a region-specific narrative that distinguishes recent immigrants from Mexico from individuals whose families have resided in New Mexico for generations and who emphasize their Spanish heritage. Our analyses of NM-HL and 13 Central and South America HL samples suggest that genetic substructure is ubiquitous in admixed populations in the Americas, and that this substructure may be related to variation in social identity. This substructure may confound attempts to identify disease-causing genes and to understand the social causes of variation in health outcomes and social inequality.

**Supporting information**

S1 Fig. African ancestry. (TIF)

S2 Fig. Native American ancestry. (TIF)

S1 Table. Ethnic subgroup participant and parents. (XLSX)

S2 Table. Credible region analyses. (XLSX)

**Acknowledgments**

This research was funded by grants from the University of New Mexico and the National Science Foundation (BCS #0962825).

**Author Contributions**

Conceptualization: Meghan E. Healy, Deirdre Hill, Marianne Berwick, Heather Edgar, Keith Hunley.

Data curation: Meghan E. Healy, Deirdre Hill, Marianne Berwick, Keith Hunley.
Formal analysis: Meghan E. Healy, Keith Hunley.

Funding acquisition: Deirdre Hill, Marianne Berwick, Heather Edgar, Keith Hunley.


Methodology: Meghan E. Healy, Jessica Gross, Keith Hunley.

Project administration: Heather Edgar, Keith Hunley.

Resources: Deirdre Hill, Marianne Berwick, Keith Hunley.

Software: Meghan E. Healy, Jessica Gross, Keith Hunley.

Supervision: Keith Hunley.

Validation: Meghan E. Healy, Keith Hunley.

Visualization: Meghan E. Healy, Keith Hunley.

Writing – original draft: Meghan E. Healy, Keith Hunley.

Writing – review & editing: Meghan E. Healy, Deirdre Hill, Marianne Berwick, Heather Edgar, Jessica Gross, Keith Hunley.

References


CHAPTER 3: ASSOCIATIONS BETWEEN ETHNIC IDENTITY, REGIONAL HISTORY, AND GENOMIC ANCESTRY IN NEW MEXICANS OF SPANISH-SPEAKING DESCENT

Meghan Healy, Heather Edgar, Carmen Mosley and Keith Hunley

Department of Anthropology, University of New Mexico, Albuquerque, New Mexico

This is an Author’s Original Manuscript of an article published by Taylor & Francis Group in Biodemography and Social Biology on 20 Dec 2018, available online at DOI: 10.1080/19485565.2018.1545563

Abstract

This study examines associations between ethnic identity, regional history, and genomic ancestry in New Mexicans of Spanish-speaking descent (NMS). In structured interviews, we asked 507 NMS to select from a list of eight ethnic identity terms identified in previous research. We estimated genomic ancestry for each individual from 291,917 single nucleotide polymorphisms (SNPs) and compared genomic ancestry, age, and birthplace between groups of individuals who identified using each ethnic identity term. Eighty-eight per cent of NMS who identified as “Hispanic,” “Nuevomexicano/a,” and “Spanish,” on average, were born in New Mexico, as were the vast majority of their parents and grandparents. Thirty-three per cent of NMS who identified as “Mexican” and “Mexican American” were born in Mexico, as were 59 per cent of their parents and 67 per cent of their grandparents. Average Native American and African ancestry proportions in “Hispanic” (0.26, 0.02, respectively), “Spanish” (0.25,
0.01), and “Nuevomexicano/a” (0.24, 0.01) NMS were significantly lower than in “Mexican American” (0.37, 0.04) NMS. Significant age differences between older “Spanish” and younger “Nuevomexicano/a” individuals, combined with widespread use of the term “Hispanic,” may reflect ongoing nomenclature changes. Patterns of correspondence between ethnic identity, ethnic nomenclatures, and genomic ancestry reflect historical patterns of migration, colonization, and cultural change.

**Introduction**

Race and ethnicity are frequently employed as variables in biomedical and social science research that seeks to identify and eliminate the causes of social inequality and health disparity in the US. This research regularly employs a standardized racial and ethnic nomenclature developed by the federal government. While this nomenclature takes into account regional histories and is constantly shifting in response to ongoing demographic and social change (Office of Management and Budget 1997; US Census Bureau 2017; Siegel and Passel 1979), relatively little attention has been paid by social scientists and health researchers to how well it truly captures ethnic identity in different regions of the US (Bradby 2003; Doan and Stephan 2006; Hunley et al. 2018; Taylor et al. 2012). New Mexico, with the highest proportion of people of Spanish-speaking descent in the US at 47 per cent (Pew Research Center 2014; US Census Bureau 2016), provides a unique opportunity to explore ethnic identity and its social and historical correlates.

Ethnic identity is a complex concept. Definitions of ethnicity have varied since Weber began exploring the concept (Weber 1978), but fundamental to all definitions in sociology and anthropology today is the notion of self-ascription (Cornell and Hartmann
The US government recognizes a single ethnic group defined by the Office of Management and Budget as persons of “Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race” (Office of Management and Budget 1997). More fine-grained ethnic categories from the US Census Bureau and National Academy of Medicine (Institute of Medicine 2009) recognize that there is regional variation in self-ascribed ethnic identity in the US that has been molded by complex histories of colonization and migration.

When the first Spanish colonists arrived in the area that would become New Mexico in 1598, they encountered Native Americans whose ancestors had been present in North America for more than 12,000 years (Huckell 2014). Based on census and church records from the time, mating between Spaniards and Native Americans was common from the start (Brooks 2002; Tjarks 1978). During the US territorial period (1850–1912), the population of New Mexico grew as people of largely European descent migrated from other areas of the US. In the face of potential political marginalization by the new arrivals, the resident population began to emphasize its Spanish roots. Later, in response to increasing migration from Mexico, many New Mexicans began to romanticize the Spanish heritage of the long-resident population (Gonzales 1993; Wilson 1981).

As a result of the growing emphasis on Spanish heritage, the notion of comparatively recent Mexican vs. early colonial-period Spanish heritage became embedded in the institutions of the state (Hunley et al. 2018) and in the terms that New Mexicans of Spanish-speaking descent (NMS) used to describe themselves, including, “Nuevomexicano,” “Spanish,” and “Hispano” (Lomeli, Sorell, and Padilla 2002; Nieto-Phillips 2008; Nostrand 1996). This Mexican-Spanish distinction is supported by US
Census data showing that in southern border states other than New Mexico, individuals identifying on the 2010 US Census as “Hispanic, Latino, or Spanish” (HLS) are most heavily concentrated along the border, suggesting that they or their families recently immigrated to the US from Mexico. Figure 1 shows that counties in California, Arizona, and Texas nearest to the Mexican border have relatively high proportions of residents who speak Spanish at home. In contrast, in New Mexico, counties in the northcentral portion of the state have relatively high concentrations of people who speak Spanish at home. These counties are located in a region called the San Luis Valley, where early Spanish settlers congregated in part as a result of land grants provided by the Spanish crown to encourage immigration to the region. A relatively high proportion of individuals in this area today report that their ancestors received land grants from the Spanish crown (Chávez 1984; Quintana 1974), and many have genealogical records tracing their family history to the earliest Spanish settlers. Other census results also capture this emphasis on Spanish heritage, e.g., in response to the question “Is Person 1 of Hispanic, Latino, or Spanish origin?,” many New Mexicans select the subcategory “another Hispanic, Latino, or Spanish origin” and write in terms related to Spanish origins on the line provided. In 2010, for example, 6.5 per cent wrote in “Spanish,” “Spaniard,” or “Spanish American,” compared to 0.5 per cent, 0.7 per cent, and 0.5 per cent, respectively for Arizona, California, and Texas (US Census Bureau 2016a).
Figure 1.

Top: Proportion of foreign-born NMS by county. Bottom: Proportion of people who speak Spanish at home by county. In both maps, Bernalillo County is outlined in red, and counties in the San Luis Valley in New Mexico and Colorado are outlined in yellow.
This history of the region, in combination with the geographic distribution of Spanish-language usage and the high frequency of write-in responses on the census, suggests that ethnic identity in New Mexico may be unique compared to other US regions. However, ethnic identity in other regions has also been molded by region-specific historical and social factors (De León 2003; Duany 1998, 2003; Maciel and Gonzales-Berry 2000; Miller 1976), implying that standardized ethnic nomenclatures in wide use at the national level may mask important information about how people view ethnic identity at more local levels. With this concern in mind, in this study, we asked 507 NMS to identify the ethnic identity terms that they used to describe themselves, and we examined the relationships between these terms and age, genomic ancestry, birthplace, and family ties to the US Southwest, Spain, and Mexico.

**Methods**

The University of New Mexico Office of the Institutional Review Board approved the research protocol (HRPO 10-310), and all participants provided written informed consent to participate in the study.

In 2008, in semistructured interviews, we asked 25 adult NMS to list all ethnic identity terms that NMS use to describe themselves. The seven most commonly chosen terms, in alphabetical order, were “Chicano/a,” “Hispanic,” “Latino/a,” “Mexican,” “Mexican American,” “Nuevomexicano/a,” and “Spanish.” We then conducted structured face-to-face interviews with 507 self-identified, adult NMS (Hunley et al. 2018). Individuals were recruited from advertisements at the University of New Mexico, community centers located throughout Bernalillo County, and flyers mailed to addresses in areas of Bernalillo County that contained high proportions of individuals who
identified as HLS on the 2000 census. Bernalillo County is representative of New Mexico with respect to age, income, education, English and Spanish language proficiency, and ethnic identity (table S1). Advertisements used the term “New Mexican of Spanish-speaking descent” in order to be neutral with respect to ethnic identity nomenclature. During the interviews, we asked, “With which of these groups do you identify most?” and offered the following choices: “Chicano/a,” “Hispanic,” “Latino/a,” “Mexican,” “Mexican American,” “Nuevomexicano/a,” “Spanish,” and “Other.” Participants were permitted to select multiple terms. When they did so, we asked them to rank their choices in terms of personal significance. We also asked participants to provide their sex, age, and city, county, and country of birth for themselves, both parents, and all four grandparents.

We asked individuals two questions that assessed perceptions about family ties to geographic locations. The first question was “Does your father/mother belong to an old New Mexico family?” The second question was “Were any of your ancestors colonists from Spain/ Do you have any ancestors from Mexico?”

To compare patterns of genetic diversity among various groupings of NMS, we extracted DNA from mouthwash samples. We assayed 291,917 single nucleotide polymorphisms (SNPs) from aliquots of purified DNA. We also genotyped 139 individuals from the HGDP-CEPH (Cann et al. 2002) from sub-Saharan Africa (25 Yoruba, 5 Mandenka, 5 Bantu, and 5 San), Europe (29 French, 5 Adygei, 5 Orcadian, 5 Russian, 5 Sardinian, and 5 Tuscan), and the Americas (25 Mexican Pima, 5 Maya, 5 Colombian, 5 Karitiana, and 5 Surui).
We summarized the major axes of genetic variation using principal component analyses (PCA) (Jombart 2008), and we estimated sub-Saharan African, European, and Native American ancestry using the model-based method implemented in ADMIXTURE (Alexander, Novembre, and Lange 2009). In these analyses, the sub-Saharan African, European, and Native American individuals served as proxies for the parental sources. We assumed that all partitions of individuals were equally likely a priori and set the number of clusters, K, to 3.

We used Wilcoxon rank-sum tests to assess differences in genomic ancestry between groups based on ethnic identity and family history. For the tests, two-tailed p-values were adjusted for multiple tests using the method of Holm (1979). Analyses were conducted in R (R Core Team 2014). Results for family history and genomic ancestry were displayed using violin plots, which combine conventional box plots with mirrored normal density plots (Hintze and Nelson 1998).

We tested for age differences between groups of NMS who identified using the seven ethnic identity terms and “Other” using exact Wilcoxon rank-sum tests, employing the Shift-algorithm for tied samples (Streitberg and Rohmel 1986). Age for each group was displayed using bean plots (Kampstra 2008).

Census data and birth location were plotted in R using the choroplethr package (Lamstein and Johnson 2017). Census data were accessed from the American Community Survey using the “acs.fetch” function in R and plotted with the “county_choropleth” function. We converted birthplace to latitude and longitude using the center points for city of birth, then assigned those points to counties in New Mexico and states in Mexico for mapping using the GADM spatial database (Hijman 2017).
Results

Overall, our results provide support for a historically rooted distinction between NMS who express connections to New Mexico during the colonial-period vs. NMS who express more recent connections to Mexico. The distinction is captured by the terms that NMS use to identify themselves, patterns of allele frequency variation, family ties to geographic locations, and birthplace.

Ethnic Identity Terms

Table 1 shows the primary and secondary ethnic identity terms chosen by NMS. Only 20 NMS chose a single term from the list of eight terms; 75 (15 per cent) chose two terms, 403 (79 per cent) chose three terms, and nine (2 per cent) chose more than three terms. About 44 per cent chose “Hispanic” as their primary term. As Table 1 shows, “Hispanic” individuals frequently chose “Spanish” (32 per cent) and “Nuevomexicano/a” (23 per cent) as second terms, and they rarely chose “Mexican” (4 per cent) or “Mexican American” (6 per cent). Likewise, individuals who chose “Mexican” or “Mexican American” as their primary term never chose “Spanish” as an additional term and only twice chose “Nuevomexicano/a.” About 5 per cent of NMS chose “Other” as their primary term. For the remainder of the analyses, we report results for the first-choice term only.

Table 1.

Second-choice ethnic identity terms for each first-choice term.
Demographic Characteristics

NMS did not differ appreciably from the general populations of Bernalillo County and New Mexico with respect to annual family income (table S1) (US Census Bureau 2016b, 2018). NMS were, however, more highly educated, with 97 per cent having graduated from high school compared to 88 per cent for Bernalillo County and 85 per cent for New Mexico as a whole. Additionally, a relatively low proportion of NMS were foreign born (3 per cent vs. 11 per cent for Bernalillo County and 10 per cent for New Mexico).

Median age for the NMS sample was high (49.0 years) compared to the median for New Mexicans who chose HLS on the 2010 census (36.7 years) (US Census Bureau 2018).

The age distributions for each ethnic identity group are shown in Figure 2. Within each group, age is statistically indistinguishable for males and females (results not shown). After correcting for multiple tests, the only significant difference was between older “Spanish” (mean = 54.4 years) and younger “Nuevomexicano/a” (mean = 43.3 years) (p < 0.0017).

---

Table 1. Second-choice social identity terms for each first-choice term

<table>
<thead>
<tr>
<th>First-choice social identity term (number)</th>
<th>Chic</th>
<th>Hisp</th>
<th>Lat</th>
<th>Mex</th>
<th>Mex Amer</th>
<th>Nuevo</th>
<th>Span</th>
<th>Other</th>
<th>Sum*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicano/a (53)</td>
<td>-</td>
<td>21</td>
<td>0.42</td>
<td>11</td>
<td>0.22</td>
<td>0.04</td>
<td>5</td>
<td>0.1</td>
<td>0.18</td>
</tr>
<tr>
<td>Hispanic (225)</td>
<td>29</td>
<td>0.13</td>
<td>40</td>
<td>0.19</td>
<td>9</td>
<td>0.04</td>
<td>12</td>
<td>0.06</td>
<td>49</td>
</tr>
<tr>
<td>Latino/a (17)</td>
<td>2</td>
<td>0.11</td>
<td>8</td>
<td>0.5</td>
<td>0</td>
<td>0.01</td>
<td>1</td>
<td>0.06</td>
<td>4</td>
</tr>
<tr>
<td>Mexican (16)</td>
<td>3</td>
<td>0.21</td>
<td>3</td>
<td>0.21</td>
<td>4</td>
<td>0.29</td>
<td>0</td>
<td>0.01</td>
<td>0</td>
</tr>
<tr>
<td>Mex American (38)</td>
<td>9</td>
<td>0.24</td>
<td>16</td>
<td>0.42</td>
<td>3</td>
<td>0.08</td>
<td>7</td>
<td>0.18</td>
<td>2</td>
</tr>
<tr>
<td>Nuevomexicano/a (75)</td>
<td>16</td>
<td>0.21</td>
<td>38</td>
<td>0.51</td>
<td>5</td>
<td>0.07</td>
<td>1</td>
<td>0.01</td>
<td>3</td>
</tr>
<tr>
<td>Spanish (62)</td>
<td>2</td>
<td>0.03</td>
<td>44</td>
<td>0.72</td>
<td>7</td>
<td>0.11</td>
<td>0</td>
<td>0.01</td>
<td>1</td>
</tr>
<tr>
<td>Other (27)</td>
<td>4</td>
<td>0.19</td>
<td>6</td>
<td>0.29</td>
<td>0</td>
<td>0.01</td>
<td>1</td>
<td>0.05</td>
<td>2</td>
</tr>
</tbody>
</table>

* differs from column 1 total because several individuals did not choose a second term or rated multiple terms as their first choice
Figure 2.

Age distributions for NMS groups. In each bean plot, the circle marks the mean age, the thick vertical lines mark standard deviation, and the thin vertical lines mark individual observations, with width proportional to the number of observations at that age. The contours are mirrored normal density traces.

Family Ties to Geographic Locations

Figure 3a shows the proportion of individuals from each ethnic identity group who reported having ancestors from Spain and/or Mexico. Figure 3b shows the proportion of individuals from each group who reported family ties to New Mexico during the colonial period. Both figures distinguish individuals who identified primarily as “Mexican” and “Mexican American” from individuals who identified primarily using
the other terms. In most cases, “Mexican” and “Mexican American” individuals differed most from individuals who identified as “Spanish.” As an example, of 62 individuals who identified as “Spanish,” 56 per cent (n = 35) stated that they had ancestors from Spain but not Mexico (Figure 3a). In contrast, only one individual who identified as “Mexican,” and one who identified as “Mexican American”, stated that they had ancestors from Spain but not Mexico. Additionally, 76 per cent (n = 47) of “Spanish” individuals stated that both parents belonged to families that had ties to colonial New Mexico compared to 6 per cent (n = 1) for “Mexican” and 8 per cent (n = 3) for “Mexican American” (Figure 3b).
Figure 3.

(A) Proportion of individuals with ancestors from Spain and/or Mexico. (B) Proportion of individuals reporting family ties to colonial New Mexico through each side of the family.
Genomic Ancestry

The mean European (0.71), Native American (0.27), and sub-Saharan African (0.02) ancestry proportions for the NMS sample are listed in Table 2 along with estimates from 10 other studies of admixed populations in Mexico and the US (Bonilla et al. 2004; Cerda-Flores et al. 2002; Healy et al. 2017; Klimentidis, Miller, and Shriver 2009; Martinez-Fierro et al. 2009; Moreno-Estrada et al. 2014; Risch et al. 2009; Wang et al. 2008; Young et al. 2014). All estimates were derived from autosomal genetic data, though the parental source samples, genetic markers, and estimation methods differed for each study. The studies from New Mexico and Colorado confirm our results in showing that Native American ancestry is lower in the US Southwest than in other regions of the US and Mexico, while European ancestry is higher, and African ancestry is consistently low. Figure 4a is a bar plot of individual genomic ancestry estimates for the three parental source samples and NMS. The sub-Saharan African, European, and Native American samples generally fall neatly into three clusters, though among Native Americans, several individuals in the Mayan sample (in the “Americas” cluster in Figure 4a) have high membership in the cluster associated with Europeans as a result of post-Colombian admixture (Hunley, Gwin, and Liberman 2016; Wang et al. 2007).
Figure 4.

Genomic ancestry and PCA. (A) Genomic ancestry estimates (proportion) from unsupervised analysis assuming three clusters. The plot is comprised of 646 vertical lines, one for each individual in the sample. In each grouping ancestry estimates are sorted from high to low European ancestry. (B) PC factors 1 and 2. (C) Mean and 95 per cent confidence interval ellipses for individuals who identified as “Chicano/a,” “Hispanic,” “Latino/a,” and “Other.” (D) Mean and 95 per cent confidence interval ellipses for individuals who identified as “Mexican,” “Mexican American,” “Nuevomexicano/a,” and “Spanish.”

Figure 4b contains scatterplots of the first two PC factors, which, combined, account for 6.5 per cent of the variation in allele frequencies. The PCs are significantly correlated with Native American ancestry (RPC1 = −0.75; RPC2 = 0.98) and European
ancestry (RPC1 = 0.54; RPC2 = −0.99). Both factors distinguish the three parental source
samples from one another and from NMS. NMS broadly fall between Europeans and
Native Americans for both factors, though there is substantial overlap with Europeans.

Table 2.

Comparative genomic ancestry estimates for admixed populations in Mexico and the US
(proportion (standard deviation or range)).

<table>
<thead>
<tr>
<th>Sample location</th>
<th>Native American</th>
<th>European</th>
<th>African</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMS</td>
<td>0.27 (0.08)</td>
<td>0.71 (0.11)</td>
<td>0.02 (0.10)</td>
<td>current study</td>
</tr>
<tr>
<td>New Mexico</td>
<td>0.26 (0.03-0.57)</td>
<td>0.73 (0.42-0.97)</td>
<td>0.01 (0.01-0.15)</td>
<td>Healy et al 2017*</td>
</tr>
<tr>
<td>New Mexico</td>
<td>0.33 (0.12)</td>
<td>0.62 (0.12)</td>
<td>0.06 (0.06)</td>
<td>Klimentidis et al 2009</td>
</tr>
<tr>
<td>Colorado</td>
<td>0.33</td>
<td>0.67</td>
<td>NA</td>
<td>Merrinwether et al. 1997</td>
</tr>
<tr>
<td>Colorado</td>
<td>0.34 (0.02)</td>
<td>0.63 (0.02)</td>
<td>0.03 (0.02)</td>
<td>Bonilla et al 2004</td>
</tr>
<tr>
<td>NE Mexico</td>
<td>0.40 (0.03)</td>
<td>0.55 (0.03)</td>
<td>0.05 (0.03)</td>
<td>Cerda-Flores et al</td>
</tr>
<tr>
<td>Mexico City</td>
<td>0.40</td>
<td>0.59 (0.08-0.98)</td>
<td>0.01</td>
<td>Healy et al 2017</td>
</tr>
<tr>
<td>Mexico City</td>
<td>0.40 (0.07)</td>
<td>0.57 (0.07)</td>
<td>0.03 (0.01)</td>
<td>Wang et al 2008</td>
</tr>
<tr>
<td>San Francisco</td>
<td>0.43 (0.15)</td>
<td>0.48 (0.15)</td>
<td>0.08 (0.04)</td>
<td>Risch et al 2009</td>
</tr>
<tr>
<td>NE Mexico</td>
<td>0.56 (0.27-0.81)</td>
<td>0.38 (0.17-0.71)</td>
<td>0.06 (0.01-0.12)</td>
<td>Martinez-Fierro et al</td>
</tr>
<tr>
<td>Monterrey, Mexico</td>
<td>0.57 (0.14)</td>
<td>0.38 (0.13)</td>
<td>0.04 (0.02)</td>
<td>Young et al 2014</td>
</tr>
<tr>
<td>Mexico (11 locations)</td>
<td>0.58 (0.35-0.85)</td>
<td>0.38 (0.13-0.61)</td>
<td>0.04 (0.02-0.07)</td>
<td>Moreno-Estrada et al 2014</td>
</tr>
<tr>
<td>Mexico City</td>
<td>0.62 (0.18)</td>
<td>0.32 (0.17)</td>
<td>0.06 (0.03)</td>
<td>Risch et al 2009</td>
</tr>
</tbody>
</table>

* reanalysis of data from Wang et al 2008

Figure 4c and d show the NMS portion of Figure 4b. The plots highlight variation
within and among the NMS ethnic identity groups using 95 per cent confidence interval
ellipses. Figure 4c shows that “Hispanic” and “Latino/a” PC factors range widely, while
“Chicano/a” factors are more tightly constrained. Figure 4d captures separation on PC 2
between “Mexican” and “Mexican American” vs. “Spanish” and “Nuevomexicano/a”.

Figure 5 indicates that the among-group differences in the PCA plots correspond with
among-group variation in genomic ancestry. The figure contains violin plots of ancestry
for each NMS group. The mean estimates for each NMS group are listed in table S2.
“Spanish” and “Nuevomexicano/a” NMS have significantly higher European ancestry and significantly lower Native American and African ancestry than “Mexican American” NMS (multiple-test adjusted p < 0.0017).

Figure 5.

*European, Native American and sub-Saharan African ancestry. (A) European ancestry. (B) Native American ancestry. (C) Sub-Saharan African ancestry. Each violin plot contains a conventional box plot marking median (circles), minimum, first and third quartile, and maximum values. Each box plot is surrounded by a mirrored normal density trace. Dashed lines show the mean ancestry for the full NMS sample. See also table S2.*

The pattern is similar for the “Spanish”-“Mexican” and “Nuevomexicano/a”-“Mexican” comparisons, though only the “Spanish”-“Mexican” comparison for African ancestry is significant at the multiple-test adjusted threshold (table S3). The pattern is also similar for comparisons of the other groups to “Mexican” and “Mexican American,” but only the “Hispanic”-“Mexican American” and “Chicano/a”-“Mexican” American comparisons are statistically significant.
Figure 6 contains violin plots of genomic ancestry associated with answers to the two questions about family ties to geographic locations. The plots provide further support for a historically-rooted distinction between descendants of early Spanish settlers and more recent immigrants from Mexico. NMS with ancestors from Mexico alone ("Mexico only") had significantly higher Native American and African ancestry, and significantly lower European ancestry, than NMS with ancestors from Spain only (multiple-test adjusted p < 0.0017). For family ties to colonial New Mexico, NMS who reported ties through their mothers had significantly higher European and lower Native American and African ancestry than NMS who did not report ties to colonial New Mexico ("Neither").
Figure 6.

European, Native American, and sub-Saharan African ancestry by family ties to geographic locations. (A)–(C): Ancestors from Mexico and Spain. (D)–(F): Family ties to colonial New Mexico. Dotted vertical lines show the mean ancestry for the full NMS sample.

Birthplace

Figure 7 and table S4 show the birthplaces of the 507 study participants. The vast
majority (n = 484) were born in the US Southwest or Mexico. The largest concentration of birth locations was in Bernalillo County (n = 201), which contained 32 per cent of the New Mexico population in 2010. Eighty-seven participants were born in counties located in the San Luis Valley in New Mexico and Colorado, hereafter abbreviated SLV. Other participants were born in other US states (n = 13), and other regions of Latin America (n = 7), Europe (n = 2), and Asia (n = 1). The following sections describe variation in birthplace among the seven groups (listed in alphabetical order).
Figure 7.

Chicano/a (Figure 7a). NMS who identified primarily as “Chicano/a” were frequently born in the US Southwest (94 per cent) and in New Mexico in particular (81 per cent). None were born in Mexico or the SLV. However, a small proportion of their parents were born in Mexico (4 per cent) and the SLV (5 per cent), and an even larger proportion of their grandparents (Mexico 12 per cent; SLV 11 per cent).

Hispanic and Latino/a (Figure 7b and 7c). The “Hispanic” pattern is similar to the “Chicano/a” pattern in terms of the high frequency of birthplaces in the US Southwest (94 per cent) and New Mexico (84 per cent), and the low frequency in Mexico (0 per cent). The parents and grandparents of “Hispanic” and “Latino/a” NMS were more frequently born in Mexico than the participants themselves, though the frequencies were still low (parents 3 per cent; grandparents 6 per cent). A notable difference from “Chicano/a” NMS, however, is the relatively high frequency of birth in the SLV of “Hispanic” NMS (16 per cent), and especially their parents (27 per cent) and grandparents (29 per cent). These SLV proportions are second only to those for “Spanish” NMS (see below). The “Latino/a” pattern is almost identical to the “Hispanic” pattern.

Mexican and Mexican American (Figures 7d and 7e). The birthplaces of individuals who identified primarily as “Mexican” are distinctive from all other groups. Half of “Mexican” NMS were born in Mexico and half were born in the Southwest. Only half of the latter were born in New Mexico, and, of these, none were born in the SLV.
Mexican birth was common for parents (69 per cent) and grandparents (73 per cent). New Mexican birth was less common (parents 19 per cent; grandparents 13 per cent), and SLV birth was rare (parents 3 per cent; grandparents 3 per cent).

The “Mexican American” group is intermediate between “Mexican” and the other five groups. They were less frequently born in Mexico (16 per cent) than “Mexican” NMS. They were more frequently born in the Southwest (82 per cent) and New Mexico (61 per cent) compared to “Mexican” NMS, as were their parents (Southwest 49 per cent; New Mexico 29 per cent) and grandparents (Southwest 34 per cent; New Mexico 22 per cent). They were also more frequently born in the SLV than “Mexican” individuals, though the proportions were still low (participants 5 per cent; parents 7 per cent; grandparents 6 per cent).

For “Mexican” and “Mexican American” NMS who were born in Mexico, about 50 per cent were born in the state of Chihuahua, located directly south of New Mexico (Figure 7). Chihuahua was also a common birthplace of Mexican-born parents and grandparents, though their birthplaces ranged more widely, encompassing 12 other states within Mexico.

*Spanish and Nuevomexicano/a* (Figure 7f and 7g). The birthplaces of NMS who identified primarily as “Spanish” and “Nuevomexicano/a” were also distinctive from other groups, especially “Mexican.” Almost all “Spanish” NMS were born in the Southwest (97 per cent) as were most of their parents (92 per cent) and grandparents (91 per cent). Most of these individuals were born in New Mexico, and a large proportion was born in the SLV (participants 31 per cent; parents 38 per cent; grandparents 41 per cent). No “Spanish” NMS or their parents or grandparents were born in Mexico. The
“Nuevomexicano/a” pattern was essentially identical.

Summarizing these results, “Mexican” and “Mexican American” NMS have the strongest ties to Mexico of any of the groups; “Mexican” participants themselves have strong ties to Mexico, whereas “Mexican American” NMS tend to have stronger ties through their parents and grandparents. “Spanish” and “Nuevomexicano/a” have the strongest ties to New Mexico and, within New Mexico, the SLV. “Hispanic” and “Latino/a” NMS also have strong ties to New Mexico and the SLV and relatively few ties to Mexico. These patterns hold for parents and grandparents of the groups, though they tend to be more variable, especially for the “Chicano/a” group.

Discussion

**Hispanic.** Ethnic identity nomenclatures are constantly shifting in response to complex social and political factors. The term “Hispanic,” for example, has only been in widespread use by the US government for about 50 years, and it first appeared on the US Census in 1980 (US Census Bureau 1980). Along with “Latino/a,” which first appeared on the 2000 census, “Hispanic” was ostensibly codified by the government to permit tracking of the social and economic conditions of all peoples of Spanish-speaking origins in the US.

“Hispanic” was the most common choice among our NMS study participants. NMS who used the term, as well as their parents and grandparents, were often born in New Mexico, and, more specifically, in the SLV. None of the individuals who used the term “Hispanic” were born in Mexico, and their parents and grandparents were rarely born in Mexico. “Hispanic” individuals also frequently reported family ties to colonial New Mexico, and they had relatively high European and low Native American and
African ancestry. We found no connection between the term “Hispanic” and participant age; the age of NMS who identified as “Hispanic” fell near the mean of the entire NMS sample and did not differ significantly from that of other groups.

Overall, these results indicate that “Hispanic” is a common choice among NMS who perceive connections to colonial New Mexico. As such, it does not appear to be a catchall term for all NMS.

However, a recent study of NMS conducted in southern New Mexico suggests that this may be changing (Doan and Stephan 2006). The study found that a large proportion of NMS chose the term because they believed that it might improve their chances of getting scholarships and jobs. This finding indicates that nomenclatures shift in response to perceived costs and benefits. Such shifts have the potential to mask more deeply historically-rooted aspects of ethnic identity.

**Spanish, Nuevomexicano/a, Mexican, and Mexican American.** Individuals who identified as “Spanish” and “Nuevomexicano/a” had high prevalence of birthplaces in the SLV, low prevalence of parents or grandparents born in Mexico, strong ties to colonial New Mexico, and relatively high European and low Native American and African ancestry. In contrast, individuals who identified as “Mexican” and “Mexican American” were frequently born in Mexico, seldom born in the SLV, expressed fewer ties to the colonial period, and had relatively low European and high Native American and African ancestry. Our results show that NMS who use the term “Nuevomexicano/a” are younger than NMS who use the term “Spanish.” As both groups report ties to colonial New Mexico, the age difference could reflect increasing awareness of the history of the broader region and recognition that emphasis on “Spanish-ness” in the early 20th century
reflected sociopolitical factors more than actual ancestry (Nieto-Phillips 2008). This correspondence between age and identity terms highlights the fluid nature of ethnic identity nomenclatures, even as terms remain tied to the history of the region. This finding is all the more remarkable because the rate of migration from Mexico and other regions of the US has been high for decades, and because intermarriage and cultural and linguistic exchange among resident and migrant groups has been common in the region since the colonial era (Brooks 2002; Marshall 1939; Tjarks 1978).

**Latino/a.** NMS who chose “Latino/a” were broadly similar to those who chose “Hispanic” with respect to birthplace, family ties to geographic locations, and genomic ancestry. The infrequent use of “Latino/a” puts New Mexico in contrast with other parts of the United States, specifically urban locations with high proportions of people from Latin American countries other than Mexico (Oboler 1995).

**Chicano/a.** Fifty-three NMS identified primarily as “Chicano/a.” Most of these individuals were born in Bernalillo County; none were born in Mexico, and none were born in the SLV. However, about 5 per cent of their parents and 12 per cent of their grandparents were born in Mexico, and roughly equal percentages were born in the SLV. In this sense, “Chicano/a” individuals were intermediate between “Spanish,” “Nuevomexicano/a,” and “Hispanic” on the one hand and “Mexican” and “Mexican American” on the other. Like the other ethnic identity terms, the use of “Chicano/a” is embedded in the history of the region. As in other regions, in New Mexico, it served as a rejection during the 1960s and 1970s of discrimination, racism, and economic exploitation. In New Mexico, adherents to this new identity and terminology included
native New Mexicans, many of whom had family ties to colonial New Mexico (Gonzales 1993; Maciel and Gonzales-Berry 2000).

Overall, the significant differences in genomic ancestry between several NMS groups are consistent with the existence of ethnic identity-related genetic substructure in New Mexico. These patterns of ancestry were no doubt shaped not only by initial Spanish colonization but also by interactions between groups of individuals who migrated to New Mexico from different locations throughout the US and Mexico after the colonial period. Healy et al. (2017) provide evidence for the existence of ethnic-based genetic substructure in 12 of 13 admixed populations located in Central and South America. Long-range migration has been common throughout human history (Pickrell and Reich 2014) and has no doubt played an important role in shaping patterns of social identity and mating, often producing correlations between patterns of ethnic identity and genomic diversity.

Limitations

A potentially important limitation of our study is that our sample is unlikely to be random with respect to ethnic identity among NMS as a whole because our interviews were conducted in English. In the context of interviews conducted in English in an academic setting, study participants may have been less likely to use terms that they perceive to be associated with discrimination (Doan and Stephan 2006). Additionally, our participants were more highly educated than the New Mexico population as a whole and were less frequently foreign-born (table S1). Though we did not ask participants about citizenship or resident status, all were most likely US citizens or legal US residents. Our results must be tempered by these factors, as well as the fact that individuals may choose
ethnic identity terms based on characteristics of the interviewer, including age, gender, ethnic identity, and other demographic factors (Doan and Stephan 2006).

With respect to genomic ancestry, the parental sources that we used in our analyses are only proxies for the true parental sources. The true sources derive from diverse locations in Europe, Africa, and the Americas. Additionally, different estimation methods and variation in user-defined parameters associated with those methods can lead to substantial variation in ancestry estimates. Fortunately, our inferences rely only on relative differences in ancestry proportions between individuals and groups.

We are limited in the extent to which we can connect differences in genomic ancestry to specific colonial and postcolonial migration events. This limitation in part owes to the fact that we did not collect data from Native American groups in the region.

Finally, our goal was not to test the personal narratives underlying the ethnic identity terms used by NMS. In this regard, we did not show or intend to show that European ancestry derived from Spain. Some European ancestry in NMS who perceive family ties to colonial New Mexico most certainly derives from non-Spanish Europeans who migrated during and after the US territorial period. This migration played and continues to play an important role in the formation of ethnic identity and its attendant nomenclatures in New Mexico (Gonzales 1993).

Acknowledgments

We thank our participants for their time and interest in our study. We thank Bonnie Young, Jessica Gross, Brandon Pettit, and Shannon Bermea for assistance with data collection. We thank Laura Gomez, Jack Baker, David Dinwoodie, and Sylvia Rodriguez for assistance with research design.
Funding

This work was supported by the National Science Foundation [BCS #0962825]; University of New Mexico [Cross Campus Collaboration in the Life Sciences].

References


chip_kits.html

Institute of Medicine. 2009. Race, Ethnicity, and Language Data: Standardization for
Health Care Quality Improvement [Internet]. Washington, DC: The National
Academies Press. http://www.ahrq.gov/research/findings/final-
reports/iomracereport/index.html.


Klimentidis, Y. C., G. F. Miller, and M. D. Shriver. 2009. Genetic admixture, self-
reported ethnicity, self-estimated admixture, and skin pigmentation among
Hispanics and Native Americans. American Journal of Physical Anthropology

Lamstein, A., and B. P. Johnson Package ‘choroplethr’ [Internet]. 2017. Available:
http://www.arilamstein.com/open-source

Legacy: Forms, Agencies, and Discourse. 1st ed. Albuquerque: University of New
Mexico Press.

of New Mexico. Albuquerque: University of New Mexico Press.

Marshall, C. E. 1939. The birth of the Mestizo in New Spain. The Hispanic American

Martinez-Fierro M., J. Beuten, R. Leach, E. Parra, M. Cruz-Lopez, H. Rangel-Villalobos,
doi:10.1038/jhg.2009.65.

Contreras, V. Acuña-Alonzo, et al. 2014. The genetics of Mexico recapitulates
doi:10.1126/science.1251688.

Miller, M. 1976. Mexican Americans, Chicanos, and others: ethnic self-identification and
selected social attributes of rural Texas youth. Rural Sociology 41:234–47.


Roybal, E. Joint resolution relating to the publication of economic and social statistics for Americans of Spanish origin or descent. [Internet]. H. J. Res. 1083 1974. Available: https://www.govtrack.us/congress/bills/93/hjres1083


CHAPTER 4: ASKING ABOUT ETHNIC IDENTITY: COMPARING OPEN-ENDED AND FIXED-CHOICE ETHNIC NOMENCLATURE RESPONSES IN NEW MEXICANS OF SPANISH-SPEAKING DESCENT

Abstract

Data collection on ethnic identity has moved increasingly towards the use of fixed-choice questioning, but little research has been conducted to compare the informativeness of data obtained using this format versus open-response techniques. Our study compares ethnic identity data collected using both approaches in structured interviews of 507 New Mexicans of Spanish-speaking descent (NMS). We find higher levels of concordance between identity nomenclature elicited by the two forms of questioning than in previous research and attribute this to background research that we had done previously to identify important ethnic identity terms among NMS. We also find that the data obtained from open-response and fixed-choice questioning together provide richer information than either dataset alone. Among our NMS participants, the term “Hispanic” was overwhelmingly the most commonly-used for both open-ended and fixed-choice responses, whereas “Latino/a”, was rarely named or selected despite its popularity as a panethnic term in other regions of the United States. In open responses, “Spanish” and closely-related terms were only second to “Hispanic” in frequency, but the academic construction “Nuevomexicano/a” surpassed “Spanish” in frequency of selection when it was offered on our list of terms. Participants identifying as “Mexican” or “Mexican American” showed greater strength of commitment to these terms than did others, suggesting a distinction between practical and symbolic ethnic identities in New Mexico.

We suggest that ethnic identity can be understood with greater confidence when
researchers first identify current, relevant ethnic nomenclature in the population of interest and then use both open-response and fixed-choice identity questions.

**Introduction**

Research in the social sciences has consistently shown ethnic identity to be unstable and subjective (Burton, Nandi, & Platt, 2010; Nagel, 1994). Because it is a social construction, ethnic identification is subject to the specific environment affecting an individual at the moment of identification. Some of the factors affecting ethnic identification include current social trends, whether the context is official or informal in nature, cues based on the ethnic affiliation of the person asking the question, the language being used, the person’s own phenotype, and the wording and format of the question. In terms of formatting, some interviews and questionnaires provide respondents with lists of categories and ask them to choose the term/s that best identifies them, called the fixed-choice option, while others ask respondents to categorize their own ethnicity in an open-ended fashion.

Responses to fixed-choice classification systems for ethnic identity have the advantage of being more easily tabulated and less prone to missing or unusable data than open-ended responses, but some researchers argue that accuracy and legitimacy in assessment can only be obtained through individual self-designation (Stephan & Stephan, 2000). This latter position assumes that open-ended and fixed choice are entirely different approaches and likely to produce responses that are contradictory, inconsistent, or non-comparable. On the other hand, Aspinall (2012) asserts that open-response questions can provide important insights, but only if they are paired with list-based questions that can help provide context. Other researchers propose measuring ethnicity using more nuanced
methods that take into account its dynamic and multidimensional nature (Burton et al., 2010; Phinney & Ong, 2007; Williams & Husk, 2013). Importantly, while controversy concerning the merits of open and closed modes of inquiry in surveys has persisted for nearly a century in social science research in general, few researchers have conducted empirical research to compare the utility of the data collected through the two different methodological approaches (Schuman & Presser, 1979).

Terms used to describe ethnic identity are rarely given precise definitions even when used in formal contexts such as research and policymaking, and thus are subject to different interpretations by different users (Kaplan, 2014). This issue can affect interpretations of data from both fixed-choice and open-ended ethnic identity questions: whether selecting from a fixed list or using one’s own words, respondents may pick the same term but have different understandings of its meaning, or alternatively use different terms to describe the same cultural background and descent. Despite this and other difficulties in measuring ethnicity, ethnic classification has important practical applications in health research and development and application of government-level policies intended to benefit minorities (Williams & Husk, 2013). Its social importance also underscores the obligation of social scientists to understand ethnic classification systems and their effects (Stephan & Stephan, 2000).

In this paper, we argue that the optimal approach to acquisition of data on identity construction among subjects of a given ethnic population is one that utilizes both open-ended and fixed-choice types of inquiry. If designed appropriately, the two approaches can be combined to test for both consistency and strength of identity and ethnic self-labeling.
Census data showed the “Hispanic” population of the United States to comprise 57.5 million people in 2016 (U.S. Census Bureau, 2017), constituting 17.8% of the total US population. This makes it the largest ethnic or racial minority group in the country, and its proportion of the population continues to grow. Although sometimes treated as a homogeneous entity, the group shows internal variation in its geographical, historical and cultural origins (Gimenez, 1989). Moreover, different ethnic identity terms are utilized by individuals to emphasize their connections to those origins. Researchers and policymakers have fluctuated in the terms they have used to describe people of Spanish-speaking descent in the United States and elsewhere, and in how many subcategories they recognize within the broader grouping. They also have struggled to determine ways to collect data about “Hispanic” ethnic identity without losing out on information concerning groups defined by race and national origins. Despite this, there is no available body of research comparing the information gained from open-ended vs. fixed-choice identity questions in Americans of Spanish-speaking descent.

New Mexico is the state with the highest proportion of people identifying as “Hispanic or Latino”, 46.3% of the population in 2010 (U.S. Census Bureau, 2010), and has a long and complex history of settlement, migration and political change that make it a compelling location for researching ethnic identification among those with Spanish-speaking ancestors. Additionally, social science researchers already have extensively examined the phenomenon of changes in ethnic nomenclature throughout New Mexican history (e.g., (Gómez, 2007; Gonzales, 1993; Gonzalez, 1969; Montgomery, 2002; Nieto-Phillips, 2008), and their work provides a rich backdrop for exploring the current usage of different ethnic identity terms. We have previously shown that self-identification in
New Mexicans of Spanish-speaking descent (NMS) is linked to family ties, geographic regions, and genomic ancestry (Healy et al., 2018), suggesting that their ethnic identity term choices can provide meaningful information despite the situational and changeable character of ethnic identity.

In structured interviews, we asked about ethnic identity in 507 participants recruited as “New Mexicans of Spanish-speaking descent”, first in an open-ended format and later using a fixed-choice question with a list of terms. Our list of terms was formed from the most popular terms used in open-ended responses about ethnic identity in pilot research conducted prior to this study. In this paper, we compare overall and individual-level participant responses to these two forms of inquiry. Further, we discuss the implications for the potential of open-ended and fixed-choice forms of ethnic identification to provide information for researchers and policymakers in the health and social sciences. We conclude with a discussion of important ethnic identity terms that have been utilized by NMS, historical trends in the use and significance of these terms, and how our findings contribute to the question of how to best measure ethnic identity in research. Participant comments that were documented during our interviews enable us to look at variation in the meaning of specific terms to different people, highlighting the difficulties researchers face when creating, implementing and interpreting questions about ethnic identity.

**Materials and Methods**

We obtained structured face-to-face interview responses from 507 NMS recruited from Bernalillo County, New Mexico (Healy et al., 2018) to two identity questions, one open-ended and one fixed-choice. The latter included an open-ended “Other” option.
The interviewer set the context with the following statement: “Many New Mexicans have Spanish-speaking ancestors, but these people don’t all have the same background or family history, and they don’t all see themselves as belonging to the same group of people.” The interviewer then asked, “What are the groups?” and followed up with our open-ended identity question: “Which group do you identify with?”

The fixed-choice identity question came later in the interview. The interviewer asked, “With which of these groups do you identify?” and provided a card with these responses, which were ordered alphabetically:

a. Chicano/a
b. Hispanic
c. Latino/a
d. Mexican
e. Mexican American
f. Nuevomexicano/a
g. Spanish
h. Other:_________

Interviewers asked participants to select a first-choice term, then asked if there was a second term that they identified with, then if there was a third term, and finally if there were any further terms that they would use to identify themselves. When participants chose “Other”, they were asked what term they would use instead of those provided.

To compare overall-level participant responses to the open-ended and fixed-choice identity questions, we first look at what terms were used when we asked NMS to describe their ethnic identity in an open-ended fashion, and in what frequency each term
appeared. We then compare the overall responses from the open-ended question to the fixed-choice responses, focusing on several features: 1) the proportion of open-ended responses that were present on our list of terms, 2) the frequencies of fixed-choice responses compared to the frequencies of those same terms appearing in open-ended responses, and 3) any evidence that important terms were missing from our list, i.e., terms that appeared frequently in open-ended responses that were not included in our fixed-choice question.

We then move to the individual level and compare the self-designated terms provided by participants to the terms those specific participants chose from our list of terms. We again focus on several features: 1) for those who described themselves on our open-ended question using listed terms, we examine the adherence to those term choices by checking how often they chose the same terms in the fixed-choice question, 2) for those who described themselves on our open-ended question using unlisted terms, we look for patterns between unlisted terms, or words and key elements within the terms, and the terms they did choose from our list, and 3) we check for important missed terms by looking at participants who answered the open-ended question with unlisted terms and then chose “Other” on the fixed-choice question, i.e., participants who did not find any of our listed terms appropriate for describing their ethnic identity.

Results

Overall Responses

On the open-ended question, most participants (394) described themselves using a single term due to the phrasing of the question being singular in form (“Which group do you identify with?”). There were 85 participants who used two terms, and 28 who used
three. No participants gave more than three terms. In contrast, 409 participants named three terms on the fixed-choice question, where the interviewer provided the opportunity for them to give first, second and third-choice terms. There were 20 participants who chose only one term on the fixed-choice question, 75 who chose two, and three who chose more than three terms.

Table 1 shows the terms that participants used to describe their identities in an open-ended fashion, ordered according to the frequency with which they were used overall. Additional columns show only the participants’ primary terms (the term that the participant stated was most important if a preference was stated, or the one that was provided first) on the open-ended question, and then show the tallies for both overall and first-choices when participants were asked to choose identity terms from a list.

Table 1.

*Comparison of terms and term frequencies in responses to open-ended and fixed-choice identity questions.*

<table>
<thead>
<tr>
<th>Identity Terms Used/Chosen by NMS, ordered by overall number of uses on Open-Ended Identity question</th>
<th>Open-Ended, overall uses</th>
<th>Open-Ended, 1st-choice</th>
<th>Fixed-choice, overall choices</th>
<th>Fixed-choice, 1st-choice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hispanic*</td>
<td>243</td>
<td>223</td>
<td>415</td>
<td>223</td>
</tr>
<tr>
<td>Spanish*</td>
<td>65</td>
<td>45</td>
<td>230</td>
<td>62</td>
</tr>
<tr>
<td>Chicano/a*</td>
<td>58</td>
<td>40</td>
<td>187</td>
<td>53</td>
</tr>
<tr>
<td>Mexican American*</td>
<td>34</td>
<td>23</td>
<td>101</td>
<td>38</td>
</tr>
<tr>
<td>Mexican*</td>
<td>32</td>
<td>23</td>
<td>57</td>
<td>16</td>
</tr>
<tr>
<td>Latino/a*</td>
<td>29</td>
<td>18</td>
<td>169</td>
<td>16</td>
</tr>
<tr>
<td>Spanish American</td>
<td>26</td>
<td>22</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>New Mexican</td>
<td>22</td>
<td>19</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>American</td>
<td>11</td>
<td>9</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Mestizo/a</td>
<td>10</td>
<td>8</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mexicano/a</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Norteno/a</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Nuevomexicano*</td>
<td>6</td>
<td>5</td>
<td>205</td>
<td>75</td>
</tr>
<tr>
<td>Term</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>American of Spanish descent</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Anglo</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Hispanic American</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hispano/a</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Native American</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Native New Mexican</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Italian</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Mutt</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Puerto Rican</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Spaniard</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Caucasian</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Coyote</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Heinz 57</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mexican descent</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mixed</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Native</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Northern New Mexican</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>White</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Terms used once: A human being\(^a\), African American\(^a\), American combination plate\(^a\), American of Hispanic descent\(^a\), American of Latin American descent\(^a\), American of Mexican descent\(^a\), American with Spanish and Italian roots\(^a\), Asian\(^c\), Atrisqueno/a\(^**c\), Black\(^a\), Caribena\(^b\), Caucasian Hispanic\(^a\), Caucasian Spanish\(^a\), citizen of the Earth\(^a\), Citizen of the United States\(^a\), Cuban American\(^**c\), European\(^a\), European descent\(^b\), European Hispanic\(^a\), European Spanish\(^b\), Filipino American\(^**c\), First-generation Hispanic\(^a\), French\(^a\), Half-Spanish/Half-Black\(^b\), Hispanic and Native American\(^b\), Hispanic descent\(^a\), Hispanic New Mexican\(^a\), Hispanic of New Mexican descent\(^b\), Irish\(^a\), Latin American\(^a\), Mexica\(^a\), Mexican American/Spanish\(^b\), Mezcla\(^a\), Mixed blood\(^a\), Mixture of Hispanic and White\(^**b\), Multi-racial\(^a\), Native American\(^b\), Native Mexican\(^**c\), New Mexican of Spanish ancestry\(^a\), None\(^**b\), Northern New Mexico Hispanic\(^a\), Panamanian\(^**c\), Samoan\(^a\), Sephardic\(^a\), Sephardic Jew\(^a\), Sephardim\(^b\), Spanish American as a cultural reference but with Native American, too\(^b\), Spanish/Italian/Caucasian mutt\(^**b\), Spanish-Native hybrid\(^**b\), United States citizen\(^b\)

**Totals for terms used once:**

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>31</td>
<td>31</td>
<td>23</td>
<td>9</td>
</tr>
</tbody>
</table>

**Other* totals (for open-ended, "Other" includes any term not included later in our list-based question)**

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>181</td>
<td>130</td>
<td>55</td>
<td>24</td>
</tr>
</tbody>
</table>

* Term present in our list of terms. For the list-based question, all non-list terms were supplied under the category "Other"

** Used as a first-choice term (after selecting "Other") on list-based identity question

\(^a\) Term used once on open-ended identity question (all were first choices)

\(^b\) Term used once on list-based identity question

\(^c\) Term used once on both open-ended and list-based questions
There was a high degree of overlap in the terms NMS used to describe themselves in the open-ended identity question and the terms later offered to them in the fixed-choice identity question: the six most frequent terms used on the open-ended question comprised six of the seven terms provided in the fixed-choice question. On the open-ended question, there were 636 total term uses, and 467 (73%) of those were instances of the seven terms included in the fixed-choice question. Of the 507 participants, 377 (74%) used a primary identity term that was among the seven choices in the fixed-choice question.

Figure 1 shows the frequencies of use of each of the seven terms comprising the fixed-choice list for both the open-ended and fixed-choice questions. The most common choice for both questions was “Hispanic”. In the open-ended question, 243 participants (48%) used “Hispanic” overall to describe themselves, and 223 (44%) used it as their primary identity term. On the fixed-choice question, 415 participants (82%) selected Hispanic in any order, making it clear that this term is one NMS are likely to identify with even when it’s not the first term they would choose. Like with the open-ended question, 223 (44%) picked “Hispanic” as their first choice on the fixed-choice question, although the 223 do not overlap completely, as discussed later.

Although used widely in the media and as an alternative to “Hispanic” on many forms and popularly in many urban areas of the United States, “Latino/a” was less popular in NMS. Only 29 participants (6%) used “Latino/a” when answering the open-ended question, and it had 169 overall uses but only 16 first-choice picks in the fixed-choice question.

The term “Nuevomexicano/a” did not appear often in open-ended responses (six times overall, five times as a primary term), but was chosen 75 times as a first choice and
205 times overall on the fixed-choice question. Only “Hispanic” and “Spanish” (which had 230 overall selections, 62 first choice) were selected more often overall on the fixed-choice question than “Nuevomexicano/a.”

For the open-ended question, 130 participants described themselves primarily using terms that were not among the seven terms on the fixed-choice list. On the fixed-choice question, when the option “Other” was provided, 24 participants selected it as their first choice, while the other 483 chose a term from the fixed-choice list.

Figure 1.

Comparison of frequencies of identity terms used by NMS as open-ended responses and those chosen as fixed-choice responses. The first (blue) and third (gray) bars show the overall frequencies for each term in open response and fixed-choice responses, respectively, while the second (orange) and fourth (yellow) bars show how often the term was given as a first choice.

Table 2 shows overall and individual-level comparisons in the selection of identity terms between the open-ended and fixed-choice identity question. The terms
“Hispanic”, “Chicano/a”, “Mexican American” & “Nuevomexicano/a” showed the strongest adherence at 75% or greater, indicating that the individual participants who used those terms on their own in our open-ended question also selected them when we provided the list that contained them. Of 18 participants who described themselves as “Latino/a” on the open-ended question, only nine chose it as their first choice from our list for the lowest adherence at 50%.

Table 2.

Overall & individual term changes between open-ended & list-based identity questions.

<table>
<thead>
<tr>
<th>List Term*</th>
<th>A</th>
<th>B</th>
<th>A/B</th>
<th>C</th>
<th>C/A</th>
<th>C/B</th>
<th>1- C/B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicano/a</td>
<td>40</td>
<td>53</td>
<td>0.75</td>
<td>30</td>
<td>0.75</td>
<td>0.57</td>
<td>0.43</td>
</tr>
<tr>
<td>Hispanic</td>
<td>223</td>
<td>223</td>
<td>1</td>
<td>178</td>
<td>0.8</td>
<td>0.8</td>
<td>0.2</td>
</tr>
<tr>
<td>Latino/a</td>
<td>18</td>
<td>16</td>
<td>1.13</td>
<td>9</td>
<td>0.5</td>
<td>0.56</td>
<td>0.44</td>
</tr>
<tr>
<td>Mexican</td>
<td>23</td>
<td>16</td>
<td>1.44</td>
<td>14</td>
<td>0.61</td>
<td>0.875</td>
<td>0.125</td>
</tr>
<tr>
<td>Mexican American</td>
<td>23</td>
<td>38</td>
<td>0.61</td>
<td>19</td>
<td>0.83</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Nuevomexicano/a</td>
<td>5</td>
<td>75</td>
<td>0.07</td>
<td>4</td>
<td>0.8</td>
<td>0.05</td>
<td>0.95</td>
</tr>
<tr>
<td>Spanish</td>
<td>45</td>
<td>62</td>
<td>0.73</td>
<td>29</td>
<td>0.64</td>
<td>0.47</td>
<td>0.53</td>
</tr>
<tr>
<td>Other</td>
<td>130</td>
<td>24</td>
<td>5.42</td>
<td>21</td>
<td>0.16</td>
<td>0.875</td>
<td>0.125</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>507</td>
<td>507</td>
<td>304</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Term order is alphabetical, as shown to participants during interview
A. Number of times used first in open-ended identity question
B. Number of times chosen as first choice from list
A/B: Relative popularity as open choice vs. on list
C. Number of times B matched A for an individual
C/A: Adherence to open choice
C/B: How often list choice was not new, i.e., the frequency with which participants chose a listed term that they had already used on the open-choice question
1-C/B: How often list choice was new, i.e., the frequency with which participants chose a listed term that they had not already used on the open-choice question
The use of non-list terms dropped when we asked our fixed-choice question.

“Other” was chosen only 24 times compared to 130 participants using non-list terms on
open-ended question. We therefore saw increases in the use of several list terms on the fixed-choice question. This is consistent with a known tendency for interview/survey respondents to restrict themselves to the options given on a closed question and give a much larger variety of responses to the same open-ended question (Schuman & Presser, 1979). “Chicano/a”, “Mexican American”, “Spanish”, and especially “Nuevomexicano/a” were all “attractors” on our list, meaning that they were chosen from the list as a first-choice term more often than they were named in open responses.

“Hispanic” stayed the same with 223 primary uses/first choices on both questions, indicating its cultural power. There was high overlap on the use of this term by individuals between questions: there were 178 matches, i.e., people who described themselves as “Hispanic” and then also picked it as their first choice on the list (see Table 3). This means that 45 people named “Hispanic” as their primary open response but then selected another term as their first choice in the fixed-choice question, and 45 others used another term in open response but then chose “Hispanic” from the list as their first-choice term.

Only “Latino/a” and “Mexican” showed fewer uses on the list than in open responses. People were unlikely to choose the term “Mexican” if they had not already used it in the open response: only two people who had not already called themselves “Mexican” on the open response question chose it as their first choice from the list.

People were not likely to choose Other on the fixed-choice question if they had already used a term that fell on our list on the open response question. Of 24 people who chose “Other” from the list, 21 of those had already used terms not on our list, leaving three people who had used a term that was on our list in the open response, but then
named a different, off-list term on the fixed-choice question. Two of these participants described themselves as “Chicano/a” and one as “Hispanic” on the open-ended question. The two “Chicano/a” participants then chose “Other” and gave the terms “Norteño/a” and “Spanish American” to describe themselves, while the “Hispanic” participant chose “Other” and then declined to provide a term.

Of people who chose “Spanish” on the fixed-choice question (n=62), a majority included "Spanish" or "Spaniard" in some form in their open-ended answers, but there were variations: 29 had already called themselves “Spanish”, but the remaining 33 included individuals who had used the terms “Spanish American” (9), “American of Spanish descent” (3), “Caucasian Spanish” (1), “Spaniard” (1) and “Spanish of Mexican descent” (1) when describing themselves. However, there were also individuals who chose “Other” and provided the specific terms “Spanish American” (n=6) and “American of Spanish descent” (n=1) as their first-choice terms on our list-based question as opposed to selecting “Spanish” from our list.

The use of “Nuevomexicano/a” as a primary term surged from five uses in the open-response question to 75 first choices on the fixed-choice question. Four of the five who used it in open-response adhered to the term on the fixed-choice question, while the other individual chose “Hispanic” first, “Spanish” second and “Chicano” third from the fixed-choice question without selecting “Nuevomexicano/a” at all. Of the 71 who had not named it themselves but chose it from the list, 39 had used an off-list term in the open-response question. Of those, the biggest contribution came from the 19 people who had called themselves “New Mexican” in the open response, of whom 13 selected “Nuevomexicano/a”. For list terms, the largest numbers of shifts to “Nuevomexicano/a”
on the fixed-choice question were from those who had called themselves “Hispanic” (19) and “Chicano/a” (6) in open response.

Another shift was from “Mexican” on the open response to “Mexican American” on the list. Of the 23 individuals who called themselves “Mexican” in open response, eight switched to “Mexican American” when choosing from our list, while no one did the opposite, i.e., no one went from supplying the term “Mexican American” on their own to choosing the term “Mexican” from our list.

There were 29 people who used a non-list term in open response and then chose “Hispanic” first from the list. Of these, eight were people who had identified as “Spanish American” in the open response. As shown in Table 3, 11 people also switched from using “Spanish” in the open response to “Hispanic” on our list, indicating some fluidity between the broader term “Hispanic” and the more regionally meaningful use of “Spanish” and other terms including the word “Spanish” or closely-related ones.

Of the 21 people who used an off-list term when describing themselves and then chose “Other” on the fixed-choice question, almost all used the same or similar wording from their open-ended response to what they named under “Other” on the list, especially when incorporating their secondary terms from the open-ended question. There were two cases in which “Other” appeared to identify participants who did not fit well with our criteria for sampling; one participant self-identified as “Filipino American” and another who self-identified as “White” and had no certain Spanish-speaking ancestors. On the other hand, one participant who was born in Spain and had European Spanish parents used “Hispanic”. Although some individuals who identified as “Spanish American” in open response did choose “Spanish” or “Hispanic” on the fixed-choice question, as
described above, there were three individuals who used “Spanish American” who chose “Other” from our list and again identified as “Spanish American”. This was the only term that was used more than once under “Other” on the fixed-choice question.

Table 3.

*Contingency table for open-ended vs. list-based identity terms.*

<table>
<thead>
<tr>
<th>Open-Ended Identity Term</th>
<th>List-Based Identity Term</th>
<th>Chicano/a</th>
<th>Hispanic</th>
<th>Latino/a</th>
<th>Mexican</th>
<th>Mexican Am.</th>
<th>Nuevomexicano/a</th>
<th>Spanish</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicano/a</td>
<td>30</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>0</td>
<td>2</td>
<td></td>
<td>40</td>
</tr>
<tr>
<td>Hispanic</td>
<td>6</td>
<td>178</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>19</td>
<td>12</td>
<td>1</td>
<td>223</td>
<td></td>
</tr>
<tr>
<td>Latino/a</td>
<td>3</td>
<td>3</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Mexican</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Mexican American</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>19</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Nuevomexicano/a</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>2</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>29</td>
<td>0</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>11</td>
<td>29</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>39</td>
<td>20</td>
<td>21</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>53</strong></td>
<td><strong>223</strong></td>
<td><strong>16</strong></td>
<td><strong>16</strong></td>
<td><strong>38</strong></td>
<td><strong>75</strong></td>
<td><strong>62</strong></td>
<td><strong>24</strong></td>
<td><strong>507</strong></td>
<td></td>
</tr>
</tbody>
</table>

Discussion

The identity terms that we examined in this study are discussed here in declining order of use on the open-ended identity question.

**Hispanic**

In the US today, the use of the term “Hispanic” stems from a government-level desire to identify people with origins in Spanish-speaking countries on the US Census in the 1970s, a time when Spanish surnames and Spanish language use were becoming less
uniform among the people in this group (Gómez, 2007; Jones-Correa & Leal, 1996). Advocacy groups were also invested in creating a panethnic term since the determination of the group’s size could help to maximize federal funding and political influence. In addition, Spanish-language news media strongly promoted “Hispanic” as it was considered to have marketing value (Mora, 2014). However, the term has long faced heavy criticism as being one that fails to represent any ethnic or nationality group at all (Gimenez, 1989). That reality did not prevent its rise to dominance in the 1990s both from within and without, i.e., its adoption as an ethnic identity term by many in the US, and its use as an identifier in politics, the social sciences and health research in ways that have often conflated ethnicity and race.

Within New Mexico, use of the term “Hispanic” is complicated. At a broader level, NMS group themselves together under this umbrella in contexts in which they wish to emphasize membership in this largest American minority group, which has roots in Latin America, including Mexico. Locally, though, the term “Hispanic” varies in application, sometimes used in contrast with Mexican-ness and other times being left out entirely when people are emphasizing the distinction between “Spanish” and “Mexican.” In this regard, “Hispanic” may appear useful to NMS members who would choose to avoid the often emotionally wrought question of their historical heritage, i.e., whether “Spanish” or “Mexican” (Gonzales, 1997). More than three-quarters (82%) of our participants selected “Hispanic” in any order on our list-based identity question, indicating that most NMS identify as “Hispanic.” However, our data also show that another subset of NMS do not identify as “Hispanic”, which suggests that its wide use on forms and in research recruitment is problematically leaving out a substantial number of
New Mexicans who have Spanish-speaking ancestors but who do not identify with this term.

**Spanish**

Gonzales (1993) speaks to the rise of Spanish nomenclature in the early twentieth century in NM. At that time, “Spanish-American” was preferred because it simultaneously separated the long-resident Spanish-surnamed population of NM from recent Mexican immigrants, implied that the individual is of European rather than indigenous descent, and expressed American citizenship (Campa, 1946). It additionally separated them from non-Spanish European Americans. The high frequency of appearance of “Spanish” in open responses, second only to “Hispanic,” indicates the staying power of this heritage nomenclature. Perhaps because NM has been a state for more than a century, making the last of those three less compelling, or perhaps just as a shortened form, there has more recently been a shift to the shorter “Spanish.” There were individual shifts between different terms indicating “Spanish” identity going from open response to fixed choice responses, demonstrating some nomenclature flexibility.

The term “Spanish American” is still in use, however, and as detailed in the Results section, this was the only term that appeared more than once under “Other” on our fixed-choice question. The fact that these individuals did not instead choose “Spanish” underscores how important subtle variations in identity nomenclature can be. Many of its users emphasized that the term imparted specific and important meaning because it referred both to heritage from Spanish settlers and to American citizenship. In several cases, individuals who described themselves as “Spanish American” mentioned that they had served in the US military and that identifying themselves in a way that
included the term “American” was therefore very important to them. As one participant said, “I consider myself an American primarily. Having served four years in the American military in Korea, I think we all earned the right to call ourselves American.”

Among our participants, the term “Spanish” seemed to be linked less to an interest in having a pure lineage of Spanish ancestors and more of an interest in asserting a cultural difference between their family histories and those of recent Mexican immigrants. One participant who identified on our forced-choice question as “Spanish” first and “Nuevomexicano/a” second, with no other selections, said, “I don't like ‘Hispanic,’ consider it a government term. ‘Chicana’ was cool in the 70s, but I think of it as kind of a revolutionary term, not really sure what it means. ‘Latino’ is too vague, ‘Mexican’ is definitely from Mexico, ‘Mexican American’ is someone who was born here or immigrated here of Mexican parents.”

Those who talked about their families’ stories of being “Spanish” often expressed uncertainty:

“The family story is basically that we are from Spain, we are Spanish American, we are not Mexican American, not Anglo, no specifics to when so-and-so came over, but growing up, it was pretty clear that we weren’t just regular Hispanics or Latinos. Our roots reached back somehow to Spain, and it was strongly implied that we were not of a mixed race, even though we probably are.”

“My mother used to be very vocal that we were Spaniards, we were not Mexicans. We all came in through Mexico, if I had ever told her that she would have said that was a big, fat lie. The family story was that we came from Spain, but of course that’s not true, they went to Mexico first.”

“They talk about being Spanish, from Spain, they've been here lots of generations, she talks about her ancestors coming from Spain, not from Mexico, my sister and I laugh at this, us all thinking we’re from Spain.”
Many participants mentioned family stories of Native American ancestors. However, they were also often quite emphatic about having no ancestors from Mexico, i.e., they believed their ancestors to be comprised primarily of Spanish and other European immigrants to the New World with small contributions to their lineages from Native Americans who lived in what is now the US Southwest. Some participants described specific branches of their family trees that had Native American ancestors: “Had a great-grandmother on dad's side who was full Apache. As far as we know, we're just a mixture of Hispanic and Native American.” Others gave more general statements about the likelihood of having Native American ancestors: “I was named after my great-grandmother, but what did they do, who were they? I imagine that I am Spanish, with, based on what I've been told, my mother would roll over in her grave, but I would perceive that I would have some Native American, even though we were isolated. We came in this caravan up the Camino Real, but at some point they'd have had to find someone in those groups.”

Many of our participants emphasized their deep family roots in New Mexico, many with genealogical ties from colonial NM to Spain, but they were often less concerned with being purely Spanish/European in origin than they were with asserting their distinctness from those of Mexican origin. This result is consistent with other recent research on ethnic identity in New Mexico that found claims of Spanish ancestry to be founded not in claims of being racially “White,” but in separating themselves from those of recent Mexican origins and the accompanying social hostility that this association confers, as well as signaling strong identification with U. S. nationality (Salgado, 2018).
A generation or two ago, “Spanish” or “Spanish-American” served as a common identifier for NMS, but based on our interview responses, modern sensibilities preclude the use of this term for adults younger than about 50, due to recognition that it is neither accurate nor informative:

“I haven't really met many people from Spain. If they say they're Spanish, I'm like ‘No you're not!’ and I think they're Chicano. I just don't believe them.”

“Growing up I used ‘Spanish’, then I realized that ‘Spanish’ meant European Spanish descent. I was confused.”

“When people say they are Spanish, I don’t know how much Spanish they really are.”

Chicano/a

The terms “Chicano” and “Chicana” rose to prominence as a term of self-identification during the 1960s era of widespread social movement activity (Gómez-Quiñones & Vasquez, 2014). At the national level, the Chicano Movement took hold among Americans of Mexican descent who would have identified as “Mexican American” in the previous generation, and in New Mexico as “Spanish American” (Gonzales, 1993). The Chicano Movement was rooted in young people’s frustration with the persistence of oppression of people of color in the United States and determination to do something about it (García, 1989). Accordingly, our participants repeatedly declared that the use of the term “Chicano/a” indicated something to do with rebellion:

“I just don't like that term Chicano, it was used by radicals in the 60s. The Hispanics will just quietly take discrimination, the Chicanos will yell and get their voice heard.”

“Chicana was a way for us to reclaim our roots. My grandmother really thought it was terrible, she had bad associations of that word. I joined Chicano studies, but no one else in my family claimed it.”
“The one I never identified with was Chicano. I grew up during that time, but only the radicals called themselves that. I think it’s a political thing, not a difference in the people themselves.”

The word “radical” came up regularly in association with the term, often used in a disparaging manner. In the 1960s, Chicano identity arose among NMS and reflected a rejection of Spanish American identity because of its association with whiteness (Gonzales, 1993). In defining themselves as Spanish American, NMS had asserted that their own status in New Mexico was equal to or higher than that of non-Hispanic whites (commonly referred to as “Anglos” in New Mexico). The Chicano movement, on the other hand, emphasized recognizing and seeking change from discrimination that they had faced and continued to face (Nieto-Phillips, 2008). Chicano movement participants in New Mexico affiliated themselves with a group from whom their ancestors had historically separated themselves under the assumption that they shared in the notion of greater Chicano people in the Southwest (Gonzales, 1993, 2006). Our research indicates that the term “Chicano/a” went on to be only weakly incorporated in NMS and not able to overcome or supplant a nomenclature that incorporated the term “Spanish.”

**Mexican**

During much of the Spanish colonial period, Spanish lands in the Americas were collectively termed the Viceroyalty of New Spain. Only when it gained its independence as a republic in 1821 did the country become “Mexico,” a name the Spanish had originally borrowed from the Mexica people to apply to their capital, Mexico City. At the time, the land that is now New Mexico was Mexican land, so the residents of the land were “Mexican” by nationality and citizenship. New Mexico officially gained the status
of a US Territory in 1850, although the process began with the Treaty of Guadalupe Hidalgo in 1848. During this transition, Mexican citizens living in the region could choose to retain their Mexican citizenship or become US citizens. Over three-fourths of them opted for U.S. citizenship. Their rights as US citizens, however, were limited because of the nature of the territorial government system (Gómez, 2007), preventing them from achieving full democratic representation until New Mexico gained statehood in 1912. While many in NM who had been Mexican citizens were no longer “Mexican” by nationality beginning in 1850, the term had become a way of referring to the ethnic group comprised of the former Mexican citizens and their descendants. The identifier “Mexican” was used both internally among the people themselves (Marez, 2001), and externally by the government and English-language newspapers (Nieto-Phillips, 2008). However, the term was often carefully contextualized to clarify that the “Mexicans” of NM were distinct from “Mexicans” from Mexico (Campa, 1946). Racist politics during the fight for statehood led to a shift away from NMS identifying as “Mexican” and towards identification as “Spanish” and “Spanish American” in order to emphasize their purity of blood and perceived ability to self-govern (Nieto-Phillips, 2008).

In our study, the identity term “Mexican” was less likely than any of the other terms to be selected in the fixed-choice question by an individual who had not already used this term on the open-ended question. This means that people were unlikely to pick “Mexican” from the list unless they had already used this term on their own.

All listed terms showed an increase in frequency of overall use in the fixed-choice responses compared to the open-ended ones. This is primarily due to the fact that most participants described themselves using one or two terms, while most provided three
ranked choices from the fixed-choice list, so the overall number of responses increased more than twofold. However, “Mexican” showed the smallest increase, with only 25 additional participants who picked “Mexican” as an identity term.\footnote{The term “Mexican American” also had a relatively small increase in selection, with 67 additional participants choosing it in any order from the list. The other terms on the list each were chosen more than 100 extra times overall in the fixed-choice responses: “Chicano/a” (129), “Hispanic” (172), “Latino/a” (140), “Spanish” (165) and “Nuevomexicano/a” (199).} It is reasonable to surmise that the reluctance to use “Mexican” among NMS is because the term is associated with Mexican nationality. This is due to the public prominence of Mexican immigration, addressed below.

**Mexican American**

By the 1930s, a majority of people of Mexican descent in the United States were born and raised as American citizens. Simultaneously, the country was undergoing increased urbanization and industrialization, leading to new educational and economic opportunities. These factors came together to create strong leadership among those of Mexican descent as they pushed towards greater realization of their rights as American citizens, and a need for a group identity to provide greater unity in their struggle. This force brought forth the rise of “Mexican American” identity from the 1930s to the early 1960s, termed the “Mexican-American Generation” (García, 1989). While this nomenclature appeared in New Mexico among NMS professionals, it failed to take broad or deep hold in the state (Gonzales, 1993).

Of the 23 people who described themselves using the term “Mexican American” in this study, 19 also chose “Mexican American” as their first choice from the list. This was the highest adherence to any term on the list, i.e., participants who described themselves as “Mexican American” on their own were highly likely to choose the same
term as their first choice from our list. Interestingly, of the 23 individuals who described themselves as “Mexican” on the open-ended question, 14 selected “Mexican” again first from the list, while eight others chose “Mexican American” first.

While “Mexican” and “Mexican American” are distinct terms with separate histories, they are brought together in New Mexico in being used to indicate very recent origins from Mexico. Our participants often described “Mexican” as a term for people who had themselves been born and raised in Mexico and “Mexican American” as a term for someone who was born and/or raised in the United States but with parents from Mexico. Many NMS do not consider themselves to have any ancestors from Mexico even as their homeland was once part of the Mexican Republic. Accordingly, they do not often identify as “Mexican American” even though they are often described as such by researchers and policymakers (e.g., Vargas, 2016). Our participants often defined their identities and family histories in opposition to Mexican heritage, as many NMS have done for more than a century (Gonzales, 2006).

We have shown that identification as “Mexican” or “Mexican American” in NMS is associated with having more recent ancestors from Mexico compared to the other listed terms (Healy et al. 2018). Here we show that this history plays out in the strength of commitment participants had to ethnic nomenclature. While other terms in list were selected often when participants were given a list and allowed to select multiple options, “Mexican” and “Mexican American” were mostly named in open response if participants were going to use them at all, showing a stronger commitment to this identity than others among NMS. Given the history of discrimination and differences in opportunity for Mexican immigrants in the United States, this may correspond to an important distinction
in New Mexico, in which NMS with a multigenerational history in the state use ethnic identity terms more symbolically, similar to ethnic distinctions made by “White Americans” (Waters, 1990), while those whose families have come from Mexico recently have a strong identity of being Mexican that is rooted in practical, material interests (Nagel, 1994).

In a separate question later in our interview, we asked participants to identify their race using the options offered on the most recent US Census. The two most common choices overall made up 97% of responses and were “White” and “Some other race”, with 44% of participants choosing “White” and 53% choosing “Some other race.” The 2010 US Census, by comparison, had 53% of the Hispanic/Latino population identify as “White” and 36.7% identify as “Some other race”. Interestingly, there were differences in racial identification in our study that corresponded to the identity terms chosen by participants (results for fixed-choice identity question). Those who identified as “Spanish” most often identified as “White” (54.8%), and most other terms had between 40-50% of participants identified as “White”. However, “Mexican” and “Mexican American” participants were much less likely to identify as “White” and much more likely to identify as “Some other race.” For those who identified as “Mexican”, 12.5% identified as “White” and 81.3% identified as “Some other race”. For “Mexican American” participants, those numbers were 23.7% (“White”) and 76.3% (“Some other race.”) These results further support the interpretation of symbolic versus practical ethnic identities by indicating that many NMS do view themselves as “White Americans”, but that those identifying as “Mexican” or “Mexican American” usually do not.

Latino/a
The term “Latino/a” (and more recently, the gender-neutral “Latinx”) is the most recent one on our list to have been adopted, not appearing on the US Census until 2000. Its popularity rose in the context of arguments against the use of “Hispanic” as a panethnic term for those of Spanish-speaking descent, with many activists arguing that the term implied Spanish heritage and “Caucasian” race and that this was not applicable to many people with ancestors from Latin America (D. E. Hayes-Bautista & Chapa, 1987). Additionally, it created confusion for how to classify Spanish people from Spain in the United States and prevented those from Portuguese-speaking Latin America from being identified in the same group as others of Latin American origins. While “Latino/a” has risen in use at the national level over the past two decades (Morales & Bonilla, 1993), its popularity in the Southwest has been limited. In Texas, many people of Spanish-speaking descent use the term to describe others who were not born in the US and are not of Mexican origin, i.e., those from a Latin American country other than Mexico (Dowling, 2014).

In New Mexico, “Latino” and “Latina” are often used to describe an outside group that does not fit within the predominant “Spanish”/“Mexican” narrative:

“For me, when I think Latino, you think like Puerto Rico, Cuba, the East Coast like New York.”

“Latino is someone from Latin America.”

However, the statements that NMS participants made about their conceptualization of “Latino/a” did not map well onto its use in the questions. It was not a popular first-choice term either in the open-ended or fixed-choice identity question, and if Gonzales (1993) is correct, will not be until New Mexico experiences significant
immigration from Central and Latin America.

Still, when the term was used, the respondents often had long family histories in New Mexico. The few participants in the study who did come from Puerto Rico and Cuba tended to identify specifically with their nation of origin, as is consistent with research indicating that people in the United States from Latin American countries prefer national terms for self-identification (Gimenez, 1989; Jones-Correa & Leal, 1996). Of the twelve participants who themselves were born or had at least one parent born in a Latin American country other than Mexico, none used the term “Latino/a” when answering our open-ended question. However, eight of the 12 selected “Latino/a” as one their choices (one first choice, six second choices and one third choice), much higher than the overall use of “Latino/a” in the study. Eight of the 12 chose “Hispanic” as their first choice from the list, and seven named “Hispanic” first on the open-ended question. This could indicate that these individuals have adopted the local use of “Hispanic” while in residence in New Mexico.

**Nuevomexicano/a**

The term “Nuevomexicano/a” never had a popular identity usage in New Mexico history. Rather, it is a construction promoted by academic researchers who consider it appropriate for people of Spanish-speaking descent whose roots go back to the original founding of New Mexico (e.g., Holtby, 2012). This is reflected in the fact that it was only used about 1% of the time as a response to our open-ended identity question. However, when we offered our list, which included “Nuevomexicano/a”, it increased dramatically to almost 15% use as a first choice. For people with Spanish-speaking ancestors who have long family histories in New Mexico, “Nuevomexicano/a” may represent an
opportunity to express something important about their regionalized ethnic identity while avoiding terms fraught with misinformation.

**Conclusions**

In this study, we found significant overlap between responses to open and fixed-choice identification questions, with 60% of participants using the same term in both, and others choosing closely-related terms or giving secondary terms that matched between the two questions. This was substantially higher than the 27.7% consistency found in a previous study that compared responses when participants were asked to describe their own ethnicities to their selections from a list of ethnic categories (Pringle & Rothera, 1996). Also, in our dataset, the open-ended identity terms map onto variables such as genetic ancestry, age, birthplace and deep family history in New Mexico in the same ways that we previously demonstrated for the fixed-choice terms (Healy et al., 2018). For example, European ancestry is lowest in those who described themselves as “Mexican” and “Mexican American” and highest in those who used the term “Spanish” when answering our open-ended identity question. Also, those who used the term “Spanish” were older, on average, compared to participants who used other terms.

If our interview had only included an open-ended identity question, the conclusions of our research published in other articles would still have held. However, we would have missed important effects, including the surge in the use of “Nuevomexicano/a” on the fixed-choice question, especially for people who had described themselves as New Mexican or Hispanic. The group of people who did identify as “Nuevomexicano/a” tended to be younger than those who identified as being “Spanish”, but the two groups showed similar results in both genetic ancestry and family
history and showed contrasts in these characteristics with those who identified as “Mexican” or “Mexican American.” On the other hand, if we had only used a fixed-choice identity question, we would not have caught the importance of the term “Spanish American”, the most popular term named in open responses that was not on our list. We would also have missed out on the richness of information provided by comparing the two answers; for example, the adherence to the terms “Mexican” and “Mexican American” between the two questions may point to real and previously undetected variation in the importance of ethnic identity in NMS with differing histories.

For scholars evaluating the causes of variation in sociocultural and biomedical variables, using ethnic identity data obtained from fixed-choice classification questions is more straightforward. For this reason, fixed-choice questions are most often used to inquire about ethnic identity in research as well as in other forms of data collection, such as census-taking. However, for researchers to be able to use fixed-choice questions, we emphasize the importance of conducting background research to inform what distinctions and groupings are relevant at a regional level. Also, our findings indicate that open-ended ethnic identity questions can contribute to a better understanding of identity construction when used in combination with fixed-choice identity questions.

Our results are based on a population from a single state, and we performed pilot research to identify the ethnic identity terms currently in common usage and offered those terms on our fixed-choice identity question. One tradeoff to this approach is that this necessarily limits cross-region comparisons. Nomenclatures rise and fall in the context of various sociological forces and accompanying political motivations (Gonzales,
1993), and this necessitates ongoing research to ensure the continued relevance of ethnic identity terms.

**References**


CHAPTER 5: SUMMARY AND CONCLUSIONS

Summary of dissertation findings

The focus of this dissertation was on characterizing variation in ethnic identity and genetic ancestry in New Mexicans of Spanish-speaking descent and looking at the origins and implications of this variation. In a broader sense, the goal of this research is to further our current understanding of the structure of variation in admixed populations.

In Chapter 2, we obtained microsatellite DNA data from a sample of New Mexicans who identified as Hispanic or Latino (NM-HL). We found that racial and ethnic self-identification among NM-HL emphasizes a distinction between those with recent family roots in Mexico and those with family ties to New Mexico stretching back to the Spanish colonial period. Our results indicate that there are differences in continental ancestry between these two groups, contributing to genetic substructure in NM-HL. In addition to NM-HL, we looked at matched microsatellite data for 13 Central and South American Hispanic or Latino population samples and found that genetic substructure is common in admixed populations in the Americas. Based on the more detailed findings in New Mexico, we suggest that this genetic substructure may often be linked to regionally-important variation in racial and ethnic self-identification. Recognizing and accounting for such substructure is critical in research on the genetic basis of diseases and in determining what social factors cause variation in health outcomes and social inequality.

For Chapter 3, we looked at genomic ancestry in a larger sample of New Mexicans of Spanish-speaking descent (NMS). We also collected data on ethnic self-identification, age, birthplace, and family ties to specific geographic regions, namely
Spain, Mexico, and colonial New Mexico.

Consistent with Chapter 2, we found significant differences in genomic ancestry between NMS who self-identified with some ethnic terms: those who identified as “Spanish” and “Nuevomexicano/a” had higher average proportions of European ancestry, and correspondingly lower Native American and African ancestry than did those who identified as “Mexican American”. We were able to explore this substructure further with this dataset, and we found that NMS who identified as “Hispanic”, “Spanish”, and “Nuevomexicano/a” were likely to have parents and grandparents from New Mexico and to report family ties to colonial New Mexico, while those who identified as “Mexican” and “Mexican American” often had been born in Mexico themselves or had parents and/or grandparents who were, and were unlikely to express a family connection to colonial New Mexico.

In Chapter 4, we compared responses to open and fixed-choice ethnic identification questions in NMS. We found that participants used the same ethnic identity nomenclature for both responses a majority of the time, and that open-ended responses correlated with other variables from our dataset in the same way that we demonstrated in Chapter 3. At the same time, we did find some important differences in the overall frequency of use of different terms, namely “Nuevomexicano/a” and “Spanish American”. The former was used as a first-choice term only five times in open responses but jumped to second place with seventy-five first-place selections when provided on our list of terms. As shown in Chapter 3, NMS identifying as “Nuevomexicano/a” showed similar characteristics as those who identified as “Spanish”, but were younger, indicating that the term attracts a distinct group of people even if the nomenclature is not commonly
used. The latter, “Spanish American”, was not on our list of terms on the fixed-choice question but was used as a first-choice term 22 times in open responses. The term “Hispanic” was the most common response for both questions, used as the first-choice term by 223 participants each time, with high but not complete overlap (n=179) between those who called themselves “Hispanic” in open responses and those who selected “Hispanic” first on the fixed-choice question. The popularity of “Hispanic” in open response demonstrates that this panethnic term has been adopted in everyday use by NMS. This contrasts it “Latino/a”, the other main panethnic term for people with Spanish-speaking ancestors used in the United States today, but one that has not taken hold in NMS. Participants who identified as “Mexican” and “Mexican American” showed greater adherence to these terms than others, potentially demonstrating that these identities have more practical, material importance in the lives of the individuals than do other ethnic identities in NMS. Asking about ethnic identity using both forms of inquiry allowed us to extract greater information from the data than we could have with either question alone.

**Limitations**

For Chapter 2, the questions that were asked of participants about their ethno-racial identity were not optimized to suit NMS. Particularly, the category “Mexican-Chicano-New Mexico” grouped several terms together in a way that conflicts with the commonly-held idea in New Mexico that those of recent Mexican origins are distinct from those with deeper New Mexican roots. The word “Spanish” was used twice, both in “Other specified Spanish origin (including European)” and “Spanish, Hispanic, or Latino, not otherwise specified”, and the former was heavily favored by participants over the
latter, perhaps to stress their perceived “Spanish” origins relative to the panethnic categories “Hispanic” and “Latino”. We attempted to address the appropriateness of ethnic nomenclature categories in our later research.

Also in Chapter 2, we found higher estimates of European ancestry in our sample from New Mexico than other regional studies (Bonilla et al., 2004; Klimentidis, Miller, & Shriver, 2009). This could reflect differences in sampling strategy, e.g., we did not exclude participants who had one parent outside of the sample population demographic. The DNA markers used were autosomal microsatellites as opposed to the Single Nucleotide Polymorphisms (SNPs) sampled in the other studies, and this could also be a factor, as microsatellite loci are prone to allelic dropout that could lead to errors in estimates of allele frequency and tests of Hardy-Weinberg Equilibrium. It is possible that dropout could have affected our estimates of the inbreeding coefficient, FIS.

For Chapters 3 – 4, an important limitation of our research is that our interviews were conducted only in English as opposed to English and Spanish, so our sample is unlikely to be representative of ethnic identity in NMS overall. We also conducted our interviews in an academic setting and using a structured interview, and it is possible this could have caused study participants to be less likely to use terms that they associate with discrimination (Doan & Stephan, 2006). Our participants’ education levels were higher than the overall population, they had a lower rate of being born outside of the United States (see Table S1 in Appendix A), and all were likely to have been United States citizens or legal residents, although we did not inquire about resident status. These factors also make our findings less likely to perfectly characterize the overall NMS population. Our interviewers themselves could also have influenced the ethnic identity terms chosen
by our participants according to research findings suggesting that interviewer characteristics including age, gender, and ethnic identity can impact responses (Doan & Stephan, 2006).

When we recruited participants, we used the term “New Mexican of Spanish-speaking descent” to avoid using any of the ethnic identity terms that we hoped to study. However, because the word “Spanish” is included, it is possible that NMS identifying as “Spanish” were more likely to enroll in the study than those who did not, particularly those whose identities are defined in opposition to the “Spanish” group.

Our results are based on a population primarily from a metropolitan area in a single state, and we used pilot research to determine ethnic identity terms of current relevance in the region. A tradeoff to this regional approach is that it limits comparisons to other regions. Because nomenclatures are constantly shifting due to sociological forces and political motivations (Gonzales, 1993), ethnic identity terms used in research must be subject to ongoing research to confirm their continued relevance.

We used modern African, indigenous American, and European populations as the parent populations in our estimates of genomic ancestry. These populations are only proxies for the true historical populations who were the actual ancestors of NMS. Ancestry estimates also vary depending on the methods used to produce them, meaning that cross-study comparisons are not always possible. However, the inferences from our own work depend only on the relative differences in genomic ancestry proportions between participants and groups within the study.
Conclusions

We found evidence of ethnic identity-related genetic substructure in NMS. Ethnic identity nomenclature in NMS, and genetic substructure in turn, is patterned according to families’ recent and historical ties to specific regions, especially Mexico and colonial New Mexico. The patterns of genomic ancestry that we found clearly began to take form during the Spanish colonial period, but were also influenced by migrations to New Mexico from different locations throughout the United States and Mexico in more recent times. We also found evidence suggestive of ethnic-based genetic substructure in 12 of 13 other admixed populations located in Central and South America, underscoring the rarity of long-term random mating in admixed populations in the Americas. Ethnic identity-related genetic substructure can be expected in human populations due to the important role of long-range migrations and resulting structured admixture events throughout human history (Pickrell & Reich, 2014). It is critical that researchers and policymakers dealing with admixed populations account for this substructure in order to obtain valid results in studies of disease-causing genes and create policies that best address variation in health outcomes and social inequality.

In NMS, using responses from both open- and fixed-choice identity questions provided additional ways to examine participants’ ethnic identities. Together, the data from these questions allowed the identification of a potential distinction between symbolic and practical ethnic identities in different NMS subgroups. The results of this research demonstrate that fixed-choice questions about ethnic identity can be used more effectively when background research is used to characterize current, relevant ethnic identity nomenclature at a regional level. Also, the findings suggest that the use of open-
ended ethnic identity questions can benefit research on identity construction when used in combination with fixed-choice identity questions.

References


APPENDICES

Appendix A

Supplementary Tables for Chapter 3

<table>
<thead>
<tr>
<th>Table S1. Comparisons of selected demographic data from Bernalillo County and New Mexico&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Total Population</td>
</tr>
<tr>
<td>Median Age</td>
</tr>
<tr>
<td>Median Annual Household Income</td>
</tr>
<tr>
<td>High School Grads</td>
</tr>
<tr>
<td>Foreign Born</td>
</tr>
<tr>
<td>Speak English Only</td>
</tr>
<tr>
<td>Speak Spanish</td>
</tr>
<tr>
<td>Hispanic or Latino (of any race)</td>
</tr>
<tr>
<td>Mexican</td>
</tr>
<tr>
<td>Puerto Rican</td>
</tr>
<tr>
<td>Cuban</td>
</tr>
<tr>
<td>Other Hispanic or Latino</td>
</tr>
<tr>
<td>Not Hispanic or Latino</td>
</tr>
</tbody>
</table>

<sup>a</sup>estimates from the 2010 census or 2016 ACS

<sup>b</sup>speak Spanish better than or as well as English

<table>
<thead>
<tr>
<th>Table S2. Genomic Ancestry Proportions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>n</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Chicano</td>
</tr>
<tr>
<td>Hispanic</td>
</tr>
<tr>
<td>Latino</td>
</tr>
<tr>
<td>Mexican</td>
</tr>
<tr>
<td>MexicanAmerican</td>
</tr>
<tr>
<td>Nuevomexicano</td>
</tr>
<tr>
<td>Spanish</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Weighted average</td>
</tr>
</tbody>
</table>
Table S3. Genomic Ancestry. Wilcoxon Rank Sum Test P-values

<table>
<thead>
<tr>
<th>European</th>
<th>Chicano/a</th>
<th>Hispanic</th>
<th>Latino/a</th>
<th>Mexican</th>
<th>Mex Amer</th>
<th>Nuevomex</th>
<th>Spanish</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicano/a</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latino/a</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mexican</td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mexican American</td>
<td>**</td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nuevomexicano/a</td>
<td>*</td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>*</td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05; ** p < Holm-adjusted p-value 0.0039

<table>
<thead>
<tr>
<th>Native American</th>
<th>Chicano/a</th>
<th>Hispanic</th>
<th>Latino/a</th>
<th>Mexican</th>
<th>Mex Amer</th>
<th>Nuevomex</th>
<th>Spanish</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicano/a</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latino/a</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mexican</td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mexican American</td>
<td>**</td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nuevomexicano/a</td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05; ** p < Holm-adjusted p-value 0.0039

<table>
<thead>
<tr>
<th>African</th>
<th>Chicano/a</th>
<th>Hispanic</th>
<th>Latino/a</th>
<th>Mexican</th>
<th>Mex Amer</th>
<th>Nuevomex</th>
<th>Spanish</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicano/a</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latino/a</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mexican</td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mexican American</td>
<td>**</td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nuevomexicano/a</td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05; ** p < Holm-adjusted p-value 0.0036
<table>
<thead>
<tr>
<th>Term</th>
<th>US Southwest</th>
<th>New Mexico</th>
<th>SLV</th>
<th>Mexico</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicano/a</td>
<td>0.94</td>
<td>0.81</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.94</td>
<td>0.84</td>
<td>0.16</td>
<td>0.00</td>
</tr>
<tr>
<td>Latino/a</td>
<td>0.88</td>
<td>0.82</td>
<td>0.12</td>
<td>0.06</td>
</tr>
<tr>
<td>Mexican</td>
<td>0.50</td>
<td>0.25</td>
<td>0.00</td>
<td>0.50</td>
</tr>
<tr>
<td>Mexican American</td>
<td>0.82</td>
<td>0.61</td>
<td>0.05</td>
<td>0.16</td>
</tr>
<tr>
<td>Nuevomexicano/a</td>
<td>0.97</td>
<td>0.91</td>
<td>0.28</td>
<td>0.01</td>
</tr>
<tr>
<td>Spanish</td>
<td>0.97</td>
<td>0.89</td>
<td>0.31</td>
<td>0.00</td>
</tr>
<tr>
<td>Other</td>
<td>0.89</td>
<td>0.78</td>
<td>0.30</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term</th>
<th>US Southwest</th>
<th>New Mexico</th>
<th>SLV</th>
<th>Mexico</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicano/a</td>
<td>0.92</td>
<td>0.79</td>
<td>0.05</td>
<td>0.04</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.90</td>
<td>0.80</td>
<td>0.27</td>
<td>0.03</td>
</tr>
<tr>
<td>Latino/a</td>
<td>0.88</td>
<td>0.76</td>
<td>0.24</td>
<td>0.06</td>
</tr>
<tr>
<td>Mexican</td>
<td>0.28</td>
<td>0.19</td>
<td>0.03</td>
<td>0.69</td>
</tr>
<tr>
<td>Mexican American</td>
<td>0.49</td>
<td>0.29</td>
<td>0.07</td>
<td>0.50</td>
</tr>
<tr>
<td>Nuevomexicano/a</td>
<td>0.95</td>
<td>0.87</td>
<td>0.39</td>
<td>0.01</td>
</tr>
<tr>
<td>Spanish</td>
<td>0.92</td>
<td>0.87</td>
<td>0.38</td>
<td>0.00</td>
</tr>
<tr>
<td>Other</td>
<td>0.85</td>
<td>0.76</td>
<td>0.37</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Term</th>
<th>US Southwest</th>
<th>New Mexico</th>
<th>SLV</th>
<th>Mexico</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicano/a</td>
<td>0.79</td>
<td>0.67</td>
<td>0.11</td>
<td>0.12</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.80</td>
<td>0.72</td>
<td>0.29</td>
<td>0.06</td>
</tr>
<tr>
<td>Latino/a</td>
<td>0.84</td>
<td>0.71</td>
<td>0.26</td>
<td>0.06</td>
</tr>
<tr>
<td>Mexican</td>
<td>0.20</td>
<td>0.13</td>
<td>0.03</td>
<td>0.73</td>
</tr>
<tr>
<td>Mexican American</td>
<td>0.34</td>
<td>0.22</td>
<td>0.06</td>
<td>0.61</td>
</tr>
<tr>
<td>Nuevomexicano/a</td>
<td>0.88</td>
<td>0.81</td>
<td>0.37</td>
<td>0.03</td>
</tr>
<tr>
<td>Spanish</td>
<td>0.91</td>
<td>0.84</td>
<td>0.41</td>
<td>0.00</td>
</tr>
<tr>
<td>Other</td>
<td>0.78</td>
<td>0.64</td>
<td>0.39</td>
<td>0.04</td>
</tr>
</tbody>
</table>
Appendix B

Questionnaire used for main phase of data collection (used in Chapters 3 – 4)

Questionnaire

Date and time: Month___ Day___ Year_______ Time_______

Location: _________________________

Interviewer: _________________________

Section I: Sociocultural Information

(Show NM map with counties to get county for questions #1-7. If answer is Bernalillo County, specify area using Albuquerque map. )

1. Where were you born?
   a. city: ____________
   b. county: ____________
   c. state: ____________
   d. country: ____________

2. Where did you grow up?
   a. city: ____________
   b. county: ____________
   c. state: ____________
   d. country: ____________

3. Where do you live now?
   a. city: ____________
   b. county: ____________
   c. state: ____________
   d. country: ____________
   e. zip code: ____________

4. Where was your mother born?
   a. city: ____________
   b. county: ____________
   c. state: ____________
   d. country: ____________

5. Where was your father born?
   a. city: ____________
   b. county: ____________
c. state: __________
d. country: __________

6. Where were your mother’s parents born?
   a. Grandmother:
      1. city: _________
      2. county: __________
      3. state: __________
      4. country: __________
   b. Grandfather:
      1. city: _________
      2. county: __________
      3. state: __________
      4. country: __________

7. Where were your father’s parents born?
   a. Grandmother:
      1. city: _________
      2. county: __________
      3. state: __________
      4. country: __________
   b. Grandfather:
      1. city: _________
      2. county: __________
      3. state: __________
      4. country: __________

8. How old are you? ______

9. What is your sex? ______

10. Does your father belong to an old New Mexico family (land grants)?

11. Does your mother belong to an old New Mexico family (land grants)?

12. Were any of your ancestors colonists from Spain? If so, who and when?

13. Do you have any ancestors from Mexico? If so, who and when?
14. Do you have ancestors from anywhere else? If so, where else, who and when?

(Interviewer: describe observations of how NMS divide themselves into at least two distinct groups to contextualize the next set of questions.)

15. What are the groups (within NMS)?

16. Which group do you belong to?

17. Which group do you think your mother belongs to?

18. Which group do you think your father belongs to?

19. Are you married?  □ yes  □ no

(If yes) Which group do you think your spouse belongs to?

20. Which of the other groups you identified is most similar to yours?

21. Which of the other groups you identified is most different from yours?

22. Please describe how the members of _____________________________(insert name of group identified as most different from subject’s; answer to #21) tend to differ from members of _____________________________(insert name of subject’s self-identified group; answer to #16) in these features:

   a. Skin color
      1. □ same
      2. □ different. How? ________________________________
      3. □ don’t know
      4. □ refuse

   b. Hair
      1. □ same
      2. □ different. How? ________________________________
      3. □ don’t know
      4. □ refuse

   c. Face
      1. □ same
2. ☐ different. How? ____________________________
3. ☐ don’t know
4. ☐ refuse

d. Other physical difference
   1. What feature? ________ Different how? ________________
   2. What feature? ________ Different how? ________________
   3. ☐ don’t know
   4. ☐ refuse

e. Amount of Spanish use
   1. ☐ same
   2. ☐ different. How? ____________________________
   3. ☐ don’t know
   4. ☐ refuse

f. Accent when speaking English
   1. ☐ same
   2. ☐ different. How? ____________________________
   3. ☐ don’t know
   4. ☐ refuse

g. Accent when speaking Spanish
   1. ☐ same
   2. ☐ different. How? ____________________________
   3. ☐ don’t know
   4. ☐ refuse

h. Other language difference
   1. What feature? ________ Different how? ________________
   2. What feature? ________ Different how? ________________
   3. ☐ don’t know
   4. ☐ refuse

i. Food (what people eat)
   1. ☐ same
   2. ☐ different. How? ____________________________
   3. ☐ don’t know
   4. ☐ refuse

j. Clothing (what people wear)
   1. ☐ same
   2. ☐ different. How? ____________________________
3. [ ] don’t know
4. [ ] refuse

k. Make-up

1. [ ] same
2. [ ] different. How? __________________________
3. [ ] don’t know
4. [ ] refuse

l. Other cultural difference

1. What feature? __________ Different how? __________
2. What feature? __________ Different how? __________
3. [ ] don’t know
4. [ ] refuse

23. Between the members of the group you belong to and those of the group most different from yours, does either one:

a. Experience more discrimination?
   1. [ ] yes, my group
   2. [ ] yes, other group
   3. [ ] neither
   4. [ ] don’t know
   5. [ ] refuse

b. Have more wealth?
   1. [ ] yes, my group
   2. [ ] yes, other group
   3. [ ] neither
   4. [ ] don’t know
   5. [ ] refuse

c. Have more political influence?
   1. [ ] yes, my group
   2. [ ] yes, other group
   3. [ ] neither
   4. [ ] don’t know
   5. [ ] refuse

d. Have more education?
   1. [ ] yes, my group
   2. [ ] yes, other group
3. [ ] neither
4. [ ] don’t know
5. [ ] refuse

(“The last set of questions has asked you to define what groups exist in NMS and tell us about the differences between them. Because we know that participants will give us different groups, we now will ask you the same questions using a set of group names that we’ve gathered from other participants so that we can compare everyone’s opinions about these terms. The groups are listed in alphabetical order.” For questions 23-26, ask for first choice and mark as #1. Then ask if there is a second-best choice and mark as #2. Then ask if any of the other terms also describe the person of interest, mark as #3, etc.)

24. With which of these groups do you identify?
   a. [ ] Chicano/a
   b. [ ] Hispanic
   c. [ ] Latino/a
   d. [ ] Mexican
   e. [ ] Mexican American
   f. [ ] Nuevomexicano/a
   g. [ ] Spanish
   h. [ ] Other ____________________

25. Which of these groups do you think other New Mexicans of Spanish-speaking descent would think that you belong to?
   a. [ ] Chicano/a
   b. [ ] Hispanic
   c. [ ] Latino/a
   d. [ ] Mexican
   e. [ ] Mexican American
   f. [ ] Nuevomexicano/a
   g. [ ] Spanish
   h. [ ] Other ____________________

26. Which of these groups do you think your mother would identify with?
   a. [ ] Chicano/a
   b. [ ] Hispanic
   c. [ ] Latino/a
   d. [ ] Mexican
   e. [ ] Mexican American
   f. [ ] Nuevomexicano/a
   g. [ ] Spanish
   h. [ ] Other ____________________
27. Which of these groups do you think your father would identify with?
   a. [ ] Chicano/a
   b. [ ] Hispanic
   c. [ ] Latino/a
   d. [ ] Mexican
   e. [ ] Mexican American
   f. [ ] Nuevomexicano/a
   g. [ ] Spanish
   h. [ ] Other ________________

28. (if yes to #19) Which of these groups do you think your spouse would identify with?
   a. [ ] Chicano/a
   b. [ ] Hispanic
   c. [ ] Latino/a
   d. [ ] Mexican
   e. [ ] Mexican American
   f. [ ] Nuevomexicano/a
   g. [ ] Spanish
   h. [ ] Other ________________

29. Which of the other groups on the list is most similar to your first choice?
   a. [ ] Chicano/a
   b. [ ] Hispanic
   c. [ ] Latino/a
   d. [ ] Mexican
   e. [ ] Mexican American
   f. [ ] Nuevomexicano/a
   g. [ ] Spanish
   h. [ ] Other ________________

30. Which of the other groups you identified is most different from your first choice?
   a. [ ] Chicano/a
   b. [ ] Hispanic
   c. [ ] Latino/a
   d. [ ] Mexican
   e. [ ] Mexican American
   f. [ ] Nuevomexicano/a
   g. [ ] Spanish
h. ☐ Other ______________________

31. Please describe how the members of _____________ (insert name of group from the list named as most different from subject’s; answer to #30) tend to differ from _____________ (insert name of group from the list that subject most identified with; answer to #24) in these features:

a. Skin color
   1. ☐ same
   2. ☐ different. How? ________________________________
   3. ☐ don’t know
   4. ☐ refuse

b. Hair
   1. ☐ same
   2. ☐ different. How? ________________________________
   3. ☐ don’t know
   4. ☐ refuse

c. Face
   1. ☐ same
   2. ☐ different. How? ________________________________
   3. ☐ don’t know
   4. ☐ refuse

d. Other physical difference
   1. What feature? ________ Different how? ________________
   2. What feature? ________ Different how? ________________
   3. ☐ don’t know
   4. ☐ refuse

e. Amount of Spanish use
   1. ☐ same
   2. ☐ different. How? ________________________________
   3. ☐ don’t know
   4. ☐ refuse

f. Accent when speaking English
   1. ☐ same
   2. ☐ different. How? ________________________________
   3. ☐ don’t know
   4. ☐ refuse
g. Accent when speaking Spanish
   1. ☐ same
   2. ☐ different. How? __________________________
   3. ☐ don’t know
   4. ☐ refuse

h. Other language difference
   1. What feature? ________ Different how? _____________
   2. What feature? ________ Different how? _____________
   3. ☐ don’t know
   4. ☐ refuse

i. Food
   1. ☐ same
   2. ☐ different. How? __________________________
   3. ☐ don’t know
   4. ☐ refuse

j. Clothing
   1. ☐ same
   2. ☐ different. How? __________________________
   3. ☐ don’t know
   4. ☐ refuse

k. Make-up
   1. ☐ same
   2. ☐ different. How? __________________________
   3. ☐ don’t know
   4. ☐ refuse

l. Other cultural difference
   1. What feature? ________ Different how? _____________
   2. What feature? ________ Different how? _____________
   3. ☐ don’t know
   4. ☐ refuse

32. Between the members of the group you chose from the list and those of the most different group from yours on the list, does either one:

   a. Experience more discrimination?
      1. ☐ yes, my group
      2. ☐ yes, other group
3. □ neither
4. □ don’t know
5. □ refuse

b. Have more wealth?
   1. □ yes, my group
   2. □ yes, other group
   3. □ neither
   4. □ don’t know
   5. □ refuse

c. Have more political influence?
   1. □ yes, my group
   2. □ yes, other group
   3. □ neither
   4. □ don’t know
   5. □ refuse

d. Have more education?
   1. □ yes, my group
   2. □ yes, other group
   3. □ neither
   4. □ don’t know
   5. □ refuse

33. What religion did your parents raise you?

34. What is your current religion?

35. Have you
   a. *(if female)* Had a quinceñera?
      1. □ Yes
      2. □ No
      3. □ Don't know
      4. □ No response
   b. Participated in a pilgrimage to Chimayó?
      1. □ Yes
      2. □ No
      3. □ Don't know
      4. □ No response
   c. Attended or taken part in the dance of the Matachines?
1. ❑ Yes
2. ❑ No
3. ❑ Don't know
4. ❑ No response
d. Participated in Las Posadas festivities?
   1. ❑ Yes
   2. ❑ No
   3. ❑ Don't know
   4. ❑ No response

36. Do you have any Jewish ancestors?
    a. ❑ Yes
    b. ❑ No
    c. ❑ Don't know
    d. ❑ No response

37. On a scale of 0% to 100%, with 0% being no European ancestry and 100% being pure European, what percentage of European ancestry do you think you have?

38. Of your European ancestry, what percentage of Spanish ancestry do you think you have, with 0% being no Spanish ancestry and 100% being pure Spanish ancestry?

39. On a scale of 0% to 100%, with 0% being no Native American/indigenous ancestry and 100% being pure Native American/indigenous, what percentage of Native American/indigenous ancestry do you think you have?

40. (if 37+39 doesn't add to 100%) What do you think makes up the rest of your ancestry?

41. We’ve asked about your ancestors, and now we’d also like to ask about your appearance. For these features, please tell us how you appear on a scale from completely European to completely Native Americans/indigenous:
    a. Skin color

    European | | | | | | | | | | | Native American
b. Hair

<table>
<thead>
<tr>
<th>European</th>
<th>Native American</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>European</th>
<th>Native American</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

c. Face

d. Other physical feature: _____________________

<table>
<thead>
<tr>
<th>European</th>
<th>Native American</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

42. The 2000 U.S. Census used the following categories for race. Which of these apply to you?
   a. □ American Indian or Alaska Native
   b. □ Asian
   c. □ White
   d. □ Black or African American
   e. □ Native Hawaiian or other Pacific Islander
   f. □ Some other race _______________________

43. Which of these racial categories do you think most New Mexicans would use to describe you?
   a. □ American Indian or Alaska Native
   b. □ Asian
   c. □ White
   d. □ Black or African American
   e. □ Native Hawaiian or other Pacific Islander
   f. □ Some other race _______________________

44. When you were growing up, were you expected to finish high school?

45. When you were growing up, were you expected to go to college?

46. What is your highest completed level of education?
   a. □ Some high school or less
   b. □ High school graduate/GED
c. □ Some college or technical/vocational school/Associates degree

d. □ College graduate (4-year college degree)

e. □ Some postgraduate courses/Advanced or Professional degree

f. □ unknown

47. Are you still in school? ________

48. What is your mother’s highest completed level of education?
   a. □ Some high school or less
   b. □ High school graduate/GED
   c. □ Some college or technical/vocational school/Associates degree
   d. □ College graduate (4-year college degree)
   e. □ Some postgraduate courses/Advanced or Professional degree
   f. □ unknown

49. What is your father’s highest completed level of education?
   a. □ Some high school or less
   b. □ High school graduate/GED
   c. □ Some college or technical/vocational school/Associates degree
   d. □ College graduate (4-year college degree)
   e. □ Some postgraduate courses/Advanced or Professional degree
   f. □ unknown

50. What are your mother’s parents’ highest completed levels of education?
   a. Grandmother:
      1. □ Some high school or less
      2. □ High school graduate/GED
      3. □ Some college or technical/vocational school/Associates degree
      4. □ College graduate (4-year college degree)
      5. □ Some postgraduate courses/Advanced or Professional degree
      6. □ unknown

   b. Grandfather:
      1. □ Some high school or less
      2. □ High school graduate/GED
      3. □ Some college or technical/vocational school/Associates degree
      4. □ College graduate (4-year college degree)
      5. □ Some postgraduate courses/Advanced or Professional degree
      6. □ unknown

51. What are your father’s parents’ highest completed levels of education?
a. Grandmother:
   1.  □ Some high school or less
   2.  □ High school graduate/GED
   3.  □ Some college or technical/vocational school/Associates degree
   4.  □ College graduate (4-year college degree)
   5.  □ Some postgraduate courses/Advanced or Professional degree
   6.  □ unknown
b. Grandfather:
   1.  □ Some high school or less
   2.  □ High school graduate/GED
   3.  □ Some college or technical/vocational school/Associates degree
   4.  □ College graduate (4-year college degree)
   5.  □ Some postgraduate courses/Advanced or Professional degree
   6.  □ unknown

52. What is your current occupation?

(If student) What occupation do you expect to have after finishing school?

53. What is/was your mother’s occupation?

54. What is/was your father’s occupation?

55. What were/are your mother’s parents’ occupations?
   a. Grandmother ___________________
   b. Grandfather ___________________

56. What were/are your father’s parents’ occupations?
   a. Grandmother ___________________
   b. Grandfather ___________________

57. Think of this ladder as representing where people stand in New Mexico. At the top of the ladder are the people who are the best off—those who have the most money, the most education and the most-respected jobs. At the bottom are the people who are the worst off—who have the least money, least education and the least-respected jobs or no job. The higher up you are on this ladder, the closer you are to the people at the very top; the lower you are, the closer you are to the people at the very bottom.
Where would you place yourself on this ladder?

“Please place a large “X” on the rung where you think you stand at this time in your life, relative to other people in New Mexico.”

Rung # from bottom: ______

58. Choose all of the following that describe your current daily activities and/or responsibilities:
   a. ☐ Working full time
   b. ☐ Working part-time
   c. ☐ Full-time student
   d. ☐ Unemployed or laid off
   e. ☐ Looking for work
   f. ☐ Keeping house or raising children full-time
   g. ☐ Retired

59. How much did you earn, before taxes and deductions, during the past 12 months?
   a. ☐ Less than $5,000
   b. ☐ $5,000 through $11,999
   c. ☐ $12,000 through $15,999
   d. ☐ $16,000 through $24,999
   e. ☐ $25,000 through $34,999
60. Does one or both of your parents still claim you as a dependent on their taxes?
   a. ☐ Yes
   b. ☐ No
   c. ☐ Don't know
   d. ☐ No response

For questions 61-68, subject should respond for family household that claims subject if response was “Yes” to question 60. If response was “No”, subject should respond for current household.

61. How many people are currently living in your household, including yourself?
Your parents’ number of family members in 2009-2010. Include in your parents’ household: (1) your parents and yourself, even if you don’t live with your parents, (2) your parents’ other children if your parents will provide more than half of their support between July 1, 2009, and June 30, 2010, or and (3) other people only if they live with your parents, your parents provide more than half of their support and your parents will continue to provide more than half of their support between July 1, 2009, and June 30, 2010.
   a. _____Number of people
   b. _____Of these people, how many are children?
   c. _____Of these people, how many are adults?
   d. _____Of the adults, how many bring income into the household?

62. Which best describes the building in which you/your family lives? (Include all apartments, flats, etc., even if vacant.)
   a. ☐ A mobile home
   b. ☐ A house detached from any other house
   c. ☐ A house attached to one or more houses
   d. ☐ A building with 2 apartments
   e. ☐ A building with 3 or 4 apartments
   f. ☐ A building with 5 or more apartments
   g. ☐ Boat, RV, van, etc.

63. Is your/your family’s residence:
   a. ☐ Owned or being bought by you (or someone in the household)?
   b. ☐ Rented for money?
c. ☐ Other (specify) __________________________________________

64. Do you or your family own land?
   a. ☐ Yes
   b. ☐ No
   c. ☐ Don't know
   d. ☐ No response

65. (If household size >1) Which of these categories best describes your total combined family income for the past 12 months? This should include income (before taxes) from all sources, wages, rent from properties, social security, disability and/or veteran's benefits, unemployment benefits, workman's compensation, help from relatives (including child payments and alimony), etc.
   a. ☐ Less than $5,000
   b. ☐ $5,000 through $11,999
   c. ☐ $12,000 through $15,999
   d. ☐ $16,000 through $24,999
   e. ☐ $25,000 through $34,999
   f. ☐ $35,000 through $49,999
   g. ☐ $50,000 through $74,999
   h. ☐ $75,000 through $99,999
   i. ☐ $100,000 and greater
   j. ☐ Don't know
   k. ☐ No response

66. Beyond what your employer provides, do you have any financial investments?
   a. ☐ Yes
   b. ☐ No
   c. ☐ Don't know
   d. ☐ No response

67. Do you have at least one car?
   a. ☐ Yes (Make: __________ Model: ____________ Year: ______ If subject has more than one car, ask to describe primary car he/she drives)
   b. ☐ No
   c. ☐ Don't know
   d. ☐ No response

68. Do you own a computer?
   a. ☐ Yes (#PC desktops: ___ #Mac desktops: ___ #PC laptops: ___ #Mac laptops: ___)
   b. ☐ No
69. Choose one:
   a. ☐ I speak Spanish better than I do English
   b. ☐ I speak Spanish and English equally well
   c. ☐ I speak English better than I do Spanish
   d. ☐ I do not speak Spanish

70. Choose one:
   a. ☐ My mother does not speak English
   b. ☐ My mother speaks Spanish better than English
   c. ☐ My mother speaks Spanish and English equally well
   d. ☐ My mother speaks English better than Spanish
   e. ☐ My mother does not speak Spanish

71. Choose one:
   a. ☐ My father does not speak English
   b. ☐ My father speaks Spanish better than English
   c. ☐ My father speaks Spanish and English equally well
   d. ☐ My father speaks English better than Spanish
   e. ☐ My father does not speak Spanish

72. What language did you speak in your household growing up?

73. How many full siblings do you have?

74. At what store/s do you do the most of your grocery shopping?

75. How often do you listen to English-speaking radio stations?
   a. ☐ Never
   b. ☐ Occasionally
   c. ☐ Sometimes
   d. ☐ Often

76. How often do you listen to Spanish-speaking radio stations?
   a. ☐ Never
   b. ☐ Occasionally
   c. ☐ Sometimes
   d. ☐ Often
77. What radio station/s do you listen to most?

78. How often do you watch television and movies in English?
   a. □ Never
   b. □ Occasionally
   c. □ Sometimes
   d. □ Often

79. How often do you watch television and movies in Spanish?
   a. □ Never
   b. □ Occasionally
   c. □ Sometimes
   d. □ Often

80. What television station/s do you watch most?

81. What sport do you and/or your family members enjoy watching most?

82. Which are you more likely to eat with a meal made at home?
   a. □ Rice
   b. □ Potatoes
   c. □ Both equally
   d. □ Neither

83. Which are you more likely to eat with a meal made at home?
   a. □ Corn tortillas
   b. □ Flour tortillas
   c. □ Both equally
   d. □ Neither

84. Cancer
   a. You: □ Yes* □ No □ Don’t know  *What type? __________________
   b. Mother:  □ Yes* □ No □ Don’t know  *What type? __________________

Section II. Medical History

For questions #84-90, I am going to ask about your family history of several diseases. Please tell me whether you, your parents, or close relatives have/had any of these:
c. Father: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

d. Other relatives:

Who:______________________ What type? __________________

85. Diabetes

a. You: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

b. Mother: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

c. Father: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

d. Other relatives:

Who:______________________ What type? __________________

86. Hypertension (high blood pressure)

a. You: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

b. Mother: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

c. Father: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

d. Other relatives:

Who:______________________ What type? __________________

87. Heart Attack

a. You: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

b. Mother: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

c. Father: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

Who:______________________ What type? __________________

88. Gall Bladder disease

a. You: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

b. Mother: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

c. Father: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

Who:______________________ What type? __________________

89. Oculopharygeal Muscular dystrophy (OPMD)

a. You: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

b. Mother: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

c. Father: ☐ Yes* ☐ No ☐ Don’t know *What type? __________________

Who:______________________ What type? __________________

90. Cavernous angioma/cerebral cavernous malformation/CCM
a. You: □ Yes* □ No □ Don’t know *What type? __________________

b. Mother: □ Yes* □ No □ Don’t know *What type? __________________

c. Father: □ Yes* □ No □ Don’t know *What type? __________________

d. Other relatives:
Who:________________ Who:__________
What type? __________________

91. Do you smoke cigarettes? □ Yes □ No

92. Do you use any other form of tobacco (pipe/cigars/chewing)? □ Yes □ No

93. (If yes to either 91 or 92) What is your **best estimate** of the number of days you smoked part or all of a cigarette or used another tobacco product during the past 30 days?
   a. □ 1 or 2 days
   b. □ 3 to 5 days
   c. □ 6 to 9 days
   d. □ 10 to 19 days
   e. □ 20 to 29 days
   f. □ All 30 days

94. (If yes to 91) On the days you smoked cigarettes during the past 30 days, how many cigarettes did you smoke per day, on average?
   a. □ Less than one cigarette per day/1 cigarette per day
   b. □ Less than half a pack a day (2 to 5 cigarettes per day)
   c. □ 6 to 15 cigarettes per day (about ½ pack)
   d. □ 16 to 25 cigarettes per day (about 1 pack)
   e. □ More than a pack a day

95. (If yes to 92) On the days that you used other forms of tobacco, how much?

96. (If yes to 91 or 92) How old were you when you first started using tobacco?
AGE: ______

97. Did you used to smoke cigarettes? □ Yes □ No

98. Did you used to use any other form of tobacco? □ Yes □ No

99. (If yes to 97 or 98) How many days per month did you smoke cigarettes or use tobacco?
   a. □ 1 or 2 days
   b. □ 3 to 5 days
100. (If yes to 97) On the days that you smoked cigarettes, how many did you smoke per day, on average?
   a. Less than one cigarette per day/ 1 cigarette per day
   b. Less than half a pack a day (2 to 5 cigarettes per day)
   c. 6 to 15 cigarettes per day (about ½ pack)
   d. 16 to 25 cigarettes per day (about 1 pack)
   e. More than a pack a day

101. (If yes to 98) On the days that you used other forms of tobacco, how much?

102. (If yes to 97 or 98) For how many years did you smoke or use tobacco? ___

103. Do you drink alcohol? ______

104. (If yes to 103) On average, on how many days per week do you drink alcohol? ___

105. (If yes to 103) On days that you do drink, about how many drinks do you have, on average? ___

106. How often do you see a doctor?

107. For what reason/s would your parents have taken you to the doctor growing up?
   a. Regular checkup/exam
   b. Feeling sick
   c. Emergency (injury or severe illness)
   d. Treatment for condition doctor discovered earlier
   e. Other

108. Do you have health insurance? ______
   (If student or under 25) Are you on your parents’ health insurance plan? ______

109. Which of the following best describes your current health status?
   a. excellent
   b. good
   c. fair
   d. poor

110. Which of the following best describes your mother’s current health status?
   a. excellent
   b. good
111. Which of the following best describes your father’s current health status?
   a. ☐ excellent 
   b. ☐ good 
   c. ☐ fair 
   d. ☐ poor 
   e. ☐ N/A (not living or unknown)

112. How often do you go to the dentist?
   a. ☐ At least once a year 
   b. ☐ Every 2 years 
   c. ☐ Less often than every 2 years 
   d. ☐ Whenever needed - no regular schedule 
   e. ☐ Other 

113. What was the main reason for your last visit for dental care?
   a. ☐ Went in for checkup/exam/cleaning 
   b. ☐ Something wrong/hurting/bothering 
   c. ☐ Treatment for condition dentist discovered earlier 
   d. ☐ Check/adjust appliance/orthodontia 
   e. ☐ Other 

114. What would you do if you had dental pain? ________________________ 

115. How would you describe the condition of your teeth and gums? Would you say . .
   a. ☐ excellent 
   b. ☐ good 
   c. ☐ fair 
   d. ☐ poor 

These questions are about how you feel and how things have been with you. For each question, please give the one answer that comes closest to the way you have been feeling:

   0. never 
   1. occasionally 
   2. sometimes 
   3. often 

116. In the last month, how often have you felt that you were unable to control the important things in your life?
   ☐0 ☐1 ☐2 ☐3
117. In the last month, how often have you felt nervous and “stressed”?

❏ 0  ❏ 1  ❏ 2  ❏ 3

118. In the last month, how often have you felt that you were effectively coping with irritating life hassles?

❏ 0  ❏ 1  ❏ 2  ❏ 3

119. In the last month, how often have you been angered because of things that happened that were outside of your control?

❏ 0  ❏ 1  ❏ 2  ❏ 3

120. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?

❏ 0  ❏ 1  ❏ 2  ❏ 3  ❏ 4

The next set of questions is about how you are treated by other people. We’re coming back to the groups of NMS that we talked about earlier, and asking about discrimination you may have experienced as a member of your group. The answers are the same as above; please give the one that comes closest to how often you receive the type of treatment described:

121. As a ______________ (insert self-identified category from #16), how often are you treated with less courtesy than other people?

❏ 0  ❏ 1  ❏ 2  ❏ 3

122. As a ______________ (insert self-identified category from #16), how often do you receive poorer service than other people in restaurants or stores?

❏ 0  ❏ 1  ❏ 2  ❏ 3

123. As a ______________ (insert self-identified category from #16), how often do people treat you as if they are better than you?

❏ 0  ❏ 1  ❏ 2  ❏ 3

124. As a ______________ (insert self-identified category from #16), how often do people act as if they are smarter than you?

❏ 0  ❏ 1  ❏ 2  ❏ 3

125. As a ______________ (insert self-identified category from #16), how often do you think that discrimination makes it more difficult for you to accomplish your life goals?

❏ 0  ❏ 1  ❏ 2  ❏ 3
126. How often do you think that discrimination makes it more difficult for other ______________ (insert self-identified category from #16) to accomplish their life goals?

   □ 0 □ 1 □ 2 □ 3

**Section III. Photograph responses**

*I am now going to show you some photographs of other participants in this study and ask you two questions about the person in each photograph.*

127. Of the ethnicity terms you listed at the outset (remind them), which term would you use to describe this person?

128. Which of these terms would you use?
   a. □ Chicano/a
   b. □ Hispanic
   c. □ Latino/a
   d. □ Mexican
   e. □ Mexican American
   f. □ Nuevomexicano/a
   g. □ Spanish
   h. □ Other ______________

129. Earlier, I asked you where you fall on a ladder relative to other people in New Mexico. Where would you place this person on this ladder?

   *Please place an “X” on the rung where you think this person stands at this time in his/her life, relative to other people in New Mexico.*

   Rung # from bottom: _______
Map I. New Mexico counties
Map II. Albuquerque areas
Section IV. Phenotypic Information

General health indicators

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood pressure 1 (at survey start):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood pressure 2 (mid-survey):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood pressure 3 (end of survey):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waist circumference:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hip circumference:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body fat percentage:</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Check box upon completion:**

- Facial photographs
- DNA sample
- Blood spot

Skin reflectance measurements

<table>
<thead>
<tr>
<th>Forearm</th>
<th>Hand</th>
<th>Face</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix C

Paper citations and author contributions for Chapters 2 – 3

Chapter 2


Contributions:

Conceptualization: Meghan E. Healy, Deirdre Hill, Marianne Berwick, Heather Edgar, Keith Hunley.

Data curation: M. E. Healy, D. Hill, M. Berwick, K. Hunley.

Formal analysis: M. E. Healy, K. Hunley.

Funding acquisition: D. Hill, M. Berwick, H. Edgar, K. Hunley.


Methodology: M. E. Healy, J. Gross, K. Hunley.

Project administration: H. Edgar, K. Hunley.

Resources: D. Hill, M. Berwick, K. Hunley.

Software: M. E. Healy, J. Gross, K. Hunley.

Supervision: K. Hunley.

Validation: M. E. Healy, K. Hunley.

Visualization: M. E. Healy, K. Hunley.

Writing – original draft: M. E. Healy, K. Hunley

Writing – review & editing: M. E. Healy, D. Hill, M. Berwick, H. Edgar, J. Gross, K. Hunley.
Chapter 3


Contributions

Conception and design of the research: Meghan Healy, Heather Edgar, Carmen Mosley, Keith Hunley

Data collection: Meghan Healy, Carmen Mosley

Data analysis and interpretation: Meghan Healy, Keith Hunley

Drafting the article: Meghan Healy, Heather Edgar, Carmen Mosley, Keith Hunley

Critical revision of the article: Meghan Healy, Keith Hunley

Final approval of the version to be published: Meghan Healy, Heather Edgar, Carmen Mosley, Keith Hunley
Appendix D

Supplementary Material for Chapter 2

S1 Fig. African ancestry.
S2 Fig. Native American ancestry.
S1 Table.

Ethnic subgroup participant and parents.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Mother</th>
<th>Father</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spanish</td>
<td>MCN</td>
</tr>
<tr>
<td>Spanish</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td>MCN</td>
<td>4</td>
<td>47</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant</th>
<th>Spanish</th>
<th>MCN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish</td>
<td>28</td>
<td>2</td>
</tr>
<tr>
<td>MCN</td>
<td>2</td>
<td>53</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mother</th>
<th>Father</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spanish</td>
</tr>
<tr>
<td>Spanish</td>
<td>26</td>
</tr>
<tr>
<td>MCN</td>
<td>0</td>
</tr>
</tbody>
</table>
S2 Table.

Credible region analyses.

<table>
<thead>
<tr>
<th>Population</th>
<th>African Ancestry</th>
<th>Lower CI &gt; population mean</th>
<th>Upper CI &lt; population mean</th>
<th>Total different from mean</th>
<th>Individual(African ancestry) &gt; 0</th>
<th>Individual(African ancestry) &lt; 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico City</td>
<td></td>
<td>0.05</td>
<td>0.11</td>
<td>0.16</td>
<td>0.05</td>
<td>1.00</td>
</tr>
<tr>
<td>Oriente</td>
<td></td>
<td>0.05</td>
<td>0.30</td>
<td>0.35</td>
<td>0.35</td>
<td>1.00</td>
</tr>
<tr>
<td>CVCR</td>
<td></td>
<td>0.10</td>
<td>0.20</td>
<td>0.30</td>
<td>0.15</td>
<td>1.00</td>
</tr>
<tr>
<td>Pasto</td>
<td></td>
<td>0.05</td>
<td>0.42</td>
<td>0.47</td>
<td>0.05</td>
<td>1.00</td>
</tr>
<tr>
<td>Peque</td>
<td></td>
<td>0.05</td>
<td>0.35</td>
<td>0.40</td>
<td>0.20</td>
<td>1.00</td>
</tr>
<tr>
<td>Medellin</td>
<td></td>
<td>0.15</td>
<td>0.35</td>
<td>0.50</td>
<td>0.40</td>
<td>1.00</td>
</tr>
<tr>
<td>Cundinamarca</td>
<td></td>
<td>0.00</td>
<td>0.11</td>
<td>0.11</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>RGS</td>
<td></td>
<td>0.15</td>
<td>0.60</td>
<td>0.75</td>
<td>0.15</td>
<td>1.00</td>
</tr>
<tr>
<td>Quetalmahue</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Paposo</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Catamarca</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Salta</td>
<td></td>
<td>0.00</td>
<td>0.16</td>
<td>0.16</td>
<td>0.05</td>
<td>1.00</td>
</tr>
<tr>
<td>Tucuman</td>
<td></td>
<td>0.00</td>
<td>0.21</td>
<td>0.21</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>NMH</td>
<td></td>
<td>0.02</td>
<td>0.20</td>
<td>0.22</td>
<td>0.02</td>
<td>1.00</td>
</tr>
<tr>
<td>Spanish</td>
<td></td>
<td>0.00</td>
<td>0.18</td>
<td>0.18</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>All Spanish</td>
<td></td>
<td>0.00</td>
<td>0.21</td>
<td>0.21</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>MCN</td>
<td></td>
<td>0.03</td>
<td>0.22</td>
<td>0.26</td>
<td>0.03</td>
<td>1.00</td>
</tr>
<tr>
<td>All MCN</td>
<td></td>
<td>0.04</td>
<td>0.22</td>
<td>0.27</td>
<td>0.04</td>
<td>1.00</td>
</tr>
<tr>
<td>Non-HSL White Parent</td>
<td></td>
<td>0.00</td>
<td>0.14</td>
<td>0.14</td>
<td>0.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
### European Ancestry

<table>
<thead>
<tr>
<th>Population</th>
<th>Lower CI &gt; population mean</th>
<th>Upper CI &lt; population mean</th>
<th>Total different from mean</th>
<th>Individual(European ancestry) &gt; 0</th>
<th>Individual(European ancestry) &lt; 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico City</td>
<td>0.32</td>
<td>0.42</td>
<td>0.74</td>
<td>0.95</td>
<td>0.79</td>
</tr>
<tr>
<td>Oriente</td>
<td>0.30</td>
<td>0.20</td>
<td>0.50</td>
<td>0.85</td>
<td>1.00</td>
</tr>
<tr>
<td>CVCR</td>
<td>0.05</td>
<td>0.05</td>
<td>0.10</td>
<td>1.00</td>
<td>0.95</td>
</tr>
<tr>
<td>Pasto</td>
<td>0.32</td>
<td>0.37</td>
<td>0.68</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Peque</td>
<td>0.20</td>
<td>0.20</td>
<td>0.40</td>
<td>0.95</td>
<td>1.00</td>
</tr>
<tr>
<td>Medellin</td>
<td>0.25</td>
<td>0.25</td>
<td>0.50</td>
<td>1.00</td>
<td>0.95</td>
</tr>
<tr>
<td>Cundinamarca</td>
<td>0.16</td>
<td>0.16</td>
<td>0.32</td>
<td>0.95</td>
<td>1.00</td>
</tr>
<tr>
<td>RGS</td>
<td>0.30</td>
<td>0.25</td>
<td>0.55</td>
<td>1.00</td>
<td>0.75</td>
</tr>
<tr>
<td>Quetalmahue</td>
<td>0.05</td>
<td>0.00</td>
<td>0.05</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Paposo</td>
<td>0.15</td>
<td>0.30</td>
<td>0.45</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Catamarca</td>
<td>0.07</td>
<td>0.07</td>
<td>0.14</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Salta</td>
<td>0.32</td>
<td>0.37</td>
<td>0.68</td>
<td>0.68</td>
<td>1.00</td>
</tr>
<tr>
<td>Tucuman</td>
<td>0.32</td>
<td>0.37</td>
<td>0.68</td>
<td>1.00</td>
<td>0.89</td>
</tr>
<tr>
<td>NMH</td>
<td>0.17</td>
<td>0.18</td>
<td>0.36</td>
<td>1.00</td>
<td>0.94</td>
</tr>
<tr>
<td>Spanish</td>
<td>0.26</td>
<td>0.09</td>
<td>0.35</td>
<td>1.00</td>
<td>0.91</td>
</tr>
<tr>
<td>All Spanish</td>
<td>0.13</td>
<td>0.13</td>
<td>0.25</td>
<td>1.00</td>
<td>0.96</td>
</tr>
<tr>
<td>MCN</td>
<td>0.10</td>
<td>0.22</td>
<td>0.33</td>
<td>1.00</td>
<td>0.97</td>
</tr>
<tr>
<td>All MCN</td>
<td>0.09</td>
<td>0.24</td>
<td>0.33</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Non-HSL White Parent</td>
<td>0.64</td>
<td>0.00</td>
<td>0.64</td>
<td>1.00</td>
<td>0.64</td>
</tr>
</tbody>
</table>

### Native American Ancestry

<table>
<thead>
<tr>
<th>Population</th>
<th>Lower CI &gt; population mean</th>
<th>Upper CI &lt; population mean</th>
<th>Total different from mean</th>
<th>Individual(Native American ancestry) &gt; 0</th>
<th>Individual(Native American ancestry) &lt; 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico City</td>
<td>0.26</td>
<td>0.37</td>
<td>0.63</td>
<td>0.79</td>
<td>1.00</td>
</tr>
<tr>
<td>Oriente</td>
<td>0.20</td>
<td>0.45</td>
<td>0.65</td>
<td>1.00</td>
<td>0.90</td>
</tr>
<tr>
<td>CVCR</td>
<td>0.15</td>
<td>0.10</td>
<td>0.25</td>
<td>0.95</td>
<td>1.00</td>
</tr>
<tr>
<td>Pasto</td>
<td>0.32</td>
<td>0.47</td>
<td>0.79</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Peque</td>
<td>0.20</td>
<td>0.40</td>
<td>0.60</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Medellin</td>
<td>0.30</td>
<td>0.25</td>
<td>0.55</td>
<td>0.95</td>
<td>1.00</td>
</tr>
<tr>
<td>Cundinamarca</td>
<td>0.11</td>
<td>0.26</td>
<td>0.37</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>RGS</td>
<td>0.30</td>
<td>0.40</td>
<td>0.70</td>
<td>0.60</td>
<td>1.00</td>
</tr>
<tr>
<td>Quetalmahue</td>
<td>0.00</td>
<td>0.30</td>
<td>0.30</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Paposo</td>
<td>0.20</td>
<td>0.50</td>
<td>0.70</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Catamarca</td>
<td>0.00</td>
<td>0.14</td>
<td>0.14</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Salta</td>
<td>0.21</td>
<td>0.42</td>
<td>0.63</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Tucuman</td>
<td>0.37</td>
<td>0.26</td>
<td>0.63</td>
<td>0.84</td>
<td>1.00</td>
</tr>
<tr>
<td>NMH</td>
<td>0.21</td>
<td>0.21</td>
<td>0.43</td>
<td>0.88</td>
<td>1.00</td>
</tr>
<tr>
<td>Spanish</td>
<td>0.15</td>
<td>0.21</td>
<td>0.35</td>
<td>0.91</td>
<td>1.00</td>
</tr>
<tr>
<td>All Spanish</td>
<td>0.17</td>
<td>0.04</td>
<td>0.21</td>
<td>0.96</td>
<td>1.00</td>
</tr>
<tr>
<td>MCN</td>
<td>0.24</td>
<td>0.21</td>
<td>0.45</td>
<td>0.88</td>
<td>1.00</td>
</tr>
<tr>
<td>All MCN</td>
<td>0.27</td>
<td>0.13</td>
<td>0.40</td>
<td>0.93</td>
<td>1.00</td>
</tr>
<tr>
<td>Non-HSL White Parent</td>
<td>0.00</td>
<td>0.86</td>
<td>0.86</td>
<td>0.50</td>
<td>1.00</td>
</tr>
</tbody>
</table>
REFERENCES, COMPLETE

Who do you think you are? Season 5, Episode 4. (2014). USA: TLC.


Roybal, E. Joint resolution relating to the publication of economic and social statistics for Americans of Spanish origin or descent (1974). Retrieved from available: https://%0Awww.govtrack.us/congress/bills/93/hjres1083


