The Hybridized Self: Understanding Multiracial Identity Through Narrative and Computational Methods

by

Rachel Jacobson

A dissertation accepted and approved in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in Psychology

Dissertation Committee:

Sara J. Weston, Chair

Rob Chavez, Core Member

Mariah Kornbluh, Core Member

Aaron Gullickson, Institutional Representative

University of Oregon

Fall 2025

© 2025 Rachel Jacobson

DISSERTATION ABSTRACT

Rachel Jacobson

Doctor of Philosophy in Psychology

Title: The Hybridized Self: Understanding Multiracial Identity Through Narrative and Computational Methods

Cultural psychology has long emphasized that the self is a meaning system shaped by access to sociocultural norms. For multiracial individuals, whose lives span intersecting racial and cultural contexts, this dissertation investigates how they define themselves, how they represent the self when race is not salient, and how they navigate their multiple racial identities across situations. In Study 1, I analyze open-ended narratives of ethnicity-race–salient events to examine how multiracial individuals experience agency and communion in contexts of racialization, revealing a balance between assertion and belonging. In Study 2, I examine spontaneous self-concept when race is not primed, using computational text analysis to characterize the structure and content of cognitive self-representations. In Study 3, I identify the strategies multiracial individuals use to navigate their multiple racial identities, integrating qualitative coding with computational modeling. These studies advance a cultural-psychological framework of the multiracial self as dynamic, where differences emerge across racial backgrounds. In an ongoing negotiation, people learn when and how to emphasize or downplay race, connect across difference, and sustain belonging in diverse contexts. Over time, identity work becomes less effortful and more internalized. These results position multiracial identity as a model for understanding how individuals sustain meaning and stability amid cultural and racial plurality; it also encourages recognizing "multiracial" as a prism, unified by a shared structure, yet refracting differently in each person's experience. This dissertation includes previously published (and unpublished) co- authored material.

ACKNOWLEDGMENTS

To my advisor, Dr. Sara Weston, who turns ambitious ideas into reality. You have taught me the statistical foundation of research and gifted the most invaluable course of my PhD, Bayesian Statistics, opening my eyes to the intuition of causal inference. You have pushed my ideas to be pushed even further. I am amazed by your statistical repertoire, data science skills, and sharp writing prowess. This dissertation is made possible by your support. Even then, I cannot express all you have done for me. You remind me to reach higher, think deeper, and conduct research openly. These are pillars I will carry forward with me in life.

Dr. Rob Chavez, with your vital support, this research expands towards how we perceive others across multiple identities. I can rely on your philosophical musings about "the self" and ability to creatively bridge methods across disciplines. Your existential skepticism towards psychological approaches keeps me questioning everything, and that's a good thing.

Dr. Sanjay Srivastava, I am indebted to your encouragement of my research ideas, which has sown the seeds of this dissertation, and for getting them off the ground. Thank you for recruiting me and believing in my potential. Having a mentor who values mixed methods, and my interdisciplinary background in anthropology, shaped the trajectory of my own research direction.

Dr. Bradley T. Hughes, thank you for your unwavering guidance, feedback, and hours-long research conversations as a collaborator, friend, and mentor. Our interpersonal perception research (and puppy-sitting for Lolli) has brought me so much joy!

To my committee members, Dr. Mariah Kornbluh, who shows thoughtfulness, community empowerment, and rigor in her research, and graciously helped with my codebook and thematic analytic blueprint for the final study of this dissertation; and Dr. Aaron Gullickson, who taught

me that there is a lot to learn about multiracial identity outside of psychology. Since then, I have read many sociology papers from multiracial experts of the field.

Other faculty I feel so much gratitude for: Lori Olsen, Dr. Sara Hodges, Dr. Kate Mills, and Dr. Elliot Berkman.

For my husband, Loqman, who inspires and reminds me that is it not a matter of *if* but *how* a research question can be answered. Who sees and accepts me as I am. Who once teasingly said that his mother country, Iran, is the geographic midpoint between Eastern Europe and the Philippines, where my mom and great grandparents emigrated from.

My journey would not have been possible without my parents, parents-in-law, brother, friends, colleagues and collaborators throughout the program. I thank Denicia Aragon Espinosa, Clare McCann, Dr. Dillon Welindt, Dr. Moriah Stendel, Chiagoziem Adindu, Dr. Pooya Razavi, Dr. Sara Lieber, Mara Jennings, Dr. Vinita Vader, Dr. Tim Valshtein, and every cultural psychology classmate from SISPP (who I confidently entrust the future of the field to), for their endless support and wisdom.

I thank my research assistants and mentees who taught me the joy of being part of a team, and the meaning of "paying it forward" as my own mentors have done for me. I thank Christina Nyamuswa, Angie Suh, Piper Ahern, Debbie Okeke, Eleanor Struve, Madi Teopaco, Blu Mackey, Taylor Culinski, Phurbu Dolma for pushing themselves to grow as researchers and students.

"A self then, is the outcome of a process, unique to life, of maintaining and perpetuating an individual form, a form that, as it is iterated over the generations, grows to fit the world around it to maintain its selfsame identity, which is forged with respect to that which it is not; anteaters re-represent previous representations of ant tunnels in their lineage, but they are not themselves ant tunnels. Insofar as it strives to maintain its form, such a self acts for itself: a self, then, whether 'skin-bound' or more distributed, is the locus of what we call agency."

Eduardo Kohn, *How the Forests Think*

Table of Contents

CHAPTER I: INTRODUCTION **15**

DEFINING MULTIRACIAL 16
THEORETICAL FRAMEWORKS OF MULTIRACIAL IDENTITY 18
NAVIGATING CULTURAL WORLDS IN MULTIRACIAL IDENTITY DEVELOPMENT 20
Ecological systems theory *21*
Narrative identity theory *22*
Self-Concept & Identity Schema *25*
Rites of Passage *26*
Multicultural Framework *27*
'Multiracial' is not a monolith. *30*
QUALITATIVE APPROACHES TO ANALYZING TEXT 31
Grounded Theory *32*
Interpretive Phenomenological Analysis (IPA) *33*
Discourse Analysis (DA) *34*
Thematic Analysis (TA) *35*
Narrative Analysis *36*
QUANTITATIVE APPROACHES TO ANALYZING TEXT 38
Preprocessing *39*
Bag-of-words approaches & TF-IDF *40*
N-gram models *41*
Static Embeddings *42*
Contextual Embeddings *43*
Semantic similarity *45*
Sentiment Analysis *47*
Topic Modeling *48*
DISSERTATION AIMS 52

CHAPTER II: **56**

STUDY 1: USING STRUCTURAL TOPIC MODELING TO UNDERSTAND ETHNICITY/RACE-RELATED NARRATIVES **56**

INTRODUCTION 56
THE INTERSECTION OF STM AND NARRATIVE IDENTITY 57
THE PRESENT STUDY 60
METHOD 61
PARTICIPANTS AND PROCEDURE 61
MEASURES 62
Human Coded Agency and Communion *63*
ANALYSIS PLAN 64
Aim 1: Identify Ethnicity-Related Topics and Label Using Top Words and Example Narratives. *64*
Aim 2: Determine Whether the Content of the STM Topics Changes as a Function of Participants' Ethnicity. *68*
Aim 3: Examine the Relationship Between Extracted Topics and Themes of Agency and Communion. *68*
RESULTS 69
AIM 1: IDENTIFY/LABEL STM TOPICS 69
Topic 1: Awareness of Cultural Difference *72*

Topic 2: Peer Dynamics ... *73*
Topic 3: Othering ... *73*
Topic 4: Reactance to Prejudice ... *73*
Topic 5: Awareness of Underrepresentation ... *74*
Topic 6: Ethnic/Cultural Celebration ... *74*
Topic 7: Experience of Cultural Shock ... *74*
Topic 8: Violence ... *74*
Topic 9: Contact with Racial Discrimination ... *75*
Topic 10: Negative Affect ... *75*
Topic 11: Lack of Belonging in Cultural Space ... *75*
Topic 12: Ethnic Stereotypes at School ... *76*
Topic 13: Gained Perspective of Structural Inequality ... *76*
Topic 14: Whiteness ... *76*
AIM 2: TOPIC VOCABULARY BY PARTICIPANTS' ETHNICITY/RACE ... 76
AIM 3: AGENCY AND COMMUNION-RELATED TOPICS ... 85
DISCUSSION ... 88
FROM SHARED ETHNIC/RACIAL EXPERIENCES TO MULTIRACIAL SELF-CONCEPT ... 90

CHAPTER III: ... 92

STUDY 2: ... 92

SELVES IN CONTEXT: OPEN-ENDED SELF-DESCRIPTIONS REVEAL MULTIRACIAL SELF-CONCEPT ... 92

INTRODUCTION ... 93
THEORETICAL AND DEMOGRAPHIC RELEVANCE OF MULTIRACIAL SELF-CONCEPT ... 94
MULTIRACIAL IDENTITY MODELS ... 95
THE SELF-CONCEPT AS A SITE OF IDENTITY CONSTRUCTION ... 96
CONCEPTUAL GAPS IN EXISTING RESEARCH ... 98
OPPORTUNITY FOR PERSON-CENTERED, COMPUTATIONAL METHODS ... 99
WHY SELF-CONCEPT STRUCTURE MATTERS FOR MULTIRACIAL PSYCHOLOGY ... 100
OVERVIEW OF THE PRESENT STUDY ... 101
METHODS ... 103
PARTICIPANTS ... 103
MEASURES ... 105
Self-concept. ... *105*
Multiracial identity. ... *105*
Multiracial identity integration. ... *106*
ANALYSES ... 107
Content: What themes and categories emerge in multiracial individuals' self-concept descriptions? ... *107*
Complexity: Do multiracial individuals construct a self-concept characterized by higher semantic coherence (lower complexity) and dialecticism? ... *108*
Centrality: How does racial identity feature in self-concept? ... *110*
Similarity: How similar are multiracial individuals' self-concepts to those of their parent racial groups, and does multiracial identity integration influence similarity? ... *110*
TRANSPARENCY & OPENNESS ... 117
RESULTS ... 118
CONTENT: WHAT THEMES AND CATEGORIES EMERGE IN MULTIRACIAL INDIVIDUALS' SELF-CONCEPT DESCRIPTIONS? ... 118

COMPLEXITY: DO MULTIRACIAL INDIVIDUALS CONSTRUCT A SELF-CONCEPT CHARACTERIZED BY HIGHER SEMANTIC COHERENCE AND DIALECTICISM? 128
CENTRALITY: HOW DOES RACIAL IDENTITY FEATURE IN SELF-CONCEPT? 129
SIMILARITY: HOW SIMILAR ARE MULTIRACIAL INDIVIDUALS' SELF-CONCEPTS TO THOSE OF THEIR PARENT RACIAL GROUPS, AND DOES MULTIRACIAL IDENTITY INTEGRATION INFLUENCE SIMILARITY? 132
DISCUSSION 134
INTERSECTIONALITY BEYOND RACIAL IDENTITY 135
FROM CONCEPTUALIZING THE SELF TO NAVIGATING HYBRIDIZATION 136

CHAPTER IV: 137

STUDY 3: 137

NAVIGATING MULTIRACIAL IDENTITY ACROSS ENVIRONMENT, CULTURES AND PSYCHOLOGICAL PROCESSES 137

INTRODUCTION 137
MEASURING NAVIGATION STRATEGIES 139
SOCIOCULTURAL FACTORS 140
IDENTITY PROCESSES 141
WELL-BEING 142
LIMINAL LIFE EVENT 143
THE PRESENT STUDY 144
METHODS 145
MEASURED VARIABLES 146
Navigation Strategies. *146*
Sociocultural Factors *147*
Identity processes. *148*
Life Event *150*
Well-being. *150*
ANALYSIS PLAN 151
RQ1: What are navigation strategies endorsed by participants? *151*
OPEN-ENDED NAVIGATION STRATEGIES 151
CLOSE-ENDED NAVIGATION STRATEGY PROFILES 153
INDIVIDUAL-LEVEL MODEL 155
COMMUNITY-LEVEL MODEL 156
RQ2.2: Do navigation strategies vary by how integrated, alternating, and explored Multiracial individuals' racial identities are? *158*
RQ3: Are navigation strategies associated with well-being? *159*
RQ4: How are navigation strategies applied when describing a liminal life event that engages multiple racial identities? *160*
RESULTS 162
DESCRIPTIVES 162
RQ1: NAVIGATION STRATEGY PROFILES 164
RQ2.1: INDIVIDUAL AND COMMUNITY PREDICTORS OF NAVIGATION STRATEGY CLASS MEMBERSHIP 168
Individual-Level Predictors *169*
Community-Level Predictors *170*
RQ2.2: NAVIGATION STRATEGY CLASS MEMBERSHIP PREDICTS IDENTITY PROCESSES 172
RQ3: NAVIGATION STRATEGY CLASS MEMBERSHIP AS A PREDICTOR FOR WELL-BEING 174
NAVIGATION STRATEGY ENDORSEMENT AS A CONTINUOUS VARIABLE 176

RQ4: How are navigation strategies applied when describing a liminal life event that engages multiple racial identities? 179
Discussion 188
The Nature of Navigation Strategies 188
Insights from Open-Ended Strategies 189
Sociocultural Dimensions 190
Psychological Correlates: Integration, Exploration, and Well-Being 192
Navigation as Meaning-Making in Liminal Life Events 193
Study 3 Takeaways 195

CHAPTER V: GENERAL DISCUSSION 197

Applying Natural Language Processing to Measure Identity 197
Multiracial Identity Content 200
Monoracial-Multiracial Continuum 200
Multiracial Identity: Dual-Minority and Part-White 205
Conclusion. 208
Holding Multiplicity: Intersectional and Dialectical Features of Identity 209
Intersectionality 209
Dialecticism 210
Revisiting Identity Development Frameworks 210
1. Reframing what "integration" means in multiracial identity. 211
2. There are a psychological cost and developmental function of identity work. 211
3. Situating identity development within culture, not just psychology. 213
4. Multiracial identity is a model for understanding selves in pluralistic societies. 214

APPENDIX A 217

APPENDIX B 223

APPENDIX C 239

REFERENCES 253

LIST OF FIGURES

Figure Page

Figure 1.1. Semantic textual similarity between sentences 47

Figure 2.1. The four standardized metrics used to determine candidate models. 70

Figure 2.2. Correlations between the 14-topic model and the 5-topic model. 71

Figure 2.3. FREX words with Experience of Cultural Shock between multiracial and minority monoracial groups..... 82

Figure 2.4. The estimates between the STM topics with agency and communion..... 86

Figure 2.5a & 2.5b. Agency and communion scores predicting topic proportions. 87

Figure 3.1. Recruitment of eligible multiracial and monoracial participants. 104

Figure 3.2 Bipartite matching across two individuals' self-descriptors. 112

Figure 3.3. The racial grouping framework for self-concept similarity 115

Figure 3.4. Self-concept content composition by racial group 119

Figure 3.5. The 14 highest probability words for the STM topics..... 122

Figure 3.6. The dialectical self by racial group 129

Figure 3.7. Proportion of participants who mention race-related self-descriptors. 130

Figure 3.8. Network plots using logistic regression where the likelihood of one content category in self-concept is associated with the presence of race in self-concept. 131

Figure 4.1. Demographic breakdown of multiracial sample by ethnicity/race..... 146

Figure 4.2. Top 5 close- and open-ended navigation strategies by ethnicity/race..... 164

Figure 4.3. Endorsement of close-ended navigation strategies by latent class..... 167

Figure 4.4. Co-occurrence of latent class for close-ended strategies with strategies described by participants (in proportions)..... 168

Figure 4.5. Posterior distribution across identity process constructs 173

Figure 4.6a & 4.6b. Posterior distributions of class-specific means and pairwise class differences of strategy endorsement 175

Figure 4.7. Posterior distribution for navigation strategy endorsement by identity processes ... 179

Figure 4.8. The types of rites of passage varies by navigation strategy endorsement. 185

Figure 4.9. Association between rite of passage and valence of the narrative 187

Figure A.1. A correlation matrix between the 14 topics within the 14-topic model 217

Figure A.2a & A.2b. Agency and communion scores predicting topic proportions. 218

Figure B.1. Generational status and loci within the multiracial sample 226

Figure B.2. Query for GPT to classify identity content using Python code 226

Figure B.3. Determine the optimal number of topics in STM using four metrics 227

Figure B.4. Correlation matrix between topics within the 6-topic STM model 228

Figure B.5. Word clouds for the four GMM clusters derived from the text embeddings. 233

Figure B.6. Mean STM topic prevalence across GMM clusters 236

Figure B.7. Multiracial identity integration by ethnicity/race 238

Figure C.1. LCA using open- and close-ended strategies 245

Figure C.2. Number of close- and open-ended navigation strategies 246

Figure C.3. Reflected race category by ethnic/racial combination 246

Figure C.4. Multiracial individuals who endorsed self-identifying as “monoracial” 247

Figure C.5. Mean number of strategies endorsed per LCA class in 3-class model 248

Figure C.6. Four metrics to select optimal number of topics 248

Figure C.7. Word clouds for types of rites of passage described by ethnicity/race 249

Figure C.8. Rite of passage topic prevalence by identity processes 251

LIST OF TABLES

Table Page

Table 2.1. STM topics in the 14-topic model with exemplary narratives..... 72

Table 2.2. The five FREX words with each topic across the ethnic/racial groups. 78

Table 2.3. The mean agency and communion score, and topic proportions across ethnicity..... 84

Table 3.1. Descriptive statistics by multiracial and monoracial identity. 104

Table 3.2. Participants' self-concepts and their average cosine similarity score..... 112

Table 3.3. Example data for parent category similarity models 117

Table 3.4. Labeled STM topics in the 6-topic model with self-concept..... 123

Table 3.5. Self-concept categories, STM topic prevalence, DCS..... 126

Table 3.6. Best fitting model by racial group framework. 133

Table 4.1. Close-ended navigation strategies. 147

Table 4.2. Frequency of endorsed close- and open-ended navigation strategy by count 162

Table 4.3. Model fit indices for two- to six-class latent class solutions 165

Table 4.4. Predictors of navigation strategy class membership based on characteristics..... 170

Table 4.5. Posterior estimates of identity process means stratified by latent class 174

Table 4.6. Posterior estimates of well-being stratified by latent class..... 176

Table 4.7. Posterior estimates from Bayesian linear regressions predict associated outcomes.. 177

Table 4.8. Labeled STM topics in the 6-topic model with excerpt of most representative narratives. 183

Table A.1. Labeled STM topics in the 5-topic model with exemplary narratives. 219

Table A.2. Self-identified ethnic group membership by pan-ethnic group. 220

Table B.1. Ethnoracial combinations by pan-ethnic group..... 223

Table B.2. Self-concept categories: Significant mean contrasts from post-hoc analysis Holm's correction across ethnicity/race 227

Table B.3. Frequent and exclusive words by cluster for Spectral, AHC, GMM, and EGA from the LLM text embeddings 229

Table B.4. Internal validity measures across cluster approaches 231

Table B.5. Highest-probability self-concepts for each GMM cluster 233

Table B.6. Contradictory pairs: Significant mean contrasts from Holm's post-hoc analysis 236

Table B.7. Self-concept and semantic coherence within participants 238

Table C.1. Codebook for applied open-ended navigation strategies 239

Table C.2. Simulated narrative generation and refinement 243

CHAPTER I
INTRODUCTION

Multiracial individuals represent one of the least studied populations in psychology, despite their critical importance for demonstrating that race functions as a social construct rather than a biological reality (Gaither, 2015; Shih & Sanchez, 2005). Multiracial individuals demonstrate fluidity in their racial identification, allowing it to shift over time and vary based on their specific racial background and social environment (Sanchez et al., 2009; Townsend et al., 2009). Identification is shaped by internal preferences or familial heritage and by external factors such as phenotype, racial ambiguity, and the extent to which others racialize or assign race to the individual (Doyle & Kao, 2007; Khanna, 2011). These dynamics generate tension between self-definition and externally imposed classification, with consequences for psychological well-being and belonging (Chavez & Sanchez, 2010; Franco & O'Brien, 2018).

Understanding how individuals make meaning of their experiences sets the stage for examining broader questions of *identity*—how people come to define who they are, what they value, and where they belong (Cross & Cross, 2008). Identity development is a major task in one's lifespan and is essentially building a 'meaning-making system' (Oyserman & Destin, 2010). In decoding this meaning-making system, content is central to how the individual navigates everyday choices in life (Cross & Cross, 2008). Identity construction is inherently a social process, integrating culturally-shared expectations and story forms with personal experiences. This presents unique challenges for multiracial people.

As a group, multiracial individuals encompass many ethnic/racial backgrounds, partially due to an increase in interracial marriages and partially fueled by growth in immigration from outside of Europe (Törngren & Sato, 2021). People who are multiracial are worthy of study not only because they are an underrepresented population, but because they hold conceptual

significance: they challenge traditional classification systems. Rather than fitting neatly within fixed racial boxes, they often move along the blurred edges of these categories, drawing attention to the arbitrariness within social systems designed to sort people. Earlier research has often focused on part-white multiracial individuals (Poston, 1990; Root, 1996; Rockquemore & Brunsma, 2002; Renn, 2008). Today's multiracial population reflects increasingly diverse racial and ethnic combinations, experiencing multiple, and sometimes contrasting, sociocultural norms. These individuals actively navigate and blend cultural frameworks in ways that challenge static notions of identity. Many multiracial individuals are perceived by others as mixed, and they increasingly claim the label for themselves—finding in it a sense of centrality, belonging, and empowerment—laying the foundation for a growing body of research into their lived experiences and identity development.

It is within this context that my dissertation examines how multiracial people construct and navigate their multiple racial identities. By focusing on *individual differences* in identity development, how people define themselves, organize these self-definitions, and apply strategies in everyday life, I aim to illuminate the diversity and complexity within this population. Hence, I investigate how multiracial individuals actively engage in meaning-making as agents who creatively reshape cultural frameworks in the process of becoming themselves.

Defining Multiracial

In a national context where identity is both personal and politically charged, the act of categorizing people along racial lines carries weight. For multiracial individuals, whose existence defies singular classification, imposed labels can feel more like erasure than recognition. Institutions strive to define them and individuals themselves must navigate labels

that often fall short of lived experience. Measuring the multiracial population thus remains difficult (Starr & Pao, 2024).

The multiracial population is not a singular, bounded group but rather an evolving category shaped by context. Any operational definition must therefore grapple with these tensions. In this dissertation, the term "multiracial" will represent individuals who either self-identify as biracial or multiracial, identify with more than one racial identification, or select different races for their parents. This definition is informed by previous literature across psychology (Albuja et al., 2019; Albuja et al., 2022; Gaither, 2015), linguistics (Tsai et al., 2024), and sociology (Bean & Lee, 2009 ; Harris & Sim, 2002; Morning, 2000) and will follow participants' self-identification in defining our sample.

Defining who counts as multiracial thus requires attention to individual identification and to the shifting definitions of racial groups themselves. Race is a social construct (Omi & Winant, 1986) that evolves in response to shifting sociopolitical conditions and policy agendas in the United States. For instance, the one-drop rule historically codified racial boundaries by classifying anyone with Black ancestry as Black, effectively erasing mixed-race identities. In recent decades, communities such as Middle Eastern and North African (MENA) and Latine groups have advocated for recognition as distinct racial categories in the U.S. Census (Abdulrahim et al., 2012 ; Chávez-Moreno, 2021), a change finally enacted in March 2024. As definitions of racial groups change, the "multiracial" boundaries also shift; researchers who study this population therefore need to recognize this dynamic landscape and the diverse experiences it encompasses.

Theoretical Frameworks of Multiracial Identity

Multiracial identity models have evolved substantially, beginning with early frameworks such as the "marginal man" theory (Stonequist, 1935), which pathologized multiracial identity as inherently conflicted, deviant, or marginal, positioning individuals as socially alienated and psychologically fragmented due to their liminal racial status. This framework was later replaced with the *biracial identity development model*, which emphasizes identity formation as a potentially healthy and linear process moving from personal awareness to group integration (Poston, 1990), and the *ethnic identity development model*, which suggested that exploration and commitment are key processes in ethnic identity development and could be applied to multiracial youth (Phinney, 1990). These models shifted the discourse by framing identity not as inherently problematic, but as developmental and potentially integrated, even in the presence of complexity.

Later, ecological models further expanded the conceptual lens, emphasizing that multiracial students construct identity in relation to context, including institutional, peer, and family environments, which allowed individuals to identify monoracially, multiracially, or situationally across contexts (Renn, 2008). Similarly, a typology of racial identity options for Black–White biracial individuals, ranging from singular Black identity to transcendent (de-emphasis of race) or "border" identities, is grounded in empirical observations of how racial identity is socially embedded and responsive to societal cues (Rockquemore & Brunsma, 2002).

These later frameworks recognized that multiracial identity is neither fixed nor linear but contingent on environmental affordances, social validation, and the individual's own interpretive agency. Much of this foundational work, however, has centered on individuals of Black–White or Asian–White descent, reflecting historical and demographic patterns in the U.S. While this research has offered critical insights, it overlooks the lived realities of individuals with less

frequently studied racial combinations, such as Asian–Latine, Asian–Black, Latine–Black, or more than two races, whose experiences are equally complex and worthy of scholarly attention.

The ecological model emphasizes the importance of nested social contexts (e.g., family, peers, and institutions) in shaping multiracial identity, and a cultural approach extends this framework by illuminating the *symbolic and interpretive processes* embedded within those settings. There is value in measuring how cultural meaning is internalized or contested in everyday life. From transmission to negotiation, this would emphasize how individuals draw from multiple, and sometimes conflicting, cultural repertoires as they navigate identity. In doing so, incorporating cultural meaning invites attention to the everyday ways that multiracial individuals make sense of themselves within, and sometimes in tension with, dominant cultural narratives.

Culture, in this context, refers to the shared norms, values, practices, and symbols that give meaning to everyday life, as a way of being (Markus & Hamedani, 2019). This includes both visible customs—such as rituals and language use—and more implicit dimensions like family roles, and worldviews, all of which are transmitted across generations (Geertz, 1973; (Swidler, 1986). Scholars have variously defined culture as symbolic systems (Keesing, 1974), normative guides to interaction (Swidler, 1986), and mechanisms for enacting social meaning. Typically race, ethnicity, and culture are distinguished from one another in research, with race referring to phenotype and hierarchical societal structure; ethnicity to shared ancestry and traditions; and culture to learned meaning systems (Gonzales-Backen, 2013; Richeson & Sommers, 2016). However, when it comes to the lived experience of individuals, the separation between culture, race, and ethnicity is not clearcut (Umaña-Taylor et al., 2014; Song, 2016). Rather, these dimensions are experienced simultaneously within an individual and can activate as

individuals encounter new contexts or navigate multiple group memberships. Therefore, understanding identity requires attending to the ways culture, race, and ethnicity interweave in lived experience, resisting overly compartmentalized frameworks that obscure their mutual constitution.

Overall, a cultural framework of identity extends the ecological question of *where* individuals are embedded and asks *how* those environments are internalized into the self. For instance, we might ask: How do they make meaning of conflicting cultural norms? What scripts do they internalize about belonging, family, and identity performance? We might consider the development of a *cultural self-schema*—an internalized cognitive-affective model that guides how people interpret experiences and understand their place in the world (Markus & Kitayama, 1991). Just as bicultural individuals may develop distinct self-schemas for each cultural frame, multiracial individuals may construct hybrid or negotiated schemas that blend, juxtapose, or compartmentalize the cultural influences in their lives. In this way, identity expands beyond context into an internally-organized system for navigating the world, one that is deeply shaped by the presence (or absence) of culture.

Navigating Cultural Worlds in Multiracial Identity Development

The study of multiracial identity offers a crucial window into the processes by which people navigate, interpret, and integrate cultural worlds. Understanding multiracial identity requires attention to the lived, interpretive dimensions of multiracial experience. As an overview, this work should draw upon a set of theoretical frameworks—ecological systems theory, narrative identity theory, schema-based models of the self, rites of passage, multiracial and multicultural identity development literature—to deepen our understanding of how identity is formed across multiple levels of experience. Rather than approaching these frameworks as

isolated or competing models, the frameworks capture how multiracial identity unfolds across societal, interpersonal, and individual domains.

Ecological systems theory

Identity emerges through interactions between individuals and the cultural ecologies in which they are embedded (Bronfenbrenner, 1992). The ecological model offers a valuable framework for understanding these multilayered contexts. Individuals are situated within both proximal contexts (e.g., family, peers, community) and distal contexts (e.g., local, national, global), which provide support, pose challenges, or exert forms of influence. For example, parents may seek to prepare their children for microaggressions or hybridized practices that occur within immediate social environments. However, understanding the experiences faced by multiracial individuals necessitates attention to macrosystemic structures (Rogers et al., 2021). The broader societal architecture that shapes parental preparation includes pervasive ideologies concerning ethnicity, heritage, and belonging. In the United States, two such ideologies (anti-Blackness and colorblindness) are embedded within the macrosystem (Moffitt et al., 2022). Crucially, interpersonal and systemic racism are not synonymous; individuals may vocally oppose interpersonal racism while simultaneously perpetuating systemic inequities (Bonilla-Silva, 1997). The personal narratives of multiracial individuals offer insights into how these layered forms of racism are experienced.

Accordingly, contextual variables must be considered alongside personal narratives to consider the systemic structures that shape multiracial identity development. While Törngren et al., (2021) provides a framework that integrates individual factors (e.g., ethnoracial identification, language proficiency, phenotype) and contextual factors (e.g., socio-ethnic stratification, heritage contact, societal attitudes), little empirical work has examined how these

factors interact to shape the navigation strategies multiracial individuals use during multiracial identity navigation.

National structures play a critical role in shaping these strategies. In several European countries and in non-European nations such as Japan, colorblind policies have contributed to the emergence of large populations of multiracial individuals who identify as "neither-nor" (Törngren & Sato, 2021; Törngren, 2018). In many contexts, the absence of inclusive cultural scripts compels multiracial individuals to devise individualized or hybrid strategies, whereas monoracial groups may rely on collective narratives reinforced across multiple ecological levels. As a result, the multiracial label may reflect exclusion rather than dual identification (Jones & Rogers, 2022; Giamo et al., 2012). Empirical work that contextualizes navigation strategies is essential to understanding how systemic forces, cultural scripts, and individual meaning-making intersect in multiracial identity development.

Narrative identity theory

With two or more ethnoracial identities to draw on, the question of how multiracial people form a coherent sense of identity remains important yet unexplored, especially from multiracial participants' own voices. One form of identity that allows people to find a sense of unity and purpose in their words is narrative identity (McAdams, 2001). Narrative identity is a distinct component of personality that represents a psychosocial construction of the personal past, present, and conjectured future. This theoretical construct allows for described life events and experiences to become incorporated into one's self concept.

Two levels of memory processing shape narrative identity. First is the selection of events that are deemed significant within one's identity-defining narrative (McAdams, 2001). Second is a level of processing that involves forming a narrative about a significant memory to the present

self; this narrative either elucidates, transforms, or confuses identity. Both memory processes are socially-influenced, integrating culturally-shared expectations and story forms with personal experiences.

Multiracial individuals face unique challenges in constructing an identity narrative. A key tension is the need to reconcile identities that are often socially perceived as incompatible or mutually exclusive (Roccas & Brewer, 2002). Additionally, dominant cultural narratives may lack scripts for multiracial identity, making it harder for individuals to find story forms that reflect their lived experience. This absence can complicate efforts to craft a coherent and affirming personal narrative, particularly when social expectations and interpersonal feedback pressure individuals to prioritize or choose among identities.

The way people construct narratives is intricate: personal history and experience provide the raw material, shaped by cultural story forms and an individual's personality and narrative style (McAdams & McLean, 2013; Syed & McLean, 2021). Umaña-Taylor and colleagues (2014) describe ethnoracial identity as a process in which people actively negotiate and internalize their racial and ethnic identities. Because multiracial people are in a unique position to describe their identities with flexibility based on context, recent narrative research has adopted an ecological perspective by incorporating macrosystems (Dunlop & Wilkinson Westberg, 2022; Moffitt et al., 2022). Thematic features within narratives reveal a collective, interconnected culture of which the individual is a part. By integrating what is tacitly evident in narratives into a more structural analysis, this approach offers a richer understanding of how multiracial individuals navigate between personal meaning-making and sociocultural structures, highlighting the dynamic ways ethnicity and race shape, and are shaped by, narrative identity.

As for life experiences salient for multiracial individuals, race- and ethnicity-relevant events, such as experiences of inclusion, exclusion, discrimination, or cultural affirmation, were often described in identity narratives (Syed & Azmitia, 2008). Those who experience discrimination may also report a stronger sense of commonality and solidarity with other multiracial people (Giamo et al., 2012). This sense of connection is echoed in findings that many multiracial individuals express a desire for their racial identity to be accurately perceived by others (Remedios & Chasteen, 2013). These recurring themes of marginalization and belonging fall under the concept of identity content, which refers to the attitudes and beliefs individuals hold about their ethnic-racial groups and how those groups relate to others (Syed & Azmitia, 2010).

While a growing body of research has explored how individuals use narratives to construct meaning around their racial or ethnic identity, there remains a significant gap in understanding how multiracial people draw on fundamental motivational themes, particularly agency and communion, when constructing identity narratives. *Agency* (the drive toward autonomy, mastery, and self-definition) and *communion* (the drive toward connection, belonging, and relational harmony) are central to narrative identity development (McAdams, 2001). These two constructs reflect core dimensions of the human experience and align closely with other foundational psychological constructs such as individuality and connectedness (Grotevant & Cooper, 1986) and independent versus interdependent self-construals (Markus & Kitayama, 1991). Yet we know little about how individuals balance these motives in contexts where race and ethnicity are made salient, where rigid racial categories can constrain agency or where belonging is complicated by ambiguous or shifting group boundaries. Social life scenarios, such as navigating peer groups, family dynamics, schools, workplaces, or community spaces, often

force multiracial individuals to reconcile these tensions. To date, narrative identity research has rarely examined how multiracial people actively use agency to resist imposed categories or use communion to bridge cultural norms in everyday interactions. These negotiations reveal identity construction as both a personal and sociocultural process – strategies of meaning-making that remain understudied in the multiracial literature (Svensson & Syed, 2019; Syed & Azmitia, 2008).

Self-Concept & Identity Schema

While motivational forces like agency and communion guide multiracial individuals through negotiated racial identity in social contexts, other dimensions guide identity when external racial cues are absent. Understanding self-concept in these moments offers insight into how multiracial identity extends beyond racial integration alone and encompasses broader domains of meaning. In cultural psychology, "I" is at the center of a person's experiences and a referent for agency (Markus & Hamedani, 2019). When self-construals are widely shared and reinforced through ideas, institutions, and interactions, they form cultural blueprints: mental models of how to think, feel, and act. Yet for multiracial individuals, these blueprints may be limited, contradictory, or entirely absent due to their ambiguous or fluid racial status (Jackson et al., 2012 ; Shih & Sanchez, 2005). As a result, they may experience marginalization or feelings of disconnection, while also developing a heightened capacity to reconcile divergent cultural expectations (Cheng & Lee, 2009; Gaither, 2015).

These conditions may foster the emergence of a differentiated self-concept—one that reflects a sense of uniqueness or distinctiveness from monoracial peers (Rockquemore & Brunsma, 2002; Shih & Sanchez, 2009). Yet little research has examined which dimensions of the self come to the forefront in the absence of racial cues. Do multiracial individuals rely more

heavily on values, roles, or personal traits to define themselves? Exploring this question offers a broader understanding of identity development beyond the centrality of race or ethnicity. Cultural models of self and schema theory (Markus & Kitayama, 1991; Oyserman & Destin, 2010) offer a lens through which to understand not just *what* identities people hold, but *how* those identities are mentally structured and mobilized in everyday life. For multiracial individuals, these schemas may be shaped by both internal negotiation and the external cultural cues that signal belonging or exclusion.

Rites of Passage

Of course, self-concept does not develop in isolation. How multiracial individuals frame their identity is intertwined with how they engage with cultural scripts during major life transitions. These moments, often structured by societal narratives and rites of passage, present unique challenges and opportunities for meaning-making across sociocultural contexts. Anthropologists define a *rite of passage* as the liminal crossing of borders into a culturally different realm, where rites can be those of initiation or detachment (van Gennep, 1960 ; Turner, 1969). For multiracial individuals, narratives may draw upon culturally defined scripts yet often require adaptation when such scripts fail to account for their complex lived realities. Identity is constructed through culturally specific tools, discourses, and master narratives (Hammack, 2008; Syed & Svensson, 2014), and these narratives are themselves shaped by broader cultural environment. For example, a multiracial person may interpret a transformative event—such as leaving home, entering college, attending a funeral, or visiting a parent's homeland—through the lens of a culturally scripted rite of passage. Such transitions represent liminal spaces in which identity is tested, redefined, and, at times, asserted. Bauer & McAdams (2004) emphasize the importance of life transitions. During these transitions, individuals are more likely to reevaluate

their personal narratives and integrate new experiences into their sense of self. They argue that how people interpret and make meaning of these transitional periods can influence the trajectory of their identity over time. Little is known about the strategies that multiracial individuals employ in these contexts. These are often the very places where narrative, cultural structure, and the self most visibly intersect.

Multicultural Framework

The porous boundaries between race, ethnicity, and culture encourage the application of a multicultural framework. Multiracial individuals often hold access to multiple cultural reference systems and exhibit the capacity to engage with each meaningfully (Albuja et al., 2022). *Multicultural* individuals exhibit cultural competence across two or more cultural systems, often demonstrated through their ability to fluidly shift between languages, norms, and values depending on context (Benet-Martínez et al., 2014). This is particularly exhibited by individuals raised in households, communities, or peer networks that expose them to distinct, and sometimes contrasting, cultural traditions. By recognizing culture within identity, the cultural approach moves beyond categorization to illuminate the internal processes by which multiracial individuals give meaning to their experiences. Yet, few empirical studies systematically examine how multiracial individuals draw from, are socialized into, or selectively engage with the cultural traditions of their heritage backgrounds (Brittian et al., 2013; Brown et al., 2007; Villegas-Gold & Tran, 2018).

Building on multicultural identity research, scholars have increasingly noted key distinctions between multiracial individuals versus monoracial multicultural individuals. Although both groups may engage with multiple cultural traditions, their experiences diverge in

three important ways: racial assignment, within-family multiethnic exposure, and ethnic legitimacy.

First, *racial assignment*: the racial dimension of identity introduces social meanings that are externally imposed and hierarchically consequential—meaning a multiracial person may contend with racial hierarchical structure within themselves. For instance, multiracial individuals have described experiences with family members, friends, strangers who used microaggressions, derogatory slurs, or fetishization (Johnson, 2024; Nadal et al., 2013). Externally imposed racial hierarchies can shape their identity within systems of racial stratification. In contrast, family members can also socialize multiracial people and transcend racial hierarchical tensions (Childs et al., 2021; Csizmadia & Atkin, 2022). To illustrate this, multiracial individuals recall moments in which extended relatives emphasized pride in one's non-overlapping heritage or treated them with unconditional acceptance irrespective of phenotype. Relatives can thus express support by rejecting colorist assumptions or refuse to hierarchically rank parts of their identity. These affirming moments demonstrate how family relationships can act as protective against broader societal invalidation.

Within multiracial households, children are raised under different socialization strategies; some families emphasize learning about racial identities, while others de-emphasize race altogether (Shih et al., 2007). For example, mothers of biracial Black/White children who openly discussed race fostered greater flexibility in racial identity choice and promoted higher well-being (Morrison, 1995). Societal perceptions further complicate these dynamics. Multiracial individuals can be phenotypically ambiguous and are often misclassified (Nicolas et al., 2019). Consequently, they may be acculturated in their identities yet still excluded because of their non-prototypical appearance. Such patterns illustrate how external categorization, societal valuation,

and racial hierarchies intersect in shaping the multiracial experience who are multicultural (Axt et al., 2014).

Second, *within-family multiethnic exposure*: while many biculturals move between cultures tied to nation-states or geographic boundaries, multiracial individuals often inhabit hybrid cultural environments *within* a single family or household, where contrasting norms may be present simultaneously and embodied relationally (e.g., navigating one parent's family-centered orientation and another's individualist framework). Some individuals describe jointly celebrating traditions from both sides of the family and embracing their multiple racial and ethnic identities in an integrated way. Others, however, describe attending cultural events separately or practicing customs independently. When occasions are kept separate, this can give some individuals the opportunity to compare their identities. For others, it creates tension such that the individual feels like they have to choose between them. These within-family dynamics emphasize how multiracial individuals do not simply alternate between discrete cultural frameworks but can engage in a continuous, relational negotiation of identity within their most intimate spaces.

Third, *ethnic legitimacy*: Multiracial individuals often face challenges to being recognized as legitimate members of a particular ethnic group: not because of how they look (as in racial assignment), but because they may not meet community-specific standards of cultural authenticity, though this is often additive with their appearance. These standards often center on language ability, knowledge of traditions, or adherence to behavioral norms. While this issue also affects monoracial biculturals (e.g., someone wondering "Am I Korean enough?"), for multiracial individuals, the question may be "Am I Korean at all?" For example, a Korean-Mexican woman raised solely by her Mexican father may identify as Mexican American—not

due to a rejection of her Korean heritage, but because her cultural grounding lies fully in one community.

This example underscores an important point: not all multiracial individuals are raised with equal access to the cultural traditions of their heritage. Some grow up in monocultural households, where one parent, or one set of cultural norms, dominates the transmission of identity-relevant knowledge. A cultural approach remains relevant because identity contains the perceived meanings and absences tied to one's background, not only exposure to practices. For example, as described above, the Korean-Mexican woman raised by a single father may still experience cultural longing, silencing, or strategic passing tied to the unrepresented part of their identity. Moreover, even in predominantly monoracial settings, multiracial individuals may be racialized in culturally specific ways that shape how they experience themselves and are treated by others.

'Multiracial' is not a monolith.

Lastly, while the term *multiracial* may serve as a unifying category, it must not obscure the heterogeneity of experiences encompassed within it. Scholarship on dual-minority multiracial experiences remains limited, with few quantitative (Albuja et al., 2020) and qualitative (Childs et al., 2021; Hordge-Freeman & Veras, 2020; Ortiz, 2017) studies addressing this important topic. The common conflation of multiracial identity with White ancestry obscures the specific challenges faced by multiracial communities subject to racial injustice (Garay & Remedios, 2021). Dual-minority individuals often contend with intersecting forms of classism, colorism, and marginalization linked to non-White ancestry. Their experiences may differ markedly from those of majority-minority multiracial individuals who, in certain contexts, may benefit from White privilege, though such privilege is typically contingent on phenotypic appearance (Albuja

et al., 2018; Gaither et al., 2016). Regional variation within the United States further shapes multiracial experiences, as individuals residing in areas with high multiracial representation (e.g., Hawai‘i, California) report different patterns of identity negotiation than those in regions where multiracial populations are less visible (Meyers et al., 2022; Pauker et al., 2018). Additional factors, including generational status and phenotypic ambiguity, further contribute to the complexity of multiracial identity development (Norman & Chen, 2020). Advancing our understanding of these processes requires research that centers the diversity of multiracial experiences and the systemic conditions that shape them. To translate these cultural and narrative approaches into empirical inquiry, it is necessary to adopt methodological tools that capture the richness of multiracial identity as it is expressed in open-ended data.

Qualitative approaches to analyzing text

Understanding the lived experiences and identities of individuals can benefit from rich, in-depth analysis of textual data. As described earlier, a cultural framework to multiracial identity development positions self-definition as a meaning-making process shaped by shared story forms, such as rites of passage, and by thematic dimensions like agency and communion. From this perspective, self-concept and identity are constructed and expressed in language, whether through narratives or concise, open-ended self-descriptions. For multiracial individuals, these written and spoken accounts capture how cultural contexts, life transitions, and social positioning intersect in the process of defining “who I am.”

Because these identity expressions take linguistic form, qualitative analysis becomes a natural methodological bridge between cultural theory and meaning-making processes, especially when researchers seek to center participants’ voices and interpret context. These approaches vary in their epistemological assumptions and analytic aims, from uncovering

thematic patterns in participants' accounts to constructing conceptual frameworks grounded in data. However, they all share a commitment to interpreting language as a vehicle for psychological insight. Below, I outline several key qualitative methods used in psychological research to analyze textual data, ranging from close linguistic examinations to theory-driven narrative interpretation.

Grounded Theory

A foundational approach to qualitative inquiry is Grounded Theory, which offers a systematic method for generating theory directly from data through iterative coding, bridging with theoretical ideas, and constant comparison (Glaser, 1992; Strauss & Corbin, 1998). Its focus is more on theory generation than pattern recognition. Rather than beginning with a preexisting theoretical framework, Grounded Theory allows researchers to build inchoate theory from the ground up, using data as the raw material for conceptual development. It is especially useful when investigating underexplored processes or social phenomena for which no clear explanatory model yet exists. The method involves open coding (identifying emergent concepts), axial coding (linking concepts into categories), and selective coding (constructing a core theory around a central process), while also incorporating theoretical sampling to refine and saturate developing categories (Strauss & Corbin, 1998).

One well-known example in psychology is Charmaz's (1995) study of identity work among individuals with chronic illness. Through in-depth interviews, Charmaz inductively developed a theory of "loss of self," capturing how participants experienced diminished social roles, disrupted routines, and stigmatized identities. Her analysis explained *how* and *why* these experiences coalesced into a shared process of identity erosion. While the study demonstrated

Grounded Theory's capacity to move from individual accounts to a broader understanding of the mechanisms underlying psychosocial change, the approach has several limitations.

Grounded Theory often requires extensive iterative analysis over many cycles. There is also a risk of premature closure, where categories are forced to fit emerging theory rather than remaining grounded in participants' accounts. While Glaser's original formulation emphasized a more discovery-oriented stance, later adaptations (e.g., Charmaz, 2006) acknowledge the co-constructed nature of data, introducing tensions around epistemology and objectivity. Finally, Grounded Theory's focus on process and abstraction can sometimes distance the analysis from participants' rich, contextualized narratives, making it less suitable for research questions that prioritize depth over generalizability.

Interpretive Phenomenological Analysis (IPA)

In contrast to Ground Theory, IPA focuses more directly on participants' lived experiences and how they make meaning of them, emphasizing idiographic depth over theoretical abstraction (Smith et al., 1999; Smith & Osborn, 2003). Central to IPA is the double hermeneutic: the researcher is engaged in making sense of the participant, who is themselves engaged in making sense of their own experiences. IPA assumes an approach that is particularly concerned with the subjective, situated experience of individuals and how they interpret complex psychological phenomena.

While IPA is best suited for small, homogeneous samples, typically fewer than 10 participants, its generalizability is restricted away from cross-case analysis. Moreover, IPA requires a high degree of interpretive sensitivity, and the line between describing and over-interpreting participants' accounts can be difficult to manage. Finally, IPA is bounded by

theoretical frameworks as a method, which may constrain openness to alternative ways of understanding participants' accounts (Braun & Clarke, 2006).

Discourse Analysis (DA)

Discourse Analysis shifts the lens toward how language itself constructs experience, identity, and social reality (Potter and Wetherell, 1987; Willig, 2003). The subjective experience is de-emphasized in favor of language as constitutive. Discourse Analysis is a family of approaches that investigate how language, in speech, text, or other communicative forms, produces meaning and enacts social practices (Gee, 2014). Discourse analysts examine less the content and more the framing of what is said, what assumptions it draws upon, and what social functions it serves. DA is particularly useful when the research aim is to explore how personal accounts are shaped by culturally available discourses or institutional contexts. For example, Wetherell & Potter (1992) work on race talk in New Zealand examined how interviewees drew on socially shared repertoires to position themselves as non-prejudiced while simultaneously reproducing racial hierarchies. Both IPA and Discourse Analysis share a concern for how meaning is constructed, but they differ in scope. IPA's theoretical grounding in phenomenology and hermeneutics prioritizes the individual's lived experience, whereas Discourse Analysis draws on broader social constructionist and critical theories to situate those experiences within the cultural, historical, and institutional discourses that shape them.

As a limitation, Discourse Analysis gives less weight to the enduring, personal significance of experiences as felt and remembered. Its focus on framing can fragment the continuity of a person's account and oversee individual differences. Moreover, its orientation toward context-specific variability can make it challenging to distill broader, recurring patterns of meaning across different speakers or settings.

Thematic Analysis (TA)

Thematic Analysis is a qualitative approach that turns toward identifying patterns across a dataset. It offers a flexible but systematic way to distill recurring meanings while still situating them within the contexts in which they arise. TA enables researchers to both work inductively from the data or apply existing frameworks deductively, making it especially valuable for theory-informed research that still leaves space for discovery (Braun & Clarke, 2006).

TA can be approached from different epistemological positions: an essentialist stance treats themes as reflecting an underlying reality (similar to IPA), a constructivist stance views themes as products of social and cultural contexts (similar to DA), and a contextualist stance bridges the two. Themes may be identified at the semantic level, capturing what is explicitly stated, or at the latent level, which interprets the assumptions, conceptualizations, or ideologies that shape those statements. Braun and Clarke (2006) use the analogy of a 3D jelly to distinguish the two: semantic analysis examines the surface of the jelly, its shape, color, and texture, while latent analysis probes what ingredients, mold, and processes gave the jelly its form. For example, at the semantic level, a participant might say, "I never talked about race in my family," which would be coded as the absence of racial discussion. At the latent level, this could be interpreted as reflecting a colorblind ideology shaped by historical or cultural narratives about race. Latent themes require deeper interpretation, which often leads to lower interrater agreement compared to semantic themes, where meaning is explicit and subtext is minimal.

In practice, TA has taken several forms. *Reflexive TA* treats the researcher's subjectivity and reflexivity as analytic resources, rejecting the idea that coding should be purely mechanical or standardized (Braun & Clarke, 2006; 2021). *Coding reliability TA* emphasizes coder agreement and replicability, often using multiple coders and intercoder reliability metrics to ensure consistency. *Codebook TA* occupies a middle ground between the two, using a structured

codebook to guide analysis but without being rigidly focused on measuring coding reliability. This adaptability makes TA well suited for research aiming to link individual accounts to broader conceptual categories while preserving the specificity of participants' perspectives.

As an important consideration, Braun and Clarke emphasize that themes do not simply "emerge" from data; they are actively sought out by the researcher, shaped by theoretical assumptions about what the data represent and how they reflect the world. This acknowledgement emphasizes positionality of the research team. The process should be transparent about what is being done, by whom, why, and how. While prevalence across a dataset may be relevant, it is not the deciding factor for defining a theme; analytic relevance to the research question is. The flexibility of TA is in its ability to move between inductive pattern-finding and deductive application of theoretical frameworks, making it a useful bridge to other qualitative approaches. Its inductive dimension parallels the theory-building aims of Grounded Theory, while its deductive, theory-informed analyses resonate with the theoretical foundations of IPA and DA. This dual capacity provides a natural transition into Narrative Analysis, which, like TA, seeks to identify patterns across accounts but does so through the lens of story structure, temporal sequencing, and the functions narratives serve in meaning-making identity.

Narrative Analysis

Narrative Analysis focuses on the stories people tell about their lives, treating these accounts as descriptions of events and as structured, meaning-rich constructions that reveal how individuals understand themselves and their worlds. This approach is especially powerful for studying identity because it integrates psychological dimensions such as self-continuity, agency, communion, and evaluative framing of life events—core processes through which people make sense of who they are across time (McAdams, 1993; McLean & Syed, 2015). By examining both

what is told and how it is told, researchers can trace the ways individuals organize past experiences, interpret challenges, and position themselves within broader cultural narratives.

Human coding remains a vital tool in narrative research, supported by decades of work establishing rigorous codebooks and training protocols that enhance reliability and anchor analyses in psychological theory (McAdams, 1993; Syed & McLean, 2016). For example, human coding can detect complex narrative themes like redemption—stories that transform adversity into growth or contribution (McAdams et al., 2001)—as well as behavioral strategies tied to managing conflict or navigating identity (Dunlop & Tracy, 2013; McLean & Pratt, 2006). These themes are often expressed in complex, culturally situated ways, requiring attention to both explicit statements and implicit patterns of justification or action.

Most narrative research involves human coders, who use a codebook outlining various themes and categories to guide the assessment of particular narrative themes of interest (e.g., (Adler & Poulin, 2009; Singer, 2004; Syed & Nelson, 2015). As prefaced earlier, human coding offers advantages, such as the flexibility to choose themes and the ability to compare studies using standard manuals for theoretical consistency (McLean et al., 2020).

Narrative psychology, as outlined by McAdams (2013), examines how people craft and internalize life stories to integrate past experiences, present self-understanding, and future aspirations into a coherent whole. These personal narratives are not neutral records of events; they are meaning-making frameworks shaped by both individual agency and cultural scripts. Like IPA, narrative psychology centers on subjective experience, but it extends this focus by embedding discrete moments within the broader arc of a life story. It shares with discourse analysis an interest in the sociocultural contexts that inform storytelling, yet it treats the

narrative's temporal structure, its plot, turning points, and evaluative framing, as the primary unit of analysis.

These five qualitative approaches highlight complementary approaches to meaning-making in psychological text data: from generating theory grounded in participants' accounts, to interpreting lived experience, to situating talk within sociocultural contexts, to mapping shared thematic patterns, and finally to capturing the unfolding arc of personal narratives. While each offers unique analytic depth, their application, particularly with large datasets or lengthy text, can be time- and labor-intensive, requiring extensive human coding and interpretive engagement. In such cases, augmenting qualitative insight with computational tools can extend analytic reach. For example, analyzing hundreds of narratives or open-ended responses may be made more efficient through automated approaches that capture granular details at scale. Recent advances, such as open-access topic modeling methods, create new opportunities to complement in-depth qualitative interpretation with scalable, systematic analysis of large volumes of text data.

Quantitative approaches to analyzing text

In recent years, text analysis has significantly advanced, offering new tools for psychological research. These methods allow for systematic, reproducible, and scalable analysis of narrative data that would otherwise be unmanageable through traditional manual coding. For studies focusing on multiracial populations, where participants' narratives can be rich and complex, computational tools provide an efficient way to capture and compare patterns across subgroups, such as dual-minority and majority-minority multiracial individuals, while preserving nuance within their responses.

Preprocessing

Before applying computational models, textual narratives must be transformed into a format suitable for quantitative analysis. Preprocessing ensures that raw responses are represented in a consistent, analyzable form (Denny & Spirling, 2017; Manning et al., 2008). This is important in studies where narratives may include culturally specific terms and heterogeneous linguistic patterns that need to be preserved while removing noise. Core preprocessing steps often include:

- Converting all characters to lowercase and standardizing misspelled words, to ensure that semantically identical words are treated as the same token,
- Tokenization, or splitting text into words or sequences of words, which provides the fundamental units for computational models,
- Stopword removal, which filters out high-frequency function words (e.g., *the*, *and*, *of*) that contribute little to semantic content in most analytic contexts,
- Stemming or lemmatization, which reduces words to their root or dictionary form (e.g., *running* → *run*), consolidating variants into a single feature, and
- Handling of special characters and punctuation.

Once preprocessed, the text can be represented depending on the analytic approach: as sparse, high-dimensional frequency vectors for bag-of-words; as sequences of word pairs or triples for n-gram models; as dense semantic vectors for static and contextual embeddings; as structured prompts for LLMs; or as document-term matrices for topic modeling. In all cases, the quality and interpretability of downstream results depend heavily on the choices made during preprocessing, underscoring the importance of this stage in the analytic pipeline.

Bag-of-words approaches & TF-IDF

Bag-of-words (BoW) models are fundamental approaches to representing text, converting documents into unordered collections of words and quantifying them through simple frequency counts or binary presence indicators. In this representation, word order is ignored, *the cat sat* and *sat the cat* are treated as identical, allowing for seamless statistical processing at the cost of losing syntactic and sequential information.

A closely related family of approaches uses predefined dictionaries or lexicons to map words to psychologically or semantically meaningful categories. For example, the Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2003) uses an extensive, hand-crafted dictionary to match words in a text to psychologically meaningful categories. These categories are organized into broad domains such as emotional processes (e.g., *happy*, *angry*), cognitive processes (e.g., *think*, *because*), and social processes (e.g., *friend*, *talk*), as well as more specific subcategories like positive and negative emotion or first-person singular pronouns. The software scans the text word by word, compares each to its dictionary, and then outputs the proportion of total words that fall into each category, producing a psychologically interpretable profile of the text. Researchers can also create custom dictionaries tailored to specific domains or research questions, such as political rhetoric, health communication, or cultural identity (Benoit et al., 2018; Rice & Zorn, 2021). These dictionary-based methods leverage human-curated semantic groupings to provide interpretable measures that can be directly linked to theoretical constructs.

In both frequency-based and dictionary-based BoW models, the text is transformed into a sparse, multi-dimensional vector, with each dimension corresponding to a unique word or dictionary category. This simplicity makes BoW approaches computationally efficient and easy to implement, particularly for small or domain-specific datasets. However, their disregard for

word order and context means they cannot capture multi-word expressions, subtle semantic shifts, or long-range dependencies, which are limitations that models, such as n-gram language models, incorporate with short-range sequential information.

N-gram models

While bag-of-words models treat each word as an independent unit, n-gram models extend this approach by representing short sequences of adjacent words, most commonly bigrams (two-word sequences) and trigrams (three-word sequences). This allows them to capture local ordering information and short-range dependencies that single-word models overlook. For example, the bigram *climate change* is treated as a single feature distinct from the separate words *climate* and *change*, enabling the model to better preserve meaning in fixed expressions and collocations.

The idea of modeling text through local sequences dates to foundational work on information theory (Shannon, 1948), in which the entropy of English is estimated by predicting the next symbol in a sequence and seeing how much uncertainty remained as the context length grew. This approach laid the groundwork for modern statistical language modeling by formalizing the Markov assumption, in which the probability of a word depends only on a fixed number of preceding words (its $n - 1$ history), rather than the entire sequence.

In practice, n-gram features are represented as sparse frequency counts in a high-dimensional vector space, with each unique sequence assigned its own position in the feature set. This representation can substantially improve performance on language modeling, text classification, and speech recognition. However, the approach scales poorly: the number of possible n-grams grows exponentially with n, and most sequences occur infrequently, leading to sparsity and data inefficiency.

To address these challenges, implementations typically limit n to small values (e.g., 2 or 3) and use smoothing techniques to reassign some probability mass from frequent sequences to unseen ones. N-gram language models are often pruned to reduce size and improve efficiency, by retaining only those n-grams whose counts exceed a specified threshold or by applying entropy-based criteria to remove less informative sequences (Stolcke et al., 1998). While n-gram models excel at modeling short, rigid patterns, they are limited in their ability to capture long-range dependencies or generalize to novel combinations of words: capabilities that more advanced distributed representations, such as static and contextual embeddings, are designed to handle.

Static Embeddings

Whereas bag-of-words and n-gram models represent text as sparse frequency counts, either of individual words or fixed-length word sequences, static word embeddings take a more sophisticated approach by mapping each word to a dense, low-dimensional vector that captures its semantic relationships to other words (Ethayarajh, 2019; Peters et al., 2018). Bag-of-words and n-gram models treat each word form as independent, meaning that *cat* and *feline* are unrelated unless they co-occur in the same contexts. In contrast, static embeddings are learned from large text corpora so that words appearing in similar contexts are placed close together in a continuous vector space, allowing models to generalize meaning beyond exact word matches.

This shift from discrete counts to continuous vector spaces increases complexity in two ways. First, the representation is dense, every word's meaning is encoded across many dimensions rather than in a single count value, making it possible to capture graded similarity. Second, the embeddings are learned parameters of a predictive model, rather than fixed counts,

which requires training on large datasets with optimization techniques rather than simple frequency tabulation.

A widely used static embedding approach is Word2Vec (Mikolov et al., 2013), which trains shallow neural networks to predict either a target word from its surrounding context (Continuous Bag-of-Words) or surrounding words from a target word (Skip-gram). Another prominent method, GloVe (Pennington et al., 2014), learns embeddings by factorizing a global word co-occurrence matrix, thus combining the advantages of matrix factorization and local context window methods.

While static embeddings successfully encode semantic relationships (e.g., *king* − *man* + *woman* ≈ *queen*), they assign the same vector to every instance of a word, regardless of context. Thus, *bank* in "river bank" and *bank* in "bank loan" share an identical embedding, a limitation addressed by contextual embedding methods discussed in the next section.

Contextual Embeddings

Contextual embeddings represent words so that their meaning changes depending on the words around them (Johnson et al., 2024). Unlike static embeddings, which give every word the same vector no matter where it appears, contextual embeddings create a unique vector each time a word occurs. This means the representation of *bank* in "river bank" will be entirely different from *bank* in "bank loan," because the model interprets the surrounding context before deciding how to represent it.

Widely used contextual embedding models are part of a broader class of systems known as large language models (LLMs). LLMs are artificial intelligence models trained on massive collections of text, from books and news articles to websites, to learn the statistical relationships between words, phrases, and larger language patterns. By analyzing these patterns, LLMs can

represent words, sentences, and documents to reflect their meaning in specific contexts. This training allows them to capture subtle shifts in meaning, long-range relationships between words, and language use that other methods miss. For example, contextual embeddings can understand that “cold” in “cold weather” is about temperature, while “cold” in “she gave me a cold look” is about emotion, producing different vectors for each case.

Examples of prominent LLMs that produce contextual embeddings include BERT (Devlin et al., 2019) and RoBERTa (Liu et al., 2019), which are open-source and often used for tasks such as classification, clustering, and semantic similarity. Because their architectures, training details, and model weights are publicly available, they can be fine-tuned for specialized applications and are widely used in academic and applied NLP research. ELMo (Peters et al., 2018) is another open-source model, historically important as an early contextual embedding approach, though generally outperformed by newer transformer-based models. In contrast, proprietary or API-only models, including the *GPT* family (Radford et al., 2019) and *OpenAI embeddings*, are, for the most part[1], not available for direct download or modification. GPT is best known for generating text, but it can also produce embeddings that capture the meaning of entire sentences or documents. OpenAI embeddings are optimized for measuring semantic similarity and powering applications like search, recommendation systems, and clustering. Although API-only models can limit reproducibility and flexibility, they often deliver strong performance in zero-shot and few-shot scenarios (Brown et al., 2024).

When the model reads a sentence during training, it assigns each word a *vector* (a list of numbers) that represents its meaning in that specific context by looking at all the surrounding words in the sentence and adjusting the vector to reflect how that word is being used at that

[1] At the time of this dissertation.

moment. Once trained, the model can take any sentence you give it and instantly generate these context-specific vectors for each word, or for an entire sentence or document.

Semantic similarity

Contextual embeddings (e.g., with transformer models or LLMs) offer distinct advantages for comparing how individuals frame their identities or experiences, particularly when contrasted with approaches such as LIWC or human coding. While LIWC quantifies language through predefined dictionaries of word categories, LLMs generate high-dimensional vector representations that encode the holistic meaning of participants' responses. This allows capturing the underlying conceptual and contextual similarity between narratives, even when individuals use different words or linguistic styles to express comparable ideas (Cer et al., 2018). Empirical evaluations of sentence embeddings, such as those by Perone et al. (2018), show that models using contextualized embeddings (e.g., ELMo, InferSent) outperform traditional bag-of-words or dictionary-based approaches on semantic textual similarity tasks, highlighting their capacity to capture fine-grained meaning relationships. This is especially valuable in the study of multiracial identity, where individuals may draw on self-descriptions that do not map neatly onto fixed dictionaries or pre-defined coding schemes. By enabling gradient comparisons that move beyond word matches, LLMs provide new insights into how people frame their identities and experiences.

In cultural modeling research, embedding-based measures of similarity have allowed researchers to explore whether and how participants frame values, identities, and experiences in convergence within or between groups. For example, large language models were used to compare personal value narratives across participants, identifying patterns of alignment in language that reflected shared value systems (Benkler et al., 2022). Similarly, embedding-based

similarity has been applied to compare the moral lessons generated by large language models for narratives from different cultural backgrounds, revealing cross-cultural patterns in moral themes (Hobson et al., 2024). LLMs allow researchers to compute pairwise cosine similarity scores between participants' narratives, which can then be associated with other variables of interest, such as group membership, well-being, identity integration, or ideological attitudes, using statistical models. Cosine similarity compares the *direction* of two narratives in a multi-dimensional space, like measuring how close the arrows representing each narrative are pointing.

Each narrative is a point in a very high-dimensional space, not just 2D or 3D like a map, but hundreds or even thousands of dimensions. If you draw a line from the origin (the zero point in that space) to the point representing the narrative, you get an arrow (vector) that points in a particular direction. That direction encodes the overall meaning of the narrative. Two narratives that are semantically similar will have arrows pointing in nearly the same direction; narratives with very different meanings will have arrows pointing in very different directions. Cosine similarity measures the closeness of those directions, ignoring differences in arrow length (which might just reflect how "long" or "dense" a narrative is).

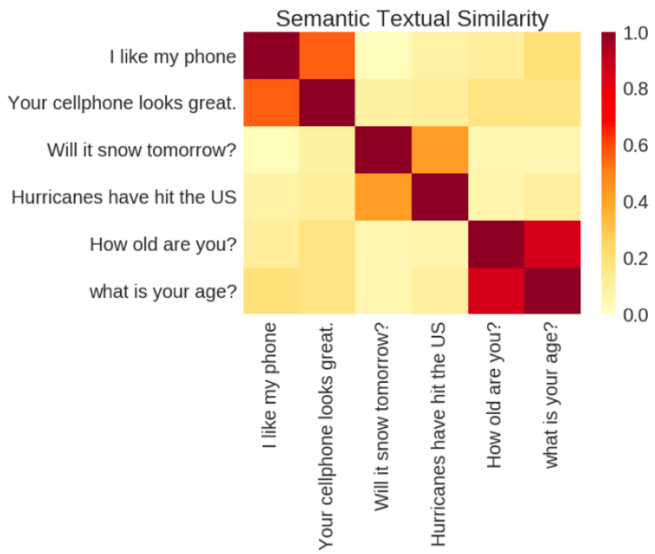
If the arrows point in the same direction, the angle between them is 0° and the cosine value is 1 (perfectly similar); if they point in completely opposite directions (180°), the cosine is -1 (maximally dissimilar); and if they're at right angles (90°), the cosine is 0 (no similarity). Computing pairwise cosine similarity entails representing each narrative as a high-dimensional embedding vector and then applying the formula:

$$\text{Cosine Similarity } (\cos \theta) = (A \cdot B) / (\|A\| * \|B\|)$$

Where $A \cdot B$ is the dot product of the vectors (n = number of embedding dimensions), and $\|A\|$ and $\|B\|$ are their Euclidean norms (magnitudes) of the vectors . The resulting value ranges from

−1 (maximally dissimilar) to 1 (identical), with 0 indicating no similarity, thereby quantifying the semantic alignment between participants' narratives in the embedding space. Figure 1 from Cer et al., (2018) illustrates semantic similarity between text. Texts around age, with different ways of framing the identical question, are more similar to each other than the phone and weather-related sentences.

Figure 1. Semantic textual similarity between sentences



The ability to generate continuous measures of linguistic proximity enables the discovery of subtle patterns: for example, individuals may frame their identity in similar ways as those living in the same geographic region or with the ethnic/racial background. Embedding-based approaches thus offer unique insight into meaning-making across participants.

Sentiment Analysis

Meaning-making is also about how people emotionally frame their experiences. While semantic similarity reveals patterns of alignment in language, the affective tone or valence, offers complementary insight into the emotional significance individuals assign to identity-

relevant events. Analyses of valence in personal narratives can illuminate individual well-being and broader cultural meaning systems that shape how identity-relevant experiences are framed (Lilgendahl et al., 2018). Recent advances in LLMs have made valence analysis more accessible and accurate across languages and cultural contexts. RoBERTa (Liu et al., 2019), a robustly optimized version of BERT, has achieved state-of-the-art performance on sentiment tasks such as SST-2[2], emphasizing that transformer-based models fine-tuned on labeled sentiment datasets can effectively capture evaluative tone. Similarly, GPT-3.5 and GPT-4 outperform traditional dictionary-based methods and match or exceed older machine learning models in detecting sentiment, discrete emotions, and offensiveness across twelve languages, including under-resourced ones (Rathje et al., 2024). While dictionary-based methods are valuable for applying theory-linked constructs in a consistent way, they are limited to pre-specified word lists and cannot account for linguistic variability, polysemy, or context-dependent meaning. Schaaff et al. (2023) found that GPT-3.5 can accurately generate and align emotional tone with user input, producing valence-consistent rephrasings and responses in both controlled and interactive settings. These advances suggest that LLM-based valence analysis could serve as a powerful tool for exploring narratives of identity, though attention to potential biases and careful validation remain essential (Burnham, 2024).

Topic Modeling

While LLMs are beneficial in their capacity to derive semantic meaning from words and position them together, a limitation is the potential for biases found across the large corpora

[2] Derived from the original Stanford Sentiment Treebank (Socher et al., 2013), SST-2 dataset is a widely used benchmark for binary sentiment classification, consisting of individual movie review sentences labeled as either positive or negative in sentiment. SST-2 is part of the General Language Understanding Evaluation (GLUE) benchmark and is commonly used to evaluate and fine-tune natural language understanding models.

(Caliskan et al., 2017; Charlesworth et al., 2021; Defranza et al., 2020). For instance, if textual data overrepresents associations between leadership and masculinity, or between criminality and certain racial groups, an LLM trained on that data may reproduce these biased associations when generating embeddings or predictions, effectively mirroring the implicit human biases embedded in the corpus. Topic modeling, on the other hand, is derived from the data itself (not an external source), and not subject to biases the content LLMs are trained on. Topic modeling works by identifying clusters of co-occurring words that represent underlying topics (Wallach, 2006). This method helps researchers understand the themes present in a large set of texts and how these themes relate to each other and to other variables of interest. As a caveat, Topic modeling is a probabilistic model that does not account for complex linguistic relationships, like word order or nuanced semantics, as LLM does.

Topic modeling can be compared to factor analysis, a technique more familiar to psychologists. Just as factor analysis identifies latent factors that explain the patterns in a set of observed variables, topic modeling identifies latent topics that explain the patterns in a set of words across multiple documents. Each topic is characterized by a group of words that frequently appear together. The researcher then interprets these groups of words to assign meaningful labels to the topics. Topic modeling provides two key insights: (1) it identifies a coherent set of topics that can summarize the entire collection of narratives and (2) it estimates the extent to which each topic is represented in any given narrative. This method allows for a nuanced analysis of textual data, making it possible to explore the relationship between identified topics and other variables, such as psychological constructs or demographic characteristics. By incorporating topic modeling into our toolkit, we can gain deeper insights from qualitative data.

Topic modeling is considered an open-vocabulary analysis, meaning that topics are estimated from the narratives themselves rather than an external, preexisting dictionary (e.g., Linguistic Inquiry and Word Count [LIWC]; Tausczik & Pennebaker, 2010). In short, dictionary methods define latent topics a priori by assigning words to topic lists (e.g., a family list might contain words like "mom," "brother," and "cousin"). Topic modeling, by contrast, estimates topics via the observed clustering of words. A benefit is that, in contrast to pre-existing dictionaries, the open-vocabulary approach honors the words of the participants themselves; in other words, the open-vocabulary approach centers the language and experiences of marginalized groups.

Among available topic modeling algorithms, structural topic modeling (STM; Roberts et al., 2014) uniquely allows for the integration of covariates into the estimation of the topics and thus may be of particular interest to personality psychologists. Within narrative psychology, STM can be used to connect human coding and topic modeling methods. For example, a researcher may code a set of narratives for previously validated themes—such as positive tone—and integrate these scores as covariates in an STM analysis. This allows the researcher to identify common content topics that arise from a specific narrative prompt and to determine whether some topics are more likely to include positive tone. An additional feature of STM analysis is the opportunity to examine the relationship of covariates to topic vocabulary. For example, a study of decision-making found that when discussing the concept of intuition, men were more likely to use the words "interest," "gamble," and "certain," while women were more likely to use "god," "middle," and "doubt" (Roberts et al., 2014). Similarly, we can examine what words multiracial individuals are likely to use across topics compared to monoracial individuals. STM, like many forms of topic modeling, uses Bayesian estimation to derive its solution. This means that STM

estimates the probability that each word belongs to each topic, and the probability that each document expresses each topic. Rather than producing a single fixed solution, STM provides distributions that quantify uncertainty in these assignments, allowing researchers to assess how confidently topics and words are associated while incorporating covariate effects.

Overall, topic modeling is ideal for questions that involve extracting content and gaining a descriptive understanding of the data at hand. Topic modeling is well suited to summarize both broad (such as "Gene," "Evolution," and "Computer," for journal topics in Science; (Blei, 2012) and narrow topics (such as "landing gear issues," "engine, oil pressure," etc., in aviation incident reports; Kuhn, 2018) and to quantify the relationship between narrator and narrative factors to these topics (Roberts et al., 2014). Topic modeling has been used to describe documents as short as tweets (Sterling et al., 2019) and as long as Pitchfork reviews (Light & Odden, 2017), making them suitable for narratives of varying lengths. Experts have not reached a consensus on the necessary number or length of documents for topic modeling. In general, we may assume that more narratives are needed when short in length. One study (using a different topic modeling algorithm known as Latent Dirichlet Allocation) found that the smallest number of essays needed to accurately recover the model parameters was 55, given a small set of topics (2–3) and essays of varying lengths (200, 250, and 500 words; Wiberg et al., 2021).

The evolution of methods for analyzing identity ranges from the interpretive richness of human coding to the efficiency and scalability of computational approaches. Human coding offers unparalleled strengths in identifying theory-informed themes, where the positionality and theoretical sensitivity of the researcher are critical. At the same time, advances in computational text analysis, from topic modeling to large language models, provide powerful tools for systematically capturing content patterns in narrative data at scale. These approaches

complement human coding by enabling analyses that measure emergent structures and linguistic proximities across large datasets. Yet, as this dissertation argues, the decision to apply human coding or automation reflects deeper epistemological commitments about what constitutes meaning and how best to study identity. By integrating these methods thoughtfully, both the complexity of individual meaning-making and the potential of computational tools can reveal new insights into how people frame their identities and experiences.

Dissertation Aims

Existing theoretical models may not capture the experiences of more diverse multiracial backgrounds; and qualitative approaches offer a powerful lens for examining identity development in ways that purely quantitative methods may overlook. In this dissertation, I aim to describe and understand identity related to ethnicity/race and, specifically, being multiracial. I will use accounts provided by participants in the form of (1) race-related narratives, (2) self-concepts organized via the Twenty Statements Test, and (3) open-ended reflections on strategies for navigating multiracial identity and rites of passage where those strategies are applied.

Based on the range of qualitative methods available, Thematic Analysis (TA) and Narrative Analysis (NA) appropriately address the research questions in this dissertation. TA offers a systematic way to identify and interpret recurring patterns across participants' accounts, balancing inductive with deductive approaches to theoretically inform constructs. NA focuses on the content, sequencing, and evaluative framing of personal stories, focusing on how individuals organize experiences into coherent identity trajectories. Narratives are particularly well suited for highlighting the experiences of those who have been overlooked, ignored, or actively silenced (Fish & Syed, 2021). While these are the primary methods employed, the analytic approach draws on elements from other qualitative traditions, the idiographic depth and lived-experience

focus of Interpretative Phenomenological Analysis, the attention to context and language use from Discourse Analysis, and the grounded, iterative engagement of Grounded Theory. This combination allows for both cross-case pattern mapping and deep, situated interpretation of identity work, aligning methodological choice with the aim of capturing how multiracial individuals construct and navigate their identities within sociocultural contexts.

The computational approach applies large language models (LLMs) and related text representation methods to trace the semantic, structural, and thematic contours of multiracial identity narratives at a scale that complements and extends the interpretive reach of qualitative analysis**.** Contextual embeddings represent each narrative in a high-dimensional semantic space, allowing for precise similarity analyses that detect conceptual alignment between accounts, even when those accounts differ in vocabulary, phrasing, or style. For this purpose, I use GPT-based models to capture semantic similarity between narratives, leveraging the model's ability to encode context and culturally specific language, and for sentiment and valence detection to quantify the evaluative tone participants assign to their identity experiences. Next, topic modeling identifies clusters of co-occurring words that signal underlying themes, offering an open-vocabulary, data-driven lens that can surface patterns beyond those anticipated in theory or detected through human coding. Lastly, a customized dictionary approach**,** developed around identity-relevant constructs, produces theory-linked measures that remain interpretable while being consistently applied across the dataset. Together, these techniques enable the detection and comparison of language patterns across subgroups, such as dual-minority and majority-minority multiracial individuals, while preserving the variability in how participants frame and express their identities. Computational methods for analyzing text offer a new toolkit for qualitative

research. *This dissertation will identify individual differences in identity development of multiracial people using both qualitative and computational methods.*

In this dissertation, I investigate how multiracial individuals actively engage in meaning-making to construct, organize, and navigate their racial identities. Rather than treating identity development as a passive response to context, I take a cultural psychological approach that emphasizes individuals as agents who internalize, reinterpret, and creatively reshape cultural material in the process of becoming themselves. Specifically, I aim to understand how multiracial people make meaning of their experiences, define who they are in their own words, and apply strategies, both broad and situated, for navigating their identities in racialized social worlds.

To pursue this objective, I examine the lived experience of multiracial individuals across three interrelated domains:

1. **Agency and communion in ethnicity/race-related experiences**

   The first study explores in general how individuals make meaning of their ethnicity/race-related experiences, with a particular focus on the themes of *agency* (autonomy, resistance, self-definition) and *communion* (connection, harmony, relational belonging). Rather than centering external appraisals or demographic contexts, this study foregrounds how participants themselves interpret moments of racial salience and how those interpretations reflect internalized schemas. This is published in *Journal of Personality* (Jacobson, Westberg, Chou, Syed, Weston, 2024).

2. **Framing the self in their own words**

   The second study investigates how multiracial individuals spontaneously define themselves through open-ended descriptors. This approach allows for the emergence of

hybrid, contradictory, and culturally specific representations of the self, without imposing fixed categories. Analyzing the structure and content of these self-definitions sheds light on how identities are internalized at the level of identity organization. This work is under review (Jacobson, Benet-Martínez, Weston, under review).

3. **Strategic navigation of multiracial identity**

   The third study examines the strategies multiracial individuals use to navigate their identities—both *globally* (across time) and *locally* (within a single, meaningful, liminal life experience engaging their multiracial identity). These strategies may include practices such as code-switching, reframing, compartmentalization, or bridging. This study highlights the embodied and context-responsive ways individuals engage with the demands of their cultural environments. This work is in prep (Jacobson & Weston, in prep).

These studies provide an account of how multiracial individuals construct identity at the intersection of internalized culture and lived experience. The dissertation offers an empirically grounded, culturally attuned account of multiracial identity development, one that centers the person as a meaning-making agent embedded in their sociocultural world.

CHAPTER II

STUDY 1: Using Structural Topic Modeling to Understand Ethnicity/Race-Related Narratives

The data used in this chapter was collected by co-author, Dr. Moin Syed. Co-authors, Dr. Dulce Westberg and Edward Chou, helped train and supervise the raters for narrative coding alongside me. I conceptualized the project, performed the analyses, and wrote the manuscript, with helpful feedback from all co-authors.

Introduction

Research on identity development in general, and ethnic/racial identity (ERI) in particular, has long emphasized the processes involved in developing a sense of identity, or the affective, behavioral, and cognitive strategies individuals use to better understand who they are and how they fit in the world (Galliher et al., 2017). In contrast, the *content* of ERI – the actual experiences of what it means to hold a particular identity – has received less attention, even though variations in such contents are important for understanding the processes (Syed, 2015).

One challenge of understanding ERI content is that this content is highly contextualized and not subject to any existing taxonomies and thus cannot be reliably assessed using rating-scale instruments. A solution is to use narratives of people's ethnicity- and race-related experiences, through which respondents provide contextualized details about their lives while recounting their personal past (Syed, 2015). These narratives are almost always human-coded, either using an existing protocol or inductively based on the data (Adler et al., 2017). This approach presents two challenges. Pragmatically, human-coding narratives is a difficult and time-consuming process. More substantively, the subjectivities associated with human-coding may lead to over-emphasizing some aspects of narratives or over-looking others. Recent

advances in natural language processing allow us to efficiently capture topics within narratives using a bottom-up approach.

In Study 1, I applied Structural Topic Modeling (STM; Roberts et al., 2013) to narratives of ethnicity- and race-related experiences among both multiracial and monoracial participants to better understand the content diversity of ERI. Specifically, I sought to identify how the contours of these experiences differ between multiracial and monoracial groups, capturing how individuals with multiple heritage backgrounds articulate belonging, difference, and meaning. I then examined the association between the prevalence of the identified topics and human-coded themes of agency and communion within the same narratives. Agency and communion are two motivational themes implicated in identity development (McAdams et al., 1996) and are considered relevant for navigating racialized and ethnicity-related experiences and expectations (Syed & McLean, 2022). Through this study, I demonstrate how integrating human-coded motivational themes with computationally derived topics can provide a richer, comparative description of how multiracial and monoracial individuals construct and emotionally frame their ethnicity/race-related experiences.

The Intersection of STM and Narrative Identity

STM analysis can help us understand the nuances of identity narratives, including those focused on ethnicity and race. For multiracial individuals, these narratives capture how multiple cultural frameworks are negotiated simultaneously, revealing how racial ambiguity or hybridity shapes meaning-making. For monoracial individuals, they illuminate how cultural continuity and group belonging are experienced within established racial boundaries.

Ethnicity/race is a salient social category in the United States (U.S.; Chae et al., 2021; Williams et al., 2020), therefore, capturing ethnicity- and race-related experiences can reveal

individuals' subjective beliefs about their place in the larger society. Past work, using human coders, identified content themes of connection to culture, awareness of cultural differences, discrimination, and underrepresentation in ethnic identity narratives (Syed & Azmitia, 2008, 2010). Moreover, narrative content varies based on ethnic group membership (Westberg, 2022; Yip et al., 2021).

For example, Syed and Azmitia (2008) found that multiracial participants were more likely than monoracial participants to tell narratives about awareness of difference and connection to culture. Relatedly, Mitchell et al. (2018) found Latine participants were more likely than other ethnicities to describe spirituality as a reason for non-typical ethnic/racial experiences. These findings suggest that multiracial narratives are especially likely to blend themes of difference and belonging, while monoracial narratives may center cultural continuity and affirmation. As evident from these examples, the content of ethnicity- and race-related experiences may vary based on ethnic group membership and represent a distinct aspect of identity. By applying STM across both groups, the current study identifies not only shared identity content but also topics uniquely prevalent in multiracial narratives.

Accordingly, *the first aim of this study was to identify the major topics extracted by STM across participants' ethnicity- and race-related narratives and to examine how these topics differ between multiracial and monoracial individuals*.

Even when a topic is equally present in narratives across groups, the words used to describe that topic may differ. This is especially important when considering how participants come from different ethnic and racial backgrounds, as words reflect the larger semantic systems of a narrator's culture (Thompson et al., 2020). Such differences in topic vocabulary provide additional nuance to the understanding of narratives. For example, when using an online peer-

support platform for mental health, Asian American participants used fewer negative words compared to White, Black/African American, and Latine participants (Loveys et al., 2018). To date, we are aware of little research that explicitly compares vocabulary within identity narratives and across ethnic groups. However – given prior work demonstrating the importance of even pronoun choice for revealing social standing (e.g., Kacewicz et al., 2014) and linguistic code-switching as a means of asserting identity (Loyd et al., 2023)– topic vocabulary is likely to hold important clues about ethnic identity. *Thus, the second aim was to examine how word choice and topic vocabulary differ across multiracial and monoracial groups, revealing how linguistic framing itself reflects culturally grounded meaning-making and identity negotiation.*

Independent of ethnic background, a narrator's psychological profile may be gleaned from their narrative, primarily through coding for motivational themes. Two motivational themes, communion and agency, capture the need for connection to a community and establishing one's purpose, respectively (McAdams et al., 1996). These themes play a crucial role in shaping one's sense of self because they act as overarching motivational frameworks for structuring ethnicity- and race-related experiences in the context of balancing group expectations and personal aspirations (Bakan, 1966; Syed & McLean, 2022).

For example, a person's sense of who they are can be based on their group memberships if they actively reflect on their social groups (Tajfel & Turner, 1979). In this case, communion-related constructs—ethnic group belonging and cultural connectedness—positively affect psychosocial outcomes by buffering experiences of discrimination (Brondolo et al., 2009) and improving well-being (Webber & Waru-Benson, 2022). Agency, on the other hand, enables individuals to adhere to their ethnic group's norms and values and to redefine what their ethnicity means to them on a personal level. For example, students who are more individuated from, yet

value attachment to, their family had higher college adjustment compared to those who did not (Kalsner & Pistole, 2003). These goals seem on the surface oppositional, yet they are interdependent, particularly within the context of people's experiences involving ethnicity (i.e., people can be both agentic and communal; (McCabe & Dinh, 2016).

Agency and communion can co-occur within the same ethnicity-related narrative (McLean & Syed, 2015). For example, an experience can be communal and agentic when the participant describes feeling cared for and accepted by other ethnic group members. Identifying ethnicity-related instances that empower and connect an individual is important for highlighting positive experiences to be fostered by communities. In contrast, other research suggests that accomplishment (agency-related) is achieved at the expense of communion (McCabe & Dinh, 2016). Understanding how multiracial and monoracial participants differ in how agency and communion co-occur can reveal distinct motivational structures of identity: whether multiracial individuals narrate empowerment through boundary-crossing and negotiation, while monoracial individuals do so through affirmation within group boundaries. *The third aim of this study was to examine how the motivational framing of agency and communion relates to topic prevalence within ethnicity- and race-related narratives, and whether these relationships differ between multiracial and monoracial individuals.* This comparison helps clarify the contexts in which participants feel valued, included, or marginalized, and how those emotional dynamics manifest differently across identity structures.

The Present Study

I led this present study where, in collaboration with researchers at University of Minnesota, we integrated STM with a narrative identity approach to examine the content of ethnicity- and race-related narratives across multiracial and monoracial individuals, linking

computationally derived topics to the motivational dimensions of agency and communion. In so doing, the methodological contribution is to introduce an additional, efficient way of analyzing content in narrative data, especially in larger quantities. In contrast, the theoretical contribution is to further elucidate the kinds of ethnicity- and race-related experiences individuals navigate, highlighting how those with multiple racial heritages organize identity meaning differently from those with a single heritage.

Specifically, I first use STM to extract topics representing the content of ethnicity- and race-related narratives (Aim 1). Next, to determine the extent to which topics derived from STM are overlapping or different across ethnic/racial groups, I examine differences in topics and vocabulary between multiracial and monoracial individuals (Aim 2). Lastly, I explore relations between topics and human-coded themes of agency and communion, testing whether these motivational dynamics operate similarly or differently across groups (Aim 3).

Method

In this study, I used existing data collected by Dr. Moin Syed between 2005 and 2012.[3] The methods, materials, exclusion criteria, and analysis plan of this study were preregistered before analyzing the data (https://osf.io/tbmd3). Ethical approval was obtained from the Institutional Review Board (IRB) at each data collection institution. This study was published in *Journal of Personality* (Jacobson et al., 2024).

Participants and Procedure

Participants were 1,184 college students recruited from psychology department participant pools at two large universities in California and one large university in the Midwest.

[3] Portions of these data have been published previously to examine a subset of the narrative data (Syed & Azmitia, 2008), as well as ethnic group typicality and theoretical models of ethnic identity related to identity coherence and psychological functioning (Mitchell et al., 2018; Syed & Juang, 2014).

Participants completed a 50–60-minute survey in a research lab on campus in exchange for course credit. From this sample, only participants who responded to the relevant narrative prompt were retained for analysis, resulting in a final sample size of 1,149 participants (M_{age} = 21.14, range = 18-57; 64% women, 35% men, 0.5% non-binary, and 0.5% unspecified).

For ethnic/racial identification, participants reported their self-identified ethnic group membership using both close-ended and open-ended formats. The close-ended responses were of seven broad ethnic/racial categories including Asian American, Black/African American, Hispanic/Latinx, White, Native American/Indigenous, Multiracial[4], and Other/Non-categorizable. Participants who did not select a close-ended ethnic category were labeled 'non-categorizable.'[5] Within our sample, 42% participants identified as White, 31% as Asian American, 8% Latinx, 5% Multiracial, 4% Black, 1.6% as Native American/Indigenous, and 6.5% as Other/Non-categorizable. In recognizing the inherent challenges in all classification systems, including ours, we adopted a categorization strategy for participant ethnicity/race when their close-ended was left blank and their open-ended was not to ensure inclusivity and to better reflect the diversity of identities who are part of these pan-ethnic groups. See the supplemental materials (Table in Appendix A) for a table with self-identified ethnic group membership by pan-ethnic group.

Measures

Ethnicity-Related Narrative Prompt

[4] Participants who selected more than one category or self-identified as Multiracial were labeled as 'Multiracial.'

[5] Within this category, I manually coded an open-ended item asking these participants to describe their ethnicity/race into one of the six aforementioned ethnic groups. For example, participants in this subset who identified as "Irish" and nothing else were categorized as "White." I manually coded 27 out of the 1149 total participants. The breakdown of these participants is the following: 8 White, 11 Asian/Asian American, 6 Latinx, 1 Black/African American, and 1 Multiracial. Still, 6.5% were not able to be categorized based on their responses and made up the Other or 'non-categorizable' category.

Participants were asked to respond to the following narrative prompt:

> *Please describe a particular time in your life, either positive or negative, that you felt aware of your race/ethnicity when you were hanging out with one or more of your close friends. Tell us how old you were when this happened, where you were, whom you were with, what happened, how you reacted. Include details that would help us see and feel as you did.*

Participants were asked to consider memories that included a close friend or family member to elicit salient experiences that involved other people. Narratives were 145 words on average (range = 4 – 628 words). One narrative (0.09%) was shorter than 5 words, 3 (0.26%) were shorter than 10 words, and 20 (1.7%) were shorter than 25 words.

Human Coded Agency and Communion

Narratives were coded for themes of agency and communion by three teams of two trained undergraduate research assistants (six coders total). Members of each team coded all narratives randomly assigned to that team; in other words, each team coded 1/3 of the narratives from the total sample, resulting in approximately 383 narratives assigned to each coder (1,184 narratives total). Coders were trained to rate themes of agency and communion using coding systems established by past researchers (McLean et al., 2019; https://osf.io/d2xts). First, coders familiarized themselves with the narrative identity approach by reading McAdams (1996), which outlined themes of agency and communion. Then, coders rated ethnicity narratives for agency (e.g., autonomy, self-mastery, control) and communion (e.g., connection, unity, togetherness) on two separate five-point scales ranging from 0 to 3 (0 = low agency/communion, 1.5 = does not contain agency/communion, 3 = high agency/communion; this was incremented as 0, 1, 1.5, 2, 3 for each theme). Since not all participants followed the instructions to describe an experience

with a close friend or family member, mention of a close friend or family member was not necessarily indicative of higher scores for communion.

Coding for each team was overseen by the first, second, and third author, respectively. In line with conventions established in past research (Adler et al., 2017; Syed & Nelson, 2015), coders met weekly with the author leading their team, during which time they were given the chance to change or maintain discrepant codes. Weekly meetings fostered reliability and validity, allowing coders to share their thoughts, and come to an agreed-upon understanding of how agency and communion manifested within narratives. In addition, the first three authors convened in a biweekly meeting separate from the coders to discuss common problems and ensure the coding process was standardized across teams. Using benchmarks suggested by Koo & Li (2016), I found moderate interrater agreement after computing the intraclass correlation coefficient across the three pairs of raters for agency (ICC = 0.71, 0.74, 0.90) and good interrater agreement for communion (ICC = 0.82, 0.75, 0.87) using two-way random effects models to assess absolute agreement (one model for each of three sets of raters and two themes).

Analysis Plan

Aim 1: Identify Ethnicity-Related Topics and Label Using Top Words and Example Narratives.

Topics were extracted using STM, a technique that estimates topics based on the distribution of words in a corpus of text (Roberts et al., 2019). Analyses were conducted in R (version 4.1.2; R Core Team, 2021) using the packages *stm* (Roberts et al., 2019) and *tm* (Feinerer, 2020). Topic modeling determines the likelihood of each topic in each document (narrative) on a probabilistic scale from 0 to 1. Agency and communion were included as prevalence covariates, meaning prevalence of topics is freed to vary across narratives as a function of these human coded themes. Estimates of these relationships – between topic

prevalence and human coded themes – can be easily extracted and statistically tested. I also included ethnicity as a content covariate, meaning that the rate of word use within a topic is freed to vary across ethnic groups. While prevalence covariates are well-suited for research questions like "Is topic 4 highly associated with emotional tone?", content covariates meaningfully examine questions like "Which words in topic 4 are exclusive to members of positive versus negative tone?"

Text data were preprocessed by removing stop words (e.g., 'the,' 'and'), numbers, special characters (e.g., '!'), and low prevalence words (appears once; Denny & Spirling, 2017).[6] Next, we determined the number of topics to extract by comparing models with varying numbers of topics (e.g., 2, 3, 4, etc.) on their model fit. I used four key metrics to compare model fit: (1) exclusivity[7], (2) variational lower-bound, (3) residuals, and (4) semantic coherence. These are recommended from prior literature (Fu et al., 2021 ; Weston et al., 2023). Exclusive words are those that are uniquely associated with a single topic; the exclusivity metric indexes the degree to which the words strongly associated with topics are exclusive to their topics. High exclusivity indicates each topic is clear and distinct from one another. Variational lower-bound indexes convergence. The estimation procedure of STM iterates until the variational lower bound is below a set threshold. All else being equal, lower values of variational lower bound indicate better fit. Residuals in this instance specifically refer to the dispersion of residuals. Recommendations for the use of this metric vary – some suggest looking for local minima, while others claim that residual dispersion lower than 1 is sufficient to indicate good fit. Lastly, semantic coherence indexes the degree to which high-probability words within a topic co-occur.

[6] This code is available on the OSF Repository: https://osf.io/7krxn/

[7] I initially preregistered to use held-out likelihood as a metric and instead used exclusivity based on recommendations from Weston et al. (2023).

This index has been found to positively correlate highly with human ratings of the coherence of topics, although there is some question as to the validity of this inference (Koltcov et al., 2019). Each of these metrics is limited in this capacity to identify a good-fitting and interpretable model. However, taken together, these four metrics can be used to identify a tractable number of candidate models for deeper investigation.

Frequently, these four metrics indicate that more than one model may provide sufficient fit to the data.[8] In those cases, I then compare estimated models in terms of topic overlap (i.e., how redundant the models are) and topic prevalence. Prevalence refers to the proportion of words within a single narrative that come from the topic. Average topic prevalence across all narratives provides an index of the proportion of words in the entire corpus that come from the topic. Topics which are especially rare (low prevalence), have low utility, as they appear in few narratives and may not generalize beyond a few key examples (Roberts et al., 2019). Overall, this comparison of multiple potential topic models allows us to determine the best set of models with little overlap and high utility. Four metrics were calculated across STM models from 3 to 20 topics. Once the narrowed number of topics were agreed upon, my co-authors and I estimated, labeled, and compared the candidate models.

I reported one optimal topic model in the final manuscript. Focusing on one model allowed for a more in-depth interpretation and optimization of the model's parameters and results. This approach facilitated a deeper understanding of the underlying themes, making it easier to draw meaningful insights on the relationship between ethnicity experiences and agency and communion across ethnicities. The choice between a parsimonious model and a more

[8] This is not dissimilar to exploratory factor analysis, in which multiple numbers of factors are supported by the fit statistics. In both cases (STM and EFA), researchers subjectively evaluate the solutions to determine whether one or several provide theoretical insight and potential application.

descriptive one hinges on the research objectives. A simpler model with few topics is advantageous when the researcher aims to replicate the model to various datasets; it distills information into broader themes that are less likely to be dataset-specific. On the contrary, a model with more topics may be ideal when the goal is to fully characterize a particular dataset. This decision is much like the choice of the number of factors in a factor analysis. Broad factors are more likely to replicate across datasets but may lack predictive power compared to narrow factors or items and may overly aggregate similar but not identical constructs.

To identify robust topics found across the models and determine which topics are unique, I examined the overlap in topics across models with a correlation matrix between the two and computed prevalence (the mean of each topic across all narratives). For robust topics, I looked for topics highly associated with one another (i.e., Topic 13 in the 14-topic model and Topic 1 in the 5-topic model). For unique topics, I looked for topics of high prevalence that did not overlap with topics across models (i.e., Topic 5 in the 14-topic model). This approach added conceptual value in that models with more topics have the benefit of capturing more granular and specific details but the limitation of including more topics that only describe a small set of narratives.

Among the STM output are two matrices. The first, the *beta matrix*, contains the term-topic probabilities. These represent the likelihood that a given term is produced by a given topic. The second, the *gamma matrix*, contains topic-document probabilities, which represent the share of a given document (narrative) that contains each topic. Together, these matrices can be used to evaluate the content of topics (from the beta matrix), the prevalence of topics (from the gamma matrix), and even the association of topic prevalence with narrator- and narrative-specific covariates (by connecting the gamma matrix with other sources of data). From these matrices, ten topic-related keywords and five topic-related narratives were extracted. To label topics, co-

authors independently looked at the ten topic-related keywords and five key narratives from the dataset most likely generated from each topic. Then, a meeting took place during which the authors came to a consensus on topic labels using narrative content and keywords. Strategies for labeling topics are thoroughly described in Weston et al. (2023). I would like to acknowledge that our interpretation of the topics may have been influenced by our backgrounds, experiences, and identities. Our research team represents two graduate students, one post-doctoral researcher, and two university faculty. Regarding ethnicity, we identify as Multiracial (n = 2), Hispanic/Latina (n = 1), White (n = 1), and Asian (n = 1).

Aim 2: Determine Whether the Content of the STM Topics Changes as a Function of Participants' Ethnicity.

To assess whether the content of STM topics changed as a function of participants' ethnicity and racial background (i.e., whether they identified as multiracial or monoracial), I examined self-identified ethnic/racial group as a content covariate. Notably, there is no significance test for these associations; thus, evaluation of Aim 2 was qualitative in nature. Instead, this covariate indicates which words for a given topic were most frequent within and exclusive to each group. In doing so, I qualitatively compared topic prevalence and word usage patterns across multiracial and monoracial participants to identify shared versus distinct narrative themes. For this aim, I provide highly frequent and exclusive words for each topic within each identity group.

Aim 3: Examine the Relationship Between Extracted Topics and Themes of Agency and Communion.

To address Aim 3, I examined the parameter estimates associated with the human-coded agency and communion covariates of topic prevalence. This allowed for the detection of which topics reflect themes of agency and communion. Model output includes coefficient estimates

(and standard errors) that communicate the association of topic prevalence to human-coded agency, communion, and their interaction.

Results

Aim 1: Identify/Label STM Topics

To determine the ideal models, I looked for models with high exclusivity, low residual value, high semantic coherence and high variational lower bound. Based on these standardized metrics, I identified two topic models: a 5-topic model and a 14-topic model (see Figure 2.1). The 14-topic model had lower residual, higher exclusivity, higher variational lower bound than the 5-topic model, indicating higher overall topic quality. As described in the Methods section, each author of this paper independently reviewed and labeled the 5- and 14-topic models based on frequent and exclusive words for each topic, words associated with each topic by ethnicity/race, and five highly-associated narratives, which are strongly linked to the specific topic, as determined by the topic assignment probabilities and the thematic content of the documents themselves. After labeling the topics, I found that the 14-topic model was more nuanced than the 5-topic model in that the topics covered a broader range of concepts and with more specificity. (See Appendix A for more detail about the 5-topic model).

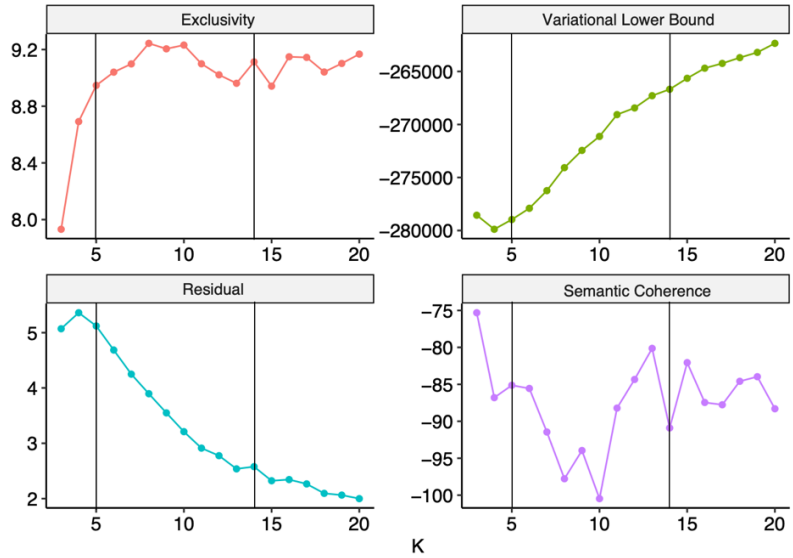


Figure 2.1. The four standardized metrics used to determine candidate models with the 5-topic and 14-topic model highlighted as vertical lines and the number of topics on the X-axis.

After labeling the STM topics, I examined correlations between topics across the two models to determine thematic overlap (see Figure 2.2). Correlations between topics were generally high, average of r = .43, indicating substantial overlap between them. Importantly, no correlations were above .70, indicating that no topic was replicated across the two solutions. Instead, this pattern of results suggests that topics in the 14-topic model likely represent subtopics or more specific instances of topics in the 5-topic model. I observed that the topics labeled *Awareness of cultural difference* in both solutions had a positive association of r = .64. Another example, the topic labeled *Peer dynamics* in the 14-topic model had a positive association of r = .60 with the labeled topic *Intra-intergroup interaction/tension* in the 5-topic model. Overall, the 5-topic model exhibited a strong association with the 14-topic model. Given that both models exhibited a significant degree of topic overlap, I report results from only one model to avoid redundancy in interpreting the similar underlying themes they capture. For the

rest of the manuscript, I describe the 14-topic model for its higher overall topic quality and the degree of overlap across the two models.

Figure 2.2. Correlations between the 14-topic model and the 5-topic model.

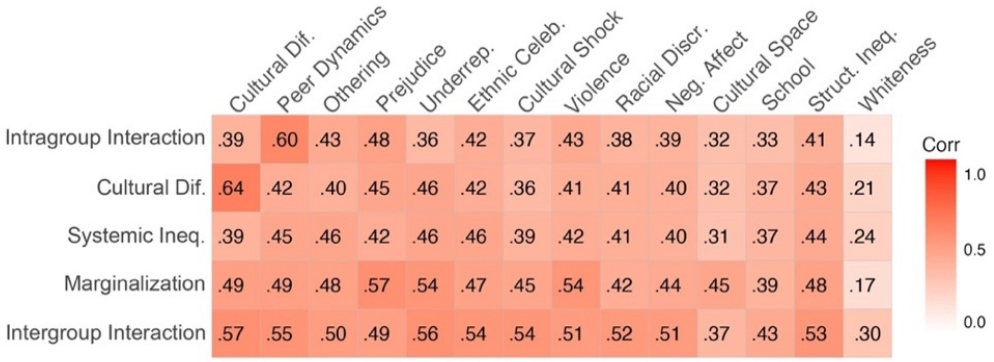


Table 2.1 illustrates the topics in the 14-topic model by prevalence, along with key narratives. After determining the 14-topic model, I explored the ethnicity-related topics. STM topics varied in prevalence, which is the proportion of the word documents associated with each topic. *Awareness of cultural difference* and *Peer dynamics* were the most prevalent across the narratives (see Table 2.1). *Othering* and *Reactance to prejudice* were the third and fourth most prevalent across narratives. The least prevalent STM topics were *Whiteness*, *Gained perspective of structural inequality*, and *Ethnic stereotypes at school*.

I examined bivariate correlations between topics in the 14-topic model to determine whether any co-occurred and found there to be little to no association between topics within the model (the range of correlations from -.13 to.10; see Appendix A for the correlation matrix). In what follows, I detail each topic by reporting the keywords and describing highly associated narratives.

Table 2.1. Labeled STM topics in the 14-topic model with exemplary narratives and prevalence represented as the mean (and median) gamma probability of documents.

Number	Label	Sample Narrative Excerpt	Topic Prevalence
1	Awareness of cultural difference	When I was with other non-native Chinese/Taiwanese teenagers, I realized that my peers were more characterized by who they actually were and not by their ethnicity, unlike in America, where I was - I'm assuming - subconsciously labeled as "Asian."	.14 (.02)
2	Peer dynamics	I felt very conscious of having brown skin where nobody else did and it kind of made me wonder if the kids were trying to make me feel bad or if they were just being annoying.	.14 (.02)
3	Othering	I knew I was different from them but I just never noticed it because they were so friendly and nice to me.	.09 (.01)
4	Reactance to prejudice	I turned around, and even though I wasn't really yelling, I told them off. I said that how could they say that we were considered as stuck up when they too had a table full of ONLY hispanic girls at their table?	.09 (.01)
5	Awareness of underrepresentation	It is a predominately white school with about a quarter of Natives going there like us, two or three black people, one asian kid, so you see, there really isn't a diversity among the student body.	.08 (.01)
6	Ethnic/cultural celebration	The conference was spread over 3 days and included workshops pertaining to the Filipino culture, an entertaining variety show, and a club afterset.	.08 (.01)
7	Experience of cultural shock	Living back home was a huge culture shock. I felt myself so caught up into the American lifestyle, however when i was living back home to where i belonged i felt that a huge part of my life was missing.	.08 (.01)
8	Violence	It turned out that the crazy black girl assaulted the Indian lady very harshly right before me with no reason, too. There were several white guys who saw the whole situation around us, but they did not take any action.	.08 (.01)
9	Contact with racial Discrimination	When we were walking to our car, two African American men were standing close by it. My two friends panicked and made us stall awhile until the two men left.	.06 (.01)
10	Negative affect	I couldn't do anything as an individual because the whole community was scorned upon by a lot of people. Of course, there were people who understood the situation, but there were those who absolutely despised us too.	.05 (.01)
11	Lack of belonging in Cultural space	I brought up the topic of LGBTA people to my mom and she said to me, only Americans have that, we Hmong people do not have LGBTA people, we dont even have a word for it. That really broke my heart. I felt really alone.	.04(.01)
12	Ethnic stereotypes at school	From everyones perception indians are never good at soccer and so no-one really considered putting me on the team.	.03(.00)
13	Gained perspective of structural inequality	Of course the education system was lacking, focus was primarily on reading/writing in school whilst trying to keep up with the other counties and their education systems. In her situation, there was nothing restraining her education.	.02(.00)
14	Whiteness	I proposed the idea that class is more important than race in the city music scene, and my guitarist countered by suggesting that the reason that I booked the shows in our band was due to my race, being that he was black and I was white.	.02(.01)

Note. Sample narratives have not been edited by the authors.

Topic 1: Awareness of Cultural Difference

Topic 1, labeled *Awareness of cultural difference*, describes the interaction of cultures either intra or interpersonally, where the individual learns the significance of their ethnicity.

Compared with Topic 3 (*Othering*), many of the narratives in this topic were neutrally valenced. The highly associated narratives depicted contrasts between Chinese/Taiwanese and American culture. The top words were “else,” “integrated,” “blended,” “lecture,” “others,” and “odd.”

Topic 2: Peer Dynamics

Topic 2, labeled *Peer dynamics*, describes exposure to ethnic/cultural norms from peers. The keywords were “latino,” “mattered,” “sixth,” “saturday,” “defensive,” “cook,” and “friends.” The highly associated narratives for Topic 2 included both positive and negative experiences. The example in Table 1 presents a negative experience where the individual, surrounded by other kids, felt aware of having brown skin.

Topic 3: Othering

Topic 3, which is labeled *Othering,* features narratives focused on feelings of confusion or frustration that stemmed from experiences of oppression, violence, or alienation. The highly associated narratives include racial profiling (being pulled over by the police) and rejection of one’s heritage from a perceived ingroup. The top keywords were “called”, "slurs”, “hit”, “point”, and “throw”. While the representative narratives for Topic 3 reflected feeling othered, not all were due to explicit acts of discrimination or prejudice and were instead sometimes characterized by awareness of difference (see Table 1).

Topic 4: Reactance to Prejudice

Topic 4, labeled *Reactance to prejudice*, features the following top words: “story,” “disappointed,” “furious,” and “roommate.” The key narratives described experiences of prejudice and related reactions, including disappointment and frustration, by an outgroup member. For example, in Table 1, the narrative described prejudice from an outgroup member who the participant deemed hypocritical.

Topic 5: Awareness of Underrepresentation

Topic 5, labeled *Awareness of underrepresentation*, reflected individual feelings of underrepresentation, either numerically and/or by society. The top words were the following: "fought," "people," "verbal," "gifted," and "tastes." In Table 1, the representative narrative includes an individual of Native American heritage who described their school with a non-diverse student body.

Topic 6: Ethnic/Cultural Celebration

Topic 6, labeled *Ethnic/cultural celebration*, involved organized celebrations of one's culture. There was often a feeling of connection, and the highly associated narratives took place in organized cultural events. The top keywords were "performers," "prominent," "dances," "relatives," and "annual." The highly associated narratives took place at a conference held by the association of Filipino-Americans, a Hispanic Scholarship Consortium, a Vietnamese new year community event, a Japanese community center, and a Filipino debutante ball.

Topic 7: Experience of Cultural Shock

Topic 7, labeled *Experience of cultural shock*, reflected unfamiliarity in navigating particular cultural contexts. The representative narratives comprise positive and negative experiences (i.e., "I was very open to meeting and interacting with other ethnicities. I wasn't sure how other ethnicities would respond to me, but I was hoping to make friends" and the excerpt in Table 1, where the person returning to Egypt describes a culture shock in a positive way). The keywords were "american," "roommate," "egypt," "world," "begin," and "complete."

Topic 8: Violence

Topic 8, labeled *Violence*, included experiences of anger and frustration (such as aggressive language and physical altercations). The highly associated words were "street,"

“walking,” “guys,” “innocent,” and “died.” In Table 1, there is a representative narrative, where the participant and friend describe being physically assaulted by a woman.

Topic 9: Contact with Racial Discrimination

Topic 9 was labeled *Contact with racial discrimination* and involved learning about or experiencing racial inequality. This includes narratives that contained mention of either witnessing or receiving discrimination. The highly associated narratives described discrimination against Black/African American participants and were primarily from the perspective of White individuals. The top keywords were “african,” “americans,” “classroom,” “trash,” “crossing,” and “slaves.”

Topic 10: Negative Affect

Topic 10, characterized by its globally negative language, is labeled *Negative affect*. The highly associated words were “unkind”, “upset[i],” “again,” “helping,” and “glendale.” The keywords were more reflective of negative affect than the highly associated narratives. However, an example narrative included a participant of Asian American descent who felt “scorned” and “despised” by members of the White community.

Topic 11: Lack of Belonging in Cultural Space

Topic 11, labeled *Lack of belonging in cultural space*, involved the individual being in an outgroup’s environment (e.g., home of a family friend, friend’s house, unfamiliar neighborhood) and feeling hurt, sad, or isolated. This was distinct from Topic 3 (*Othering*) in that narratives containing this topic mentioned the physical space of another person (i.e., a friend’s neighborhood and house) whereas Topic 3 narratives contained diffuse mention of marginalization. However, the top keywords were non-interpretable and included: “interesting,” “holds,” “debate,” “culture,” and “employee.”

Topic 12: Ethnic Stereotypes at School

Topic 12, which is labeled *ethnic stereotypes at school*, features experiences of stereotyping that occur in a school environment. The highly associated words were "studies," "teammates," "gathering," "soccer," and "merely." The key representative narratives for this topic reflected a heightened awareness of stereotypes. The highly associated narrative describes a high schooler who was stereotyped on the basis of their ethnicity as a bad soccer player and defied expectations by scoring a goal. Another narrative was about an elementary school student of Latinx descent who was placed into a "special" bilingual studies class despite speaking fluent English.

Topic 13: Gained Perspective of Structural Inequality

Topic 13, labeled *Gained perspective of structural inequality*, included encounters with inequality at the structural level. Highly associated narratives involved critical reflection, a form of critical consciousness that captures the recognition of systems of inequality and desire to advocate for social change (Watts et al., 2011). The keywords were "perspective," "writing," "analyzing," "lately," "ruin," and "immigrants."

Topic 14: Whiteness

Topic 14, labeled *Whiteness*, revolves around eye-opening experiences of White-conforming norms and privilege. The associated keywords were "class," "opener," "race," "visit," "minority," "ancestry," and "white." The representative narratives elucidate instances of white privilege.

Aim 2: Topic Vocabulary by Participants' Ethnicity/Race

Table 2 contains the group-specific vocabulary words by race for all topics, which are called frequent and exclusive (FREX) words. These refer to words that are uniquely used by a

specific group when discussing a topic. In some cases, vocabulary words simply reflect the names of the racial groups themselves (such as "Irish") but in other cases may relate to disparate experiences. For example, Latinx participants uniquely used the word "illegal" when discussing prejudice, likely reflecting politically-influenced stereotypes about race and immigration. Importantly, group-specific vocabulary words are distinct from high-probability words, which are likely to appear for a specific topic regardless of the narrator's ethnicity and are central to defining the essence of a topic based purely on word frequencies within that topic. As an example of group-specific words, for *Ethnic/cultural celebration*, the unique vocabulary of participants of Asian American descent included the words: "vietnamese," "persian," "ceremony," "keynote," and "skits." This in contrast to Black/African American participants, the unique words were "rally," "adult," "invite," and "protested." This finding suggests that types of organized events (i.e., ceremony versus rally) differed based on self-identified ethnicity.

Specifically, for Multiracial individuals, we observe language around social and peer group settings ("frat," "jokes," "everyone," "crowd," "friends," "school," "soccer"), alongside daily and public spaces ("restaurant," "dinner," "parking," "street"). Their vocabulary included emotionally evaluative terms ("disbelief," "hostile," "disheartened," "hateful," "critically") and action-oriented or defensive language ("quit," "defense," "showed," "involvement"), suggesting episodes of social tension or boundary negotiation.

Several words ("Jewish," "African," "Phillippino," "white," "minority") explicitly referenced racial or cultural categories, indicating awareness of multiple identity markers. The recurrent appearance of *jokes* across Black and multiracial individuals points to humor and social interaction as shared contexts where race becomes negotiated. References to *culture, ceremony,*

and *event* among Asian, Latine, and Multiracial individuals (“rituals”, “holiday”, “ethnic”) suggest a common emphasis on collective or celebratory experiences.

Table 2.2. The five FREX words with each topic across the ethnic/racial groups.

Topics	Black/ African American	Asian	Latine/ Hispanic	Native American/ Indigenous	Multiracial	White	Non-categorizable
Cultural Difference	scared, vanished, cheek, different, greet	store, soup, aunt, different, mom	old, send, hurt, party, everyone	phrase, chicks, exchange, teacher, everyone	frat, jokes, stereotyped, everyone, crowd	bus, minorities, scholarship, minority, everyone	sikh, bothers, interests, awkward, questions
Peer Dynamics	music, puzzled, sleeping, girlfriend, communicate	disrespectful, listing, meat, group, play	english, napa, crushed, speak, song	lions, suburban, white, jokes, names	nights, arms, assimilate, jewish, friends	jokes, persian, hurtful, friends, fun	eastern, phillipino, extreme, friends, remember
Othering	spirit, bus, secluded, water, program	game, teams, song, name, game	displayed, desk, chinese, cousin, read	patrol, plate, unit, car, questions	melody, dallas, rural, melody, ladies	teenage, awestruck, names, help, boys	old, bright, schoolmates, classmate, called
Prejudice	proof, paper, crushed, read, searched	vietnamese, mother, lady, someone, caucasian	bright, sad, illegally, racist, roommate	hawaii, paralyzed, unwelcome, biking, feel	self-esteem, send, disbelief, table, sit	assignment, russian, kiss, mother, skin	confusing, shoved, because, category, furious
Underrepresentation	shoved, scores, environment, church, bothered	esl, five, tan, classmates, eyes	letters, side, signed, cashier, salvador	king, longest, crush, nobody, school	transferred, worded, pursue, known, fact	read, dominated, songs, cousins, movie	proof, disheartened, rural, islamic, refused
Ethnic Cele	rally, invite, protested, incident, adult	ceremony, keynote, skits,	principle, listening,	event, remember, family, many, new	middles, awestruck, hostile, act, march	irish, joy, boards,	save, signed, surprise, perform

bration		event, culture	afterschool, culture, event			event, remember	ances, plenty
Cultural Shock	schedule, option, interacting, teacher, american	seat, institute, row, talk, students	mexican-american, arabs, independence, crazy, upset	reported, caucasian, switch, student, class	christmas, rituals, celebrated, holiday, ethnic	birth, slang, supportive, culture learning	spicy, egyptian, arabic, food, drink
Violence	guard, matched, drug, store, guarding	ran, crowded, crazy, scared, walking	dared, ring, assist, cop, walking	bothering, happen, shoving, store, walking	car, holocaust, truck, parking, street	car, latino, attire, walking, scared	hopes, chuckled, ruin, men, group
Racial Discr	stress, crisis, schoolmates, jokes, identity	slurs, woman, hurtful, racial, car	flight, plate, grandpa, water, presentation	guess, camp, atmosphere, suburb, friends	quit, willingness, defense, showed, african	team, game, gangster, play, basketball	nigeria, secluded, room, nervous, skin
Negative Affect	paused, matters, report, hair, wear	hmong, expensive, folks, dress, community	friends, apologized, represented, wash, cousins	military, removed, nights, map, lost	disheartened, involvement, hateful, critically, man	thrown, apartments, types, fight, city	pledge, challenge, secretary, bus, workers
Cultural Space	employee, mexican, buddy, ignored, blind	pork, dogs, taiwanese, family, food	nostalgic, resolution, conflict, religious, home	saint, pleased, holiday, nationality, irish	insight, left, restaurant, dinner, party	speak, claiming, confusion, food, interesting	blushed, intentionally, funny, family, dinner
School	middles, unwanted, color, ashamed, appear	china, social, letter, shocking, wrote	spanish, teaching, fluent, student, class	inner-city, more, fights, loud, studies	school, soccer, lot, didnt, high	jewish, observing, social, stereotypes, teammates	overheard, assist, scores, drawing, team

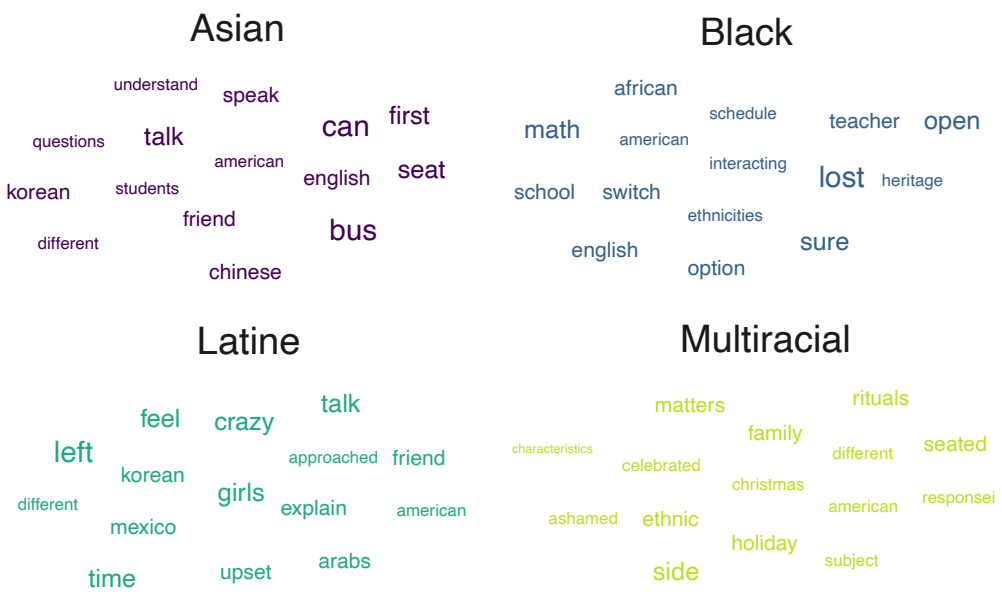
Structural Ineq	blind, buddy, dared, people, felt	india, roommate, pakistan, door, culture	education, expectations, writing, community, immigration	talking, finished, different, just, black	shit, phillipino, intentionally, refused, writing	relieved, faith, questioning, religion, feelings	shyed, ocean, serve, guy, bar
Whiteness	class, race, white, minority, opener	japanese, dads, ten, class, race	challenge, exams, latina, school, white	proposed, effecting, disagree, shows, book	class, race, white, minority, friend	class, study, race, visit, white	class, race, white, minority, different

To further illustrate the differences in topic content between multiracial and monoracial participants, Figure 2.3 has the FREX words for *Experience of Cultural Shock* across multiracial and minority monoracial participants. For Asian and Latine participants, FREX words revolved around communication-related words (i.e., “talk”, “speak”, “understand”) and emotion (i.e., “feel”, “different”, “upset”, “crazy”). This pattern suggests that cultural shock is often experienced through difficulties in mutual understanding, emotional confusion, and social misalignment. The emphasis on speech and feeling implies that cultural difference is felt in the moments of trying to connect and be understood.

In contrast, for Black participants, FREX words point to an educational setting (e.g., “teacher”, “math”, “school”, “schedule”, “english”) and negative affect (e.g., lost). Rather than being framed primarily through interpersonal miscommunication or emotion, their expressions evoke systemic encounters, navigating academic settings. The inclusion of “lost” signals an underlying sense of disorientation or alienation within these structured spaces but filtered through the lens of institutional adaptation rather than linguistic or emotional misunderstanding.

For multiracial participants, FREX words such as “rituals,” “matters,” “family,” “Christmas,” “holiday,” “ethnic,” “side,” “different,” “ashamed,” “characteristics,” and “seated” suggest that experiences of cultural shock are deeply familial and contextually embedded rather than primarily linguistic or institutional. The prominence of “family,” “rituals,” “Christmas,” and “holiday” points to moments of cultural encounter within family traditions and celebrations, where one’s sense of belonging or difference becomes especially salient. The term “side” likely indexes one ethnic or racial background (e.g., “my mom’s side,” “my dad’s side”), indicating how multiracial individuals may experience cultural shock within their own kinship networks. When paired with “different” and “ashamed,” these words evoke moments of internal tension or discomfort, suggesting that cultural shock for multiracial individuals often occurs not between external groups but within family contexts.

Figure 2.3. FREX words with *Experience of Cultural Shock* between multiracial and minority monoracial groups. The words larger in size indicate a higher probability to occur.



Despite there being differences in the content of ethnicity- and race-related narratives between multiracial and monoracial groups, multiracial participants did not significantly differ from other groups in the overall prevalence of topics identified by STM. In Table 3, I ran ANOVAs to compare topic prevalence across ethnicity/race.[9] I found that across the STM topics, topic prevalence varied across ethnicity only for Cultural Shock ($F(5) = 3.30$, $p = .006$, $\eta^2 = 0.01$). Following the Bonferroni correction, Asian and White individuals significantly differed in the prevalence of describing experiences of cultural shock ($t(1125) = 4.00$, $p = .001$, $d = 0.28$). Multiracial participants, however, did not significantly differ from any monoracial group across any of the topics, suggesting that while their narratives may have emphasized certain language

[9] Native American/Indigenous participants were excluded from the analyses examining group differences, as fewer than 50 participants identified with this group.

(e.g., family-embedded terms for cultural shock), these themes were not more or less prevalent than those of monoracial participants. Differences across ethnic groups in topic prevalence were minimal and largely specific to cultural shock. The similarity in topic distributions between multiracial and monoracial participants supports the interpretation that identity content is broadly shared across groups, even if multiracial individuals frame that content with greater emphasis on negotiation and hybridity. I also tested whether the motivational themes differed across ethnicity/race and found that there was no significant difference across groups for communion ($F(5) = 1.40$, $p = .220$, $\eta^2 = 0.01$). I did find that agency significantly differed based on ethnicity/race ($F(5) = 3.07$, $p = .009$, $\eta^2 = 0.01$). However, after using a Bonferroni correction, none of the pairwise comparisons were significant.

Table 2.3. The mean agency and communion score, and topic proportions across ethnicity.

	Group Means							
	Black/African American	Asian	Latinx/ Hispanic	Multiracial	White	Non-categor.	η^2	p-value
Themes								
Agency	1.10	1.30	1.20	1.30	1.50	1.30	.01	**.009***
Communion	0.80	1.20	1.10	1.20	1.20	1.20	.01	.220
Topics								
Cultural Difference	.10	.14	.15	.16	.15	.15	.00	.865
Peer Dynamics	.15	.13	.16	.13	.13	.18	.00	.552
Othering	.12	.09	.12	.13	.07	.10	.01	.137
Prejudice	.07	.07	.10	.08	.10	.07	.00	.548
Underrepresentation	.09	.07	.08	.05	.08	.10	.00	.812
Ethnic Celebration	.04	.09	.06	.10	.09	.05	.01	.301
Cultural Shock	.08	.12	.08	.09	.06	.07	.01	**.006***
Violence	.10	.08	.05	.12	.09	.03	.01	.091
Racial Discr	.11	.04	.05	.04	.06	.04	.01	.147
Negative Affect	.07	.06	.04	.04	.05	.06	.00	.859
Cultural Space	.03	.05	.02	.04	.05	.06	.00	.502
School	.02	.02	.04	.00	.03	.05	.01	.228
Structural Ineq	.00	.02	.03	.02	.02	.03	.00	.652
Whiteness	.01	.02	.02	.01	.02	.01	.01	.273

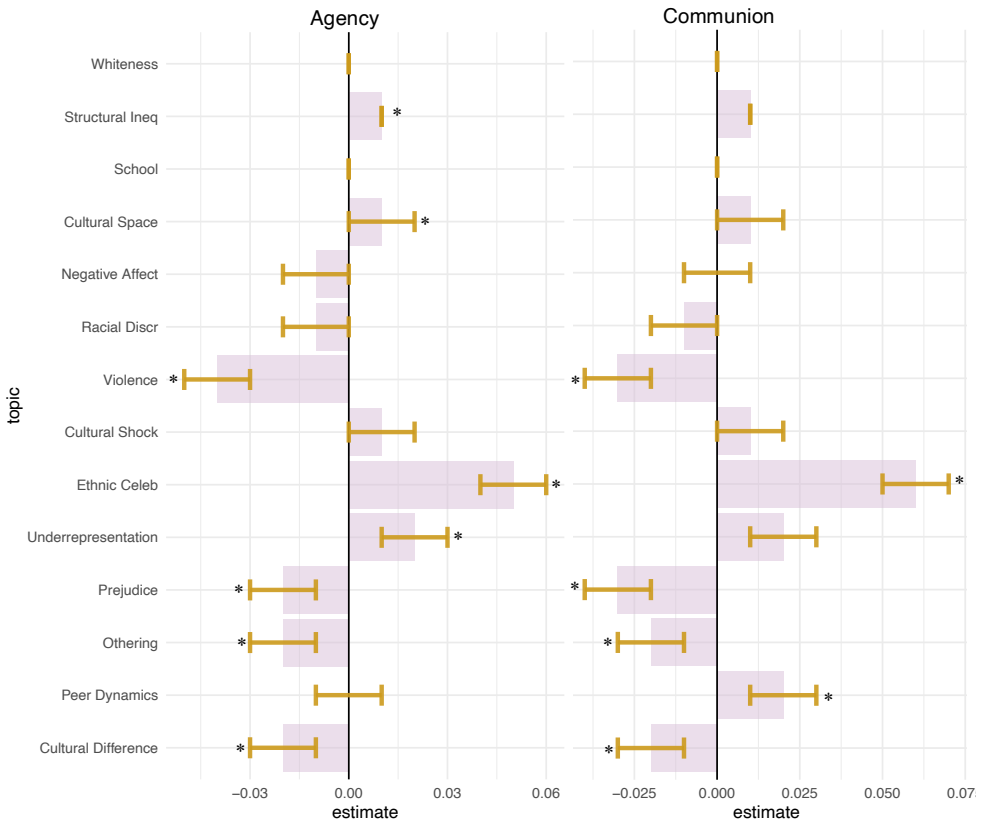
Note. The p-values and effect sizes (η^2) for the ANOVAs are presented. The p-values where the themes and topics significantly differ across ethnicity are asterisked and in bold (p-value < 0.05)

Aim 3: Agency and Communion-Related Topics

I examined the relationship between the STM and human-coded themes to determine the agency and communion-related topics. These relationships were estimated within a regression framework; I report here the slope coefficients[10] indexing the relationship between topic prevalence and theme (see Figure 4). High levels of agency were positively associated with topics *Awareness of underrepresentation* (unstandardized $b = 0.02$, $p = .020$), *Ethnic/cultural celebration* ($b = 0.05$, $p < .001$), *Gained perspective of structural inequality* ($b = 0.01$, $p = .030$, and *Lack of belonging in cultural space* ($b = 0.01$, $p = .034$). Agency was negatively associated with *Othering* ($b = -0.02$, $p = .004$), *Reactance to prejudice* ($b = -0.02$, $p = .011$), *Violence* ($b = -0.04$, $p < .001$), and *Awareness of cultural difference* ($b = -0.02$, $p = .046$).

[10] By using unstandardized coefficients, the interpretation remains in the original scale of the covariates and topics, making it easier to understand how actual changes in covariates relate to changes in topic proportions. Standardizing coefficients can obscure the relationship between the covariates and the raw data, making it less intuitive to relate the model results back to the original data.

Figure 2.4. The estimates between the STM topics with agency and communion

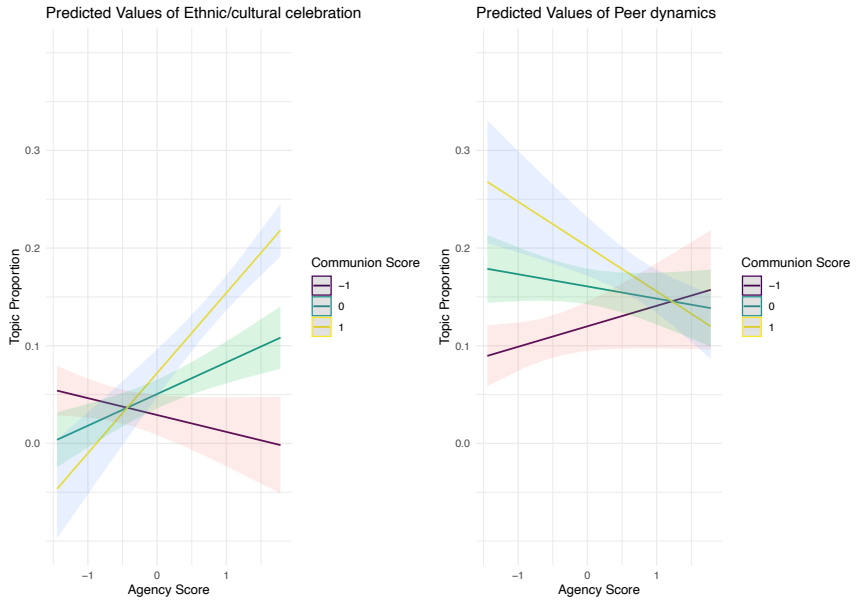


Note. The standard error is represented as the error bars. The topics that are significantly associated with each motivational theme are asterisked (p-value < 0.05).

Communion was positively associated with *Awareness of underrepresentation* (b = 0.02, p = .015), *Ethnic/cultural celebration* (b = 0.06, p < .001), and *Peer dynamics* (b = 0.02, p = .026). Communion was negatively associated with *Othering* (b = -0.02, p = .001), *Reactance to prejudice* (b = -0.03, p < .001), *Violence* (b = -0.03, p < .001), and *Awareness of cultural difference* (b = -0.02, p = .004). Notably, both agency and communion were negatively associated with *Othering*, *Reactance to prejudice*, *Awareness of cultural difference,* and *Violence*.

In examining the relationship between agency, communion, and the topics, I included the interaction between agency and communion. This interaction was significant in predicting the prevalence of *Ethnic/cultural celebration* (unstandardized, $b = 0.05$, $p < .001$). Figure 5a shows the effect of agency on topic prevalence at different levels of communion. The prevalence of *Ethnic/cultural celebration* was positively associated with agency among high communion narratives. When communion is low, there was no significant relationship between *Ethnic/cultural celebration* prevalence and agency. There was a significant interaction effect for *Peer dynamics* ($b = -0.03$, $p = .001$). In Figure 5b, agency was negatively associated with the topic *Peer dynamics* when communion was high, but unassociated with this topic otherwise. *Experience of cultural shock* ($b = 0.02$, $p = .040$) and *Lack of belonging in cultural space* ($b = -0.01$, $p = .010$) also had significant interaction effects; these effects are reported in the supplemental material (see Figure S3). Overall, the presence of communion modified the relationship between agency and topic prevalence across these four topics.

Figure 2.5a & 2.5b. Agency and communion scores predicting topic proportions for *Ethnic/cultural celebration* and *Peer dynamics*.



Discussion

The current study applied STM to narratives of ethnicity-related experiences to better understand ERI content. The 14 major topics of discussion provided insights into lived experiences of ethnicity. The most prevalent topics were *Awareness of cultural difference*, *Peer dynamics*, *Othering*, and *Reactance to prejudice*. One topic, *Experience of culture shock,* was more prevalent in narratives written by Asian American participants compared to White participants. Importantly, multiracial participants did not differ in overall topic prevalence, but their narratives featured distinctive family-centered and heritage-oriented language (e.g., "family," "holiday," "rituals," "side"), suggesting that cultural negotiation and belonging were often articulated through intra-familial contexts. In addition, topics varied in the degree to which they were agentic and communal, with *Ethnic/cultural celebration* being the most prevalent among narratives rated high in agency and communion. Briefly, I will go more into why these findings emerged and limitations for the next studies to address.

Asian American participants described *Experience of culture shock* significantly more compared to White participants. Among the relevant narratives from Asian participants, bicultural participants reporting visiting their heritage country either for the first time or after a long period away. Thus, it may be that the "Asian American" label encompasses both Asian Americans and Asian immigrants who can describe experiences of cultural shock. We would expect the latter group to potentially report experiences of initial exposure to host cultural norms (Benet-Martínez & Haritatos, 2005), while the former group may report exposure to heritage cultural norms (Altweck & Marshall, 2015).

I identified vocabulary differences across ethnic/racial groups to better understand variations in experiences. Black/African American participants used more political language ("rally", "protest") while Asian American participants discussed conferences and celebrations for

Ethnic/cultural celebration; future research might investigate how organized cultural events influence ERI. People, objects, and places surfaced while reflecting on ethnic group identity. (See the *General Discussion* for details on multiracial identity content compared to monoracial groups.)

Lastly, ethnicity-related experiences were sometimes agentic, sometimes communal, and occasionally both. Narratives high in agency often described differentiation from others, such as *Awareness of underrepresentation* and *Lack of belonging in cultural space.* In the face of isolating experiences, the individual may be positioned towards greater autonomy and independence rather than feelings of loneliness (Henning et al., 2022). The framing of agentic narratives often demonstrated acceptance and self-insight that worked to empower the individual (McAdams et al., 1996).

In tandem, highly-communal topics were *Peer dynamics* and *Awareness of underrepresentation.* Although past research found that instances of underrepresentation may contribute to feeling isolated and excluded (Ackerman-Barger et al., 2020; Campbell, 2023), the narratives in this study suggest that underrepresentation may also make salient one's connection with their own ethnic/racial group (Mehra, 2023). For emerging adults, it may be that underrepresentation narratives are highly communal because they emphasize belonging to one's ethnic group rather than disconnect from an outgroup.

In another instance, communion moderated the relationship between agency and *Peer dynamics*, such that topic proportion for peer dynamics was only negatively related with agency when narratives were rated as high in communion. Experiences involving peers often involved a source of connection (Camarena et al., 1990; McLean et al., 2018). When communal identity is high in the presence of others, this may sometimes be at the expense of autonomy (Clasen &

Brown, 1985; Connolly & McIsaac, 2009). Agentic and communal experiences solidify the individual's self-understanding as part of a larger group identity, demonstrating resilience and adaptability in navigating cultural landscapes.

From Shared Ethnic/Racial Experiences to Multiracial Self-Concept

The Study 1 takeaway is that the content and motivational dynamics of ethnicity- and race-related experiences are widely shared across racial groups, even as they are expressed through culturally distinct linguistic and contextual frames. Both agentic and communal orientations differentially appeared to be motivational anchors, shaping how individuals interpret and respond to experiences of cultural difference, celebration, or exclusion. Although Study 1 revealed that the types of ethnicity- and race-related experiences are broadly shared across groups, how such experiences are internalized into one's self-concept may differ substantially. The organization of these experiences within the self (how individuals define who they are and how their multiple group memberships relate to one another) may be more complex and nuanced for multiracial individuals. For those with multiple heritages, external encounters can become structural features of the self, requiring ongoing negotiation between cultural frameworks, social expectations, and personal meaning.

Consequently, Study 2 examines the self-concept of multiracial individuals, how they describe and organize their traits, roles, and group identities, to determine whether the structure of their self-concept (i.e., complexity, integration, and similarity to monoracial parent groups) reflects the same patterns of negotiation observed in their narratives. Study 2 addresses how shared experiences of race and ethnicity give rise to distinct forms of self-definition among multiracial individuals and across diverse ethnic/racial combinations.

Further, multiracial identity encompasses a wide diversity of ethnic and racial combinations, each carrying distinct cultural histories and social meanings. Study 2 begins to disentangle these racial backgrounds beyond the "multiracial" category. Studies 2 and 3 draw from a larger, more diverse multiracial sample, enabling us to examine how patterns of self-concept vary between multiracial and monoracial individuals *and* within multiracial groups themselves. These subsequent studies allow for a more precise understanding of how different heritage combinations shape the ways individuals internalize and structure their experiences of race and ethnicity.

CHAPTER III
Study 2:
Selves in Context: Open-Ended Self-Descriptions Reveal Multiracial Self-Concept

Co-author, Dr. Veronica Benet-Martínez, provided invaluable edits and suggestions for this manuscript under review. As first author, I was the primary contributor to this project.

Introduction

The self-concept, how individuals define who they are, has long been a cornerstone of personality and identity research (Baumeister, 1999; Oyserman et al., 2012). Scholars have conceptualized the self as a cognitive-affective system (Mischel & Shoda, 1995), a narrative structure (McAdams, 1995), and a dynamic representation of the person-in-context (Harter, 2012; Markus & Wurf, 1987) . This narrative tradition highlights the ways in which individuals actively construct meaning through stories that link past, present, and future, providing critical insight into coherence, agency, and cultural situatedness of identity. Much of this work has also emphasized either stable personality traits or predefined social identities, examining how people possess characteristics or affiliate with groups, often through closed-ended self-report questionnaires (Robins et al., 2008; Sellers et al., 1998). Though this approach has generated important insights, it can overlook the active nature of identity construction by those living it (Bamberg, 2011 ; McLean et al., 2007).

Complementing the narrative approach, a body of research has explored the self through open-ended approaches, allowing individuals to describe themselves in their own words (Kuhn & McPartland, 1954; del Prado et al., 2007). These methods capture how people spontaneously represent their identities, often surfacing patterns of meaning that are difficult to access through structured measures. Indeed, the narrative and the open-ended "Twenty Statements Test" traditions can be seen as parallel efforts to honor participants' own meaning-making rather than

imposing predefined categories. They have been especially useful in revealing the cultural content of the self and in tracing how individuals define themselves across contexts (Markus & Kitayama, 1991). Open-ended language offers a valuable window into the lived structure of identity, how people make meaning of themselves when unconstrained by predefined categories or rating scales.

However, much of this open-ended work has focused on normatively categorized groups, whose social positions typically align with stable, singular identity frameworks. Less is known about how the self is constructed by individuals whose lived experience disrupts conventional boundaries of identity. This gap is especially relevant at a time when the cultural landscape is increasingly diverse, and the need to understand identity as fluid, context-sensitive, and negotiated has become more pressing. More specifically, the rise in the multiracial population in the United States (and globally) presents a critical opportunity to explore identity construction in a group for whom traditional measures often fall short or misrepresent lived experience.

Moreover, insights gleaned from this specific population can provide insight into more general processes of self-concept construction for individuals navigating complex multiple cultural environments, such as first-generation college students bridging home and academic worlds, queer folks navigating heteronormative spaces, and religious minorities adapting to secular contexts. As Jean Phinney (1999), an early scholar of ethnic identity, said: ‘‘increasing numbers of people find that the conflicts are not between different groups but between different cultural values, attitudes, and expectations within themselves’’ (p. 27). At the same time, emerging language-based methodologies now offer powerful tools for analyzing self-descriptions at scale, tools that can illuminate subtle dimensions of self-concept such as complexity, contradiction, and hybridity. These conceptual and methodological developments

open an expansive understanding of the self—one that attends to the content of self-concept and to the cultural location and linguistic form through which it is expressed.

Theoretical and demographic relevance of multiracial self-concept

As a population whose lived experience defies fixed racial categories, multiracial individuals offer a unique lens into the dynamics of self-construction. Positioned at the intersection of racial systems, they regularly navigate social contexts that impose shifting labels, competing norms, and ambiguous expectations (Sanchez, Shih, & Garcia, 2009). These conditions often require deliberate acts of self-definition, as individuals work to assemble coherence from culturally divergent or even contradictory reference points.

In response, multiracial people must engage in deliberate acts of identity construction, assembling coherence from divergent ethno-cultural reference points. These experiences make visible psychological dynamics that are often obscured in research with more normatively categorized groups. For example, researchers highlight the self-concept as a negotiated, socially-responsive structure that is relational rather than solely intrapsychic (Markus & Kitayama, 1991). Multiracial individuals also challenge assumptions of internal consistency in the self by holding multiple, even contradictory, self-views in tension (Cheng & Lee, 2009). Studying this population thus provides a generative site for expanding general models of self-concept, surfacing dimensions such as complexity, contradiction, and integration that may be broadly relevant across diverse cultural contexts.

Second, the study of multiracial self-concept carries growing demographic and societal relevance. As racial diversity continues to expand in the United States, multiracial individuals represent an increasingly visible and understudied population (Gaither, 2015). Approximately one in seven marriages in the U.S. are interracial (Bean, 2018), and the number of people

identifying as multiracial has grown over the past two decades. Psychological research has increasingly sought to account for the diverse lived experiences of this population. Multiracial individuals often navigate multiple, and at times conflicting, cultural frameworks, each with its own norms, expectations, and definitions of belonging. These experiences raise critical questions about how people make sense of themselves, how they define who they are, and how racial identity fits into their everyday self-understanding. Understanding how multiracial individuals construct their self-concepts therefore offers not only theoretical insight but also a timely response to demographic and cultural shifts shaping the broader U.S. landscape.

Multiracial identity models

Over the past two decades, research on multiracial identity has revealed several consistent psychological patterns that shed light on how individuals construct a sense of self from multiple racial backgrounds. Multiracial individuals often experience a heightened awareness of racial context and a tendency to shift identity across settings (Hitlin et al., 2006; (Townsend et al., 2009). Many report managing competing cultural expectations or social pressures to conform to monoracial norms (Rockquemore et al., 2009), while others describe pride in their mixed heritage and the development of an integrated racial identity (Cheng & Lee, 2009; Gaither, 2015). These findings point to multiracial identity as both flexible and negotiated: an identity structure shaped by social perception, historical context, and individual meaning-making. Importantly, they suggest that multiracial individuals may cultivate self-concepts that reflect greater complexity, reflexivity, or contradiction than monoracial individuals. While prior work has largely focused on self-categorization or feelings of inclusion and belonging, we know less about how multiracial individuals define themselves in open-ended terms. Do they draw from both parent cultures when describing who they are? Are their self-concepts more likely to

reflect duality, tension, or integration? And how central is race to how they understand themselves? These questions remain underexplored and are foundational to understanding how the self is constructed in response to intersecting systems of identity.

The self-concept as a site of identity construction

While psychological models of multiracial identity have illuminated the dynamic processes by which people negotiate multiple racial backgrounds, they often stop short of examining how such identity work is encoded in the structure and content of the self-concept. Much of the existing literature focuses on antecedents (e.g., social context and racial socialization; Shih & Sanchez, 2005) and outcomes (e.g., well-being, belonging; Albuja et al., 2019; Miller et al., 2019) but less attention has been paid to the architecture of self-definition itself, or how identity is assembled in words, traits, and roles. Yet, self-concept is both a product of internal processes and a culturally-shaped system that reflects the ways individuals understand and express who they are (Adams & Markus, 2004; Markus & Kitayama, 1991). For multiracial individuals, this system may be uniquely configured.

The self-concept is the constellation of attributes, roles, group memberships, and evaluative beliefs that individuals use to make themselves intelligible to both themselves and others. Identity functions as an internalized life story that brings coherence to one's self-understanding while offering a narrative that can be shared and socially affirmed (McAdams, 1995). Similarly, Swann's (1987) identity negotiation framework emphasizes the performative and relational nature of self-definition: people construct self-views in ways that are shaped by their audience and social context. This communicative function is especially important for multiracial people, who are frequently asked to explain, justify, or translate their racial identity to

others (Franco & O'Brien, 2018). As a result, their self-concepts may reflect internal preferences or cultural knowledge and also a learned sensitivity to others' expectations and assumptions.

Self-concept can be decomposed into content and structure. *Content* refers to what individuals say about themselves—the specific traits, identities, values, and roles they claim (McCrae & Costa, 1988). For example, a person might describe themselves as "a Black woman," "hardworking," "creative," or "someone who loves my family." *Structural* features, by contrast, refer to how these self-views are organized, whether they are consistent or contradictory, simple or complex, specific or abstract. These structural qualities are not merely stylistic; they reflect deeper cognitive and cultural processes.

Two such features—dialecticism and reference to cultural templates—are especially relevant for understanding how multiracial individuals assemble a coherent self amid competing identity norms. *Dialecticism* is a tolerance for contradiction, change, and holism, and is more likely to emerge in individuals who navigate competing cultural schemas (Spencer-Rodgers & Peng, 2005). For multiracial individuals, this might be expressed through self-descriptions that incorporate opposing traits (e.g., "quiet and outgoing"), without necessarily resolving the tension between them.

Another key dimension of self-concept structure is the extent to which individuals draw on or diverge from culturally available *identity templates*. These templates are internalized scripts for what it means to be a "typical" White person, Asian American, etc., which can serve as cognitive reference points when people construct their self-descriptions (Oyserman & Markus, 1990). For example, Mexican individuals tend to anchor their self-descriptions—or identity *content*—in relational values like *simpatía* (warmth and agreeableness), respect, and familial interdependence (Ramirez et al., 2012). In contrast, those socialized in more individualist

contexts, such as the United States, emphasize self-expressive themes including autonomy, existential exploration, and personal enjoyment. Multiracial individuals may align closely with one parent group's identity model, blend elements from both, or develop an idiosyncratic profile that reflects neither. Emerging research has begun to explore how multiracial individuals locate their self-concepts in relation to others. For instance, multiracial individuals with white heritage tended to see themselves in ways more similar to their minority background, suggesting that self-definition may be more strongly shaped by experiences of marginalization than by affiliation with dominant groups (Jones & Frank, 2023). However, this study focused on predefined survey items rather than self-generated descriptions. As a result, it leaves open the question of how similarity patterns might emerge organically from individuals' own representations of the self.

Conceptual gaps in existing research

Although theories of multiracial identity routinely acknowledge its complexity, dynamism, and context sensitivity, empirical research must capture how these features manifest in the language people use to define themselves. Much of the field continues to rely on closed-ended survey items that ask individuals to check boxes or rank identification strength (Sanchez & Shih, 2018; Gaither, 2015). While valuable for large-scale data collection, such approaches inherently constrain self-definition, often forcing multiracial participants to flatten their experience into rigid categories that may not reflect how they actually experience or articulate their identities. In many studies, individuals must choose between mutually exclusive racial options or respond to items developed for monoracial populations, leaving little room for emergent, person-centered meaning-making (Rockquemore et al., 2009).

This methodological constraint has produced an uneven empirical landscape. We know, for example, that multiracial individuals often report higher identity conflict, greater identity

shifting, and more context-dependent identification than monoracial individuals (Charmaraman & Grossman, 2010; Shih & Sanchez, 2005). However, we know far less about how multiracial people actually describe themselves when invited to do so in their own words. Do they include more racial references overall? Do their self-concepts resemble those of their monoracial peers, or do they occupy a distinct semantic space?

Addressing these gaps requires a methodological shift: from measuring identity as a static category to capturing how it is constructed through language. A focus on how people say who they are offers insight into the raw materials of identity construction. This approach aligns with narrative and discursive theories of identity, which emphasize the interpretive, evolving, and communicative nature of the self (Bamberg, 2011; McAdams, 1995). It also dovetails with recent calls for person-centered methods in ethnic/racial identity research (Syed et al., 2022), especially for populations whose identities are fluid, contested, or marginalized. By analyzing the structure and content of multiracial individuals' self-descriptions, we can better understand not only how they navigate multiple cultural reference points, but how they synthesize, reconcile, or resist dominant racial frameworks in the act of self-definition.

Opportunity for person-centered, computational methods

Traditional identity research often forces a trade-off between scale and nuance: qualitative methods yield rich insights into meaning-making but are difficult to apply at scale, while quantitative methods can generalize but often flatten complexity. As mentioned in Study 1, recent advances in computational text analysis challenge this binary. Tools such as structural topic modeling, word embeddings and cosine similarity metrics enable researchers to extract semantic and structural patterns from large corpora of open-ended responses. Structural topic modeling (STM; Roberts et al., 2014) is a probabilistic method that identifies latent themes in

textual data by estimating the likelihood that certain words cluster together across documents. As in Study 1, STM allows us to include metadata, such as racial background or identity integration scores, as covariates, helping to reveal how topic prevalence varies across groups.

Word embeddings represent words or phrases as vectors in a multidimensional space, such that semantically similar words occupy nearby positions (Mikolov et al., 2013). These models are trained on large corpora and learn patterns of co-occurrence, enabling comparisons of meaning between participants' self-descriptions and theoretically relevant concepts. For example, an embedding model could reveal whether the word "family" in a participant's statement is used more similarly to "responsibility" or "joy", a nuance that would be missed in keyword-based analyses. Cosine similarity is a metric used to quantify the closeness between two text vectors, often derived from word embeddings (Cer et al., 2017). It calculates the cosine of the angle between two vectors, yielding a value between –1 and 1 that indicates how semantically similar two texts are. In self-concept research, this can be used to assess how closely a multiracial individual's self-description resembles those of monoracial peers, revealing patterns of convergence or divergence across racial groups. These methods preserve the integrity of participants' own words while allowing for systematic, replicable analysis across individuals and groups. They offer a rare combination: person-centered input and population-level analytic power.

This hybrid approach is particularly well-suited for studying multiracial self-concepts. These tools allow us to detect whether multiracial self-definitions are more semantically complex or dialectical and whether they diverge or converge with established cultural identity templates. In short, computational methods do not replace theory-driven inquiry—they complement it,

providing the scale and precision needed to uncover patterns that are otherwise inaccessible through manual coding or survey analysis alone.

Why self-concept structure matters for multiracial psychology

Structural features, such as contradiction, centrality, and similarity, reflect identity navigation. For instance, the co-occurrence of traits like "quiet" and "outgoing" may signal the ability to hold seemingly opposing qualities without resolving them into a singular narrative (Spencer-Rodgers et al., 2005). In multiracial populations, such contradictions may emerge from navigating conflicting cultural expectations or from resisting narrow identity scripts.

Other self-descriptive elements, such as listing "Black woman" alongside "Asian American," may reflect the centrality of ethnicity/race to one's identity, or an intersectional awareness of how race, gender, and other social positions shape the self (Crenshaw, 1989; Sellers et al., 1998). These mentions are not merely referential, they signal the psychological salience and social meaning of racial identity in a person's life. Finally, the degree to which a multiracial individual's self-concept is semantically similar to those of their monoracial parent groups can illuminate patterns of cultural alignment, disidentification, or hybridization. Such measures provide insight into how multiracial people position themselves vis-à-vis dominant cultural templates and social boundaries. Taken together, these features offer a multidimensional lens on multiracial selfhood, one that moves beyond questions of "what" identities people hold, toward deeper inquiries into how they hold them.

Overview of the Present Study

In the present study, I examine how multiracial and monoracial adults define themselves using open-ended self-descriptions. Leveraging computational text analysis, I apply structural topic modeling and embedding-based similarity measures to uncover thematic and semantic

patterns in self-concept content. I also quantify structural features, such as contradiction and racial centrality, using dialectical coding schema and keyword-based coding. Finally, I test whether multiracial identity integration, measured via self-report, moderates the degree of semantic similarity between multiracial individuals and their monoracial parent groups. This multimethod approach allows us to capture the complexity of identity construction at both the individual and group levels. This study seeks to describe the multiracial self-concept through four primary dimensions:

1. *Content:* What themes and categories emerge in multiracial individuals' self-concept descriptions?
2. *Complexity:* Do multiracial individuals construct a self-concept characterized by higher semantic coherence and dialecticism?
3. *Centrality [of racial identity]:* To what extent do multiracial individuals spontaneously reference their racial identity in unprimed self-concept descriptions? Does the presence of a racial identity in self-concept increase the likelihood of the appearance of other specific identity categories?
4. *Similarity:* How similar are multiracial individuals' self-concepts to those of their parent racial groups, and does multiracial identity integration influence similarity?

This study contributes to a growing effort to re-center self-definition within identity research. By examining both the content and structure of multiracial individuals' self-descriptions, this study offers a novel lens on identity as a lived, constructed experience. Our approach blends the richness of open-ended, person-centered data with the analytic power of computational methods, demonstrating how new tools can illuminate patterns that are otherwise inaccessible through traditional surveys or categorical models. In doing so, I advance theoretical understanding of

multiracial identity and expand the methodological repertoire available for studying the self in diverse populations.

Methods

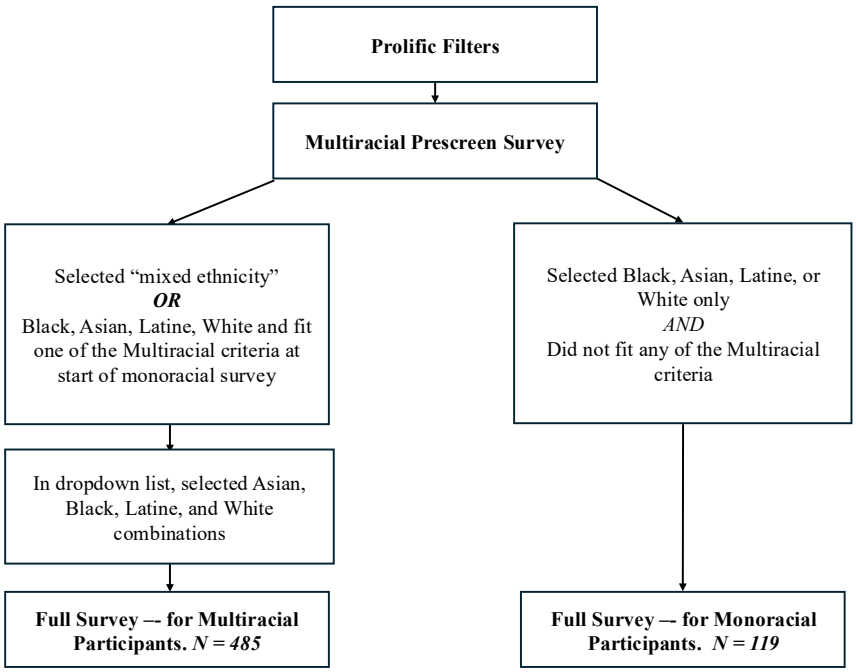
Participants

For this study, a participant was defined as multiracial if they self-identified as biracial or multiracial, selected more than one racial identification, or selected different races for their parents. These criteria were based on a definition of multiracial outlined by Tsai and colleagues (2024). I recruited a U.S. sample of multiracial participants from Prolific in July of 2024. First, using Prolific filters, I identified participants of the following racial categories: Asian, Black/African American, Latine, White, and Mixed. Second, I invited those participants to complete a pre-screening survey that included three items: (1) how many races they identified with, (2) whether their parents were of the same race, (3) and from a dropdown list, which races they considered themselves to be. I used answers to this survey to perform a quota-sampling procedure in which I attempted to recruit 100 participants with each of the following multiracial identities: of Black-White, Latine-White, Asian-White, Asian-Latine, Black-Latine and Black-Asian descent.

I also sought to compare the experience of multiracial individuals to that of monoracial individuals. To establish a clear point of comparison for multiracial participants, using Prolific filters, I recruited a U.S. sample of monoracial participants in equal numbers from four racial groups: Black/African American, Latine, Asian, and White. A target of approximately 30 participants per group was set to ensure sufficient statistical power for moderate effects while maintaining feasibility (Van Voorhis & Morgan, 2007). These participants were given the same pre-screening survey. Eight participants who self-identified as monoracial on Prolific but

reported having parents of different races in the pre-screening survey were reassigned to the multiracial participant sample.

Figure 3.1. Recruitment of eligible multiracial and monoracial participants.



Our final sample size was 604 participants. Racial demographic is the following: 107 Latine-White, 103 Asian-White, 101 Black-White, 85 Black-Latine, 46 Asian-Latine, 43 Asian-Black, 30 Asian, 30 Black/African American, 28 Latine, and 31 White participants. Breakdown across gender is 325 women, 240 men, and 39 nonbinary persons.

Table 3.1. Descriptive statistics by multiracial and monoracial identity.

Variable	Multiracial (N = 485)	Monoracial (N = 119)
Age		
18-25	127 (26.2%)	20 (16.8%)
26-35	174 (35.9%)	47 (39.5%)
36-45	129 (26.6%)	32 (26.9%)
46-55	37 (7.6%)	12 (10.1%)
56-65	2 (0.4%)	0
66+	0	1 (0.8%)

Variable	Multiracial (N = 485)	Monoracial (N = 119)
Not given	16 (3.3%)	7 (5.9%)
Gender		
Woman	273 (56.3%)	52 (43.7%)
Man	180 (37.1%)	60 (50.4%)
Nonbinary	32 (6.6%)	7 (5.9%)
Race		
Asian		30 (25.2%)
Black/African American		30 (25.2%)
Latine		28 (23.5%)
White		31 (26.1%)
Asian-White descent	103 (21.2%)	
Asian-Black descent	43 (8.9%)	
Asian-Latine descent	46 (9.5%)	
Black-White descent	101 (20.8%)	
Black-Latine descent	85 (17.5%)	
Latine-White descent	107 (22.1%)	

Measures

Self-concept. Participants completed the Twenty Statements Test (Kuhn & McPartland, 1954), in which they provided twenty self-descriptors that answer the question of "Who am I?" The survey contained the following prompt:

> "In this survey, we'll be asking you to describe yourself. I'll be using a variety of questions and prompts that we hope will collectively capture who you are. Please answer honestly and as fully as you can. We will start by simply asking: Who are you? Write down twenty different responses to the question "Who am I?" Don't worry about evaluating the logic or importance of your responses - just write the answers as they occur to you. Give yourself five minutes to complete this task."

All participants provided 20 statements, resulting in 12,080 unique self-concept descriptors.

Multiracial identity. As detailed in the introduction, the construct of multiracial is complex, socially constructed, and context specific. Therefore, throughout this study, I define multiracial

identity using several operationalizations, allowing for a more holistic and nuanced understanding of this population. Specifically, I describe and compare the following groups:

- Monoracial vs Multiracial: multiracial participants are those who self-identified as biracial or multiracial, selected more than one racial identification, or selected different races for their parents. Monoracial individuals are all others.
- Specific Racial Identification: Participants are categorized as Latine-White, Asian-White, Black-White, Black-Latine, Asian-Latine, Asian-Black, monoracial Asian, monoracial Black/African American, monoracial Latine, and monoracial White. Multiracial participants were asked to identify two of their racial identities in open-ended items. In a third open-ended item, 52 (8.6%) of the Multiracial participants disclosed other racial groups (e.g., Middle Eastern; see Appendix B). The two racial groups they chose were used to identify their parent categories. *Parent category* refers to a single-race group that is part of a multiracial individual's ancestry, providing a foundational reference point for shared cultural or familial experiences. (For example, a participant who identified as Filipino and Black was assigned parent categories of Asian and Black.)
- Majority-Minority vs Dual-Minority (Multiracial Individuals Only): Majority-minority participants (64% of the multiracial sample) are multiracial participants who identified "White" as one of their two racial identities, while dual-minority participants are those who identify as Black-Latine, Asian-Latine, or Asian-Black.

Multiracial identity integration. I assessed participants' multiracial identity integration with a modified version of the Bicultural Identity Integration scale (BIIS-1; Benet-Martínez & Haritatos, 2005). Items were adapted by replacing the word "bicultural" with "multiracial." In some cases, wording was generalized for clarity (e.g., the item, "I find it easy to harmonize ___

and American cultures," was changed to "I find it easy to harmonize my racial identities"). Participants rated these items on a scale from 1 (*strongly disagree*) to 5 (*strongly agree*). The modified BII demonstrated good internal consistency, $\alpha = .86$ ($M = 3.64$, $SD = 0.67$, range = 1.29–5.00).

Analyses

Content: What themes and categories emerge in multiracial individuals' self-concept descriptions?

I described multiracial people's self-concept using two methods. First, consistent with prior research (Jackson et al., 2019; McCrae & Costa, 1988; Tamm et al., 2024), each of the twenty TST statements from each participant was categorized into one of the following categories: loci (age, gender, race), family/relational roles, occupation, hobbies/interests, personality traits, social class, goal/aspirational, universal descriptors, physical and emotional states, ideology/attitudes, and self-evaluation. This coding was performed manually by the first author. In addition, descriptions of these categories were given to GPT-4 (OpenAI, 2024; see Python code on OSF at https://osf.io/rykc5/?view_only=944e248f5fae4bf5bbb54948aeb66641) to independently code each of the statements. There was strong agreement between the coder and GPT-4 ($\kappa = .86$; Landis & Koch, 1977). Mismatches were reconciled manually for the final categorization of descriptors by the first and senior author. I chose to include new categories of "personality types" (e.g., libra, enneagram 8; $N = 8$, 0.07%) and "skills" (e.g., multilingual, Japanese speaker; $N = 29$, 0.24%) based on the unclassified or misclassified descriptors. Given the low base rates of these two categories, I did not include them in the results. I report the distributions of all other categories across the different operationalizations of multiracial and

make statistical comparisons using independent samples *t*-tests, with a Holm correction applied within each descriptor type.

Our second method of analyzing self-concept was through *structural topic modeling (STM;* Roberts et al., 2013) using the *stm* R package (version 1.3.6.1, R version 4.3.1; Roberts et al., 2014). This method is used to uncover latent themes, or "topics," in textual data. Similar to identifying latent factors underlying observed variables with exploratory factor analysis, STM uncovers latent topics driving word distributions in text. STM may include covariates to determine whether topic prevalence varies across groups; for these analyses, I included coded self-identified ethnicity/race (e.g., Black-Latine) as a prevalence covariate. The participant's set of twenty self-descriptors are treated as a single document, and I identified an optimal number of topics across documents based on interpretability and fit metrics and labeled topics by their most representative words and self-concepts (Jacobson et al., 2024; Weston et al., 2023). Each document is a mix of topics – in other words, each participant's self-concept receives a prevalence score for each of the topics extracted, much like each participant receives a factor score for each factor following EFA. Moreover, model output allows us to assess whether topic prevalence significantly differed across racial groups, offering insights into group-specific self-concept patterns.

Complexity: Do multiracial individuals construct a self-concept characterized by higher semantic coherence (lower complexity) and dialecticism?

I used the Dialectical Coding Schema (DCS; Spencer-Rodgers et al., 2009), which measures how holistic, changing, and contradictory one's self-concept is. I created dictionaries for each aspect of the DCS. First, I counted *holistic descriptors* where the participant described themselves as a universal entity, such as "human." Second, *change descriptors* were measured

through dynamic statements (e.g., “striving,” “wants,” “struggling”—to represent becoming), temporal (e.g., “sometimes,” “now”), situational (e.g., “at” and “around”), and qualitative modification (e.g., “somewhat,” “a bit”; Spencer-Rodgers et al., 2009). Third, *contradictory DCS* was measured by assessing negation (“never,” “not”), within-statement contradiction (using all conjunctions except “and”, such as “yet,” “but”), and contradictory pairs (R code for DCS dictionary is available at https://osf.io/rykc5/?view_only=944e248f5fae4bf5bbb54948aeb66641). Contradictory pairs are the only DCS theme that was human coded. Coding was performed by two research assistants, who had access to the self-descriptors only and were blind to the research aims. For example, participants’ contradictory self-descriptors included the following: “full of faults, weak, in need of help” and “strong, unbreakable, an anchor” portrays contradiction. Another participant included: “sweet-natured” and “easily upset.” I report the distributions of DCS categories across our different operationalizations of multiracial.

Lastly, I measured the semantic coherence of participants' self-concept because, while dialecticism shows oppositional self-descriptors, semantic coherence represents how participants’ various self-descriptors—across roles, attributes, and states—organize into a cohesive whole. In other words, we’re interested in the thematic cohesiveness of individuals’ self-concepts. If a person’s self-descriptors are semantically similar (e.g., “athletic,” “sports lover”), it suggests that their self-concept is organized around closely related roles or attributes. By quantifying how these descriptors cluster or diverge in meaning, we gain insight into whether an individual’s various self-aspects are semantically complex.

To do this, I used GPT-text-embedding-3-small to convert each of participants' 20 descriptors into text embeddings, numerical representations of words or phrases that capture their meaning in dimensional space. Descriptors that are more similar will have more similar text

embeddings. For example, “I am a daughter” and “I am a wife” as text embeddings are positioned closer across the dimensional space than “I am a daughter” and “I am athletic.” For each participant, I computed pairwise cosine similarity scores between all self-descriptor pairs within their self-concept, reflecting how closely the descriptors are semantically related. The overall semantic coherence score was calculated as the average of these pairwise similarity scores, providing a single metric that quantifies the degree of semantic consistency in their self-concept. This metric ranges from -1 to 1, with higher values indicating greater similarity between words.

Centrality: How does racial identity feature in self-concept?

I counted how often multiracial participants mentioned ethnicity/race as part of their self-concept. Race was not primed, as the TST was the first question participants completed, so they were unaware of what criteria were used to determine inclusion in the study or of the other questions in the survey. I compared frequency of ethnicity/race mentions across subgroups. In addition, I created a logistic regression-based network plot to see whether including ethnicity/race in one’s self-concept was associated with the appearance of other content categories. This approach allowed us to visualize the interconnected nature of self-concept categories and assess how the presence of race relate to the likelihood of other facets appearing in an individual’s self-concept. I control for the number of categories used by participants, as participants who use more categories overall will necessarily have lower base rates of any specific category.

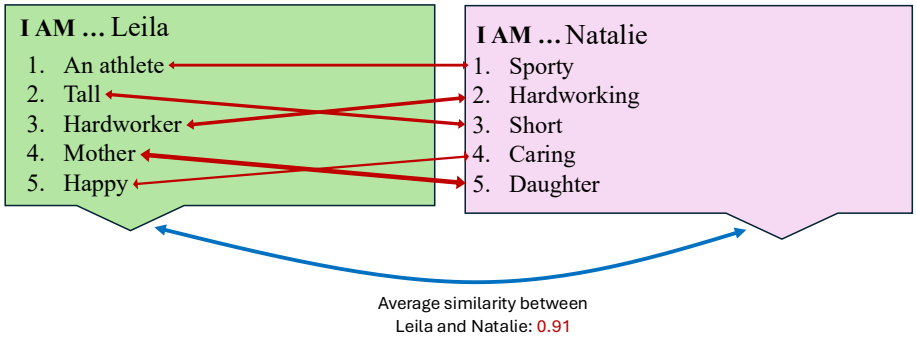
Similarity: How similar are multiracial individuals' self-concepts to those of their parent racial groups, and does multiracial identity integration influence similarity?

I aimed to understand self-concept similarity across individuals through two research sub-questions: (1) Are participants' self-concepts more similar by racial grouping (e.g., dual-minority versus majority-minority), and (2) is multiracial identity integration associated with how similar their self-concept is to members of their racial groups? These two questions capture both group-level and individual-level influences on self-concept similarity. First, looking at racial composition highlights how broad categorical distinctions might shape the way people see themselves. Given that self-concept typologies revealed common content domains (e.g., traits, emotions, ideologies, physical characteristics, and social roles), I ask whether these thematic descriptors cluster meaningfully by overarching racial identity structures. Second, examining multiracial identity integration focuses on individual differences, such that the psychological process of merging multiple identities can influence similarity to each parent racial group. Prior research found that bicultural identity integration was positively associated with similarity in personality ratings of the self to the typical individual of each cultural group (Miramontez et al., 2008). Similarly, I wanted to explore whether higher multiracial identity integration was associated with increased similarity to their parent racial groups. As such, integrated multiracial individuals may exhibit self-concepts that incorporate elements from both parent racial groups, leading to higher similarity scores.

To calculate similarity between participants' self-concepts, I used the following procedure: First, each participant's 20 self-descriptors were converted to numerical vectors (embeddings) using GPT's text-embedding-3-small model. These embeddings capture semantic meaning—similar descriptors have similar embeddings. Next, I ipsatized embeddings to control

for individual differences in how people use language (Cutler & Condon, 2023), ensuring comparisons focus on content rather than response style. Then, for each pair of participants, I used bipartite matching to pair their descriptors optimally. This algorithm finds the pairing that maximizes overall similarity across all descriptor pairs, regardless of the original order participants listed them. Finally, I computed the average cosine similarity across all matched descriptor pairs, yielding a single similarity score (0-1) for each participant pair (N_{pairs} = 182,106). For example, in Figure 2, Leila wrote "athlete" and "mother" while Natalie wrote "sporty" and "daughter," the algorithm would pair "athlete" with "sporty" (high similarity) and "mother" with "daughter" (high similarity) rather than pairing them in order; Leila and Natalie are given the score 0.91 for their similarity to each other.

Figure 3. 2 Bipartite matching across two individuals' self-descriptors, and their average similarity score.



Once all descriptors were matched, the overall similarity in self-concept was calculated as the average cosine similarity score across the matched-pairs. With a single metric of similarity for each participant-pair, I created a participant similarity matrix (e.g., each participant represented as both a row and column, and cell values were the average descriptor-pair cosine

similarity). While descriptor cosine similarity scores range from -1 to 1, with higher values indicating greater similarity, participant similarity ranged from 0 to 1. In Table 2, I include examples of two participants' self-concept with low cosine similarity and high cosine similarity.

Table 3.2. Participants' self-concepts and their average cosine similarity score.

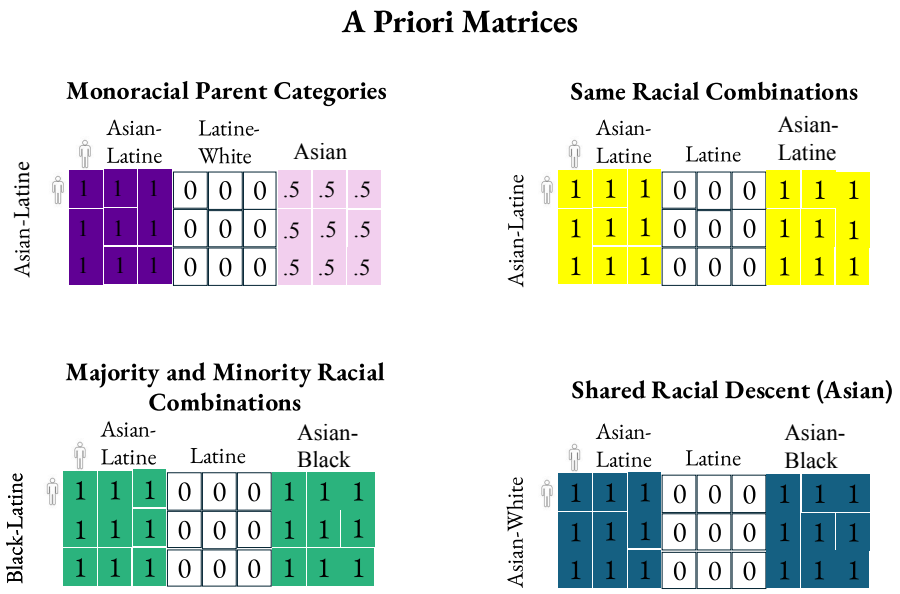
Self-concept #1	Self-concept #2	Similarity score
"a man, a lost soul, a failure, not who I want to be, someone who goes in a different direction, fed up, caring but unable to show it, a degenerate, an alcoholic, a bad partner, straight, eternally in my head, book smart but not educated, ignorant, stubborn, tired, not improving myself, retrospective, self aware, an idiot"	"a woman, a dog mom, a podcast lover, 27, a Taurus, a lover of movies, a daughter, sister and aunt, someone who laughs at everything, a filipino and black woman, a reader, a swimmer, a hockey lover, a nail artist, a makeup lover, allergic to apples"	0.10
"a daughter, a sister, Puerto Rican, a person of color, bisexual, agnostic, intelligent, strong, beautiful, a human, a student, a friend, a pet owner, short, funny, friendly, creative, indecisive, original, a person"	"a nursing student, a caregiver, an aunt, a sister, a daughter, a friend, an animal lover, a christian, a woman, black, american, puerto rican, brown, short, intelligent, creative, thoughtful, giving, intuitive, peaceful"	0.74

Similarity Q1. To assess whether racial grouping influenced self-concept similarities across participants, I compared the observed similarity matrix (derived from participants' self-concept descriptors) with four types of a priori matrices representing distinct racial grouping frameworks: monoracial parent categories, same mixed-race combinations, majority vs. minority racial combinations, and shared racial descent. The theoretical rationale underlying each matrix is that sharing *at least one* monoracial parent group may reflect overlapping cultural, familial, social referential self-descriptors (monoracial parent categories); having the exact same mixed-

race combination suggests more parallel in self-concept (same mixed-race combinations); individuals with two minority backgrounds may experience self-concept formation differently from those who have both a majority and minority background due to disparities in privilege and marginalization (majority vs. minority combinations); and possessing any portion of a racial ancestry can create a sense of shared lineage or collective identity that shapes self-concept (shared racial descent). This approach allowed us to systematically test whether similarities in self-concept are driven by broad racial distinctions, shared mixed-race identities, or specific combinations of racial groups.

These a priori matrices were constructed such that overlapping group memberships were assigned a similarity value of 1 (e.g., participants sharing the same racial category or mixed-race identity), while non-overlapping memberships were assigned a value of 0 (e.g., participants belonging to different racial groups). For the parent priori matrix, multiracial participants with monoracial individuals of overlapping parent category receive a .50 similarity score and same mixed-race a 1. For same mixed-race category, participants of overlapping racial combinations receive a 1 and those non-overlapping a 0. There are four a priori matrices for each monoracial category. For example, participants of Asian descent received a similarity score of 1 and non-Asian participants received a similarity score of 1 with each other, while those who mismatched received a similarity score of 0. The final hypotheses, those overlapping in dual-minority, majority-minority, minority, or majority would yield a similarity score of 1, whereas non-overlapping (e.g., a dual-minority and majority-minority participant) would have a similarity score of 0. The observed dissimilarity matrix was created by subtracting the similarity matrix generated using LLM from 1, as well as subtracting the hypothesized matrices from 1.

Figure 3.3. The racial grouping framework for self-concept similarity



Note. The Racial grouping framework contains a priori matrices for the hypotheses. As an example, Matrix 4 supports the hypothesis that participants of Asian descent are most similar in self-concept. Participants of Asian descent have a cosine similarity score of 1, and those who are of non-Asian descent a cosine similarity score of 1. In the illustration, individuals of Asian-Black descent on the top right are paired with individuals of Asian-White descent on the far left.

In sum, our observed similarity matrix was compared to seven a priori matrices representing distinct hypotheses: (1) monoracial parent categories, (2) same mixed-race

combinations, (3) majority vs. minority racial combinations, and (4) shared Asian descent, (5) shared Black descent, (6) shared Latine descent, and (7) shared White descent. I evaluated which hypothesis best explained the observed data by performing model comparisons using Bayesian Information Criterion (BIC). BIC provides a measure of model fit, balancing goodness of fit with model complexity. Lower BIC values indicate better fit, with penalties for overly complex models to avoid overfitting. I also used Pearson correlations to measure the strength and direction of the relationship between the observed dissimilarity matrix and each a priori dissimilarity matrix. Higher correlations support similarity in self-concept by racial grouping. Correlations were compared with the *cocor* package in R (version 1.1.4, R version 4.3.1, Diedenhofen & Musch, 2015).

Similarity Q2. I also examined whether integrated multiracial identities resemble the identities of monoracial individuals of the same race. For example, does the identity of an Asian-Black man resemble the identities of monoracial Black (or Asian) people, and is this resemblance a function of multiracial identity integration? To test this, I fit a multilevel model with parent categories (Level 1) nested within participants (Level 2). In this model I regressed participants' average cosine similarity score to individuals from each parent category onto identity integration. Table 3 illustrates how data were organized for these analyses by showing the rows for two participants. Each multiracial participant contributed two rows to the dataset, with each row representing their similarity to individuals from one of their parent racial categories (Level 1). Their multiracial identity integration score (Level 2) is how integrated the participants' racial identities are. Thus, only multiracial participants who have Black/African American heritage were compared to monoracial Black/African American participants. Similarly, multiracial participants with Latine, Asian, or White heritage were compared

exclusively to participants from those respective monoracial categories. This nested data structure ensured each participant's similarity scores were explicitly compared only within their own racial backgrounds.

Table 3.3. Example data for parent category similarity models

Multiracial Participant	Monoracial Category	Similarity	MII	Multiracial Combination
1	Black	0.49	2.29	Black-Asian
1	Asian	0.47	2.29	Black-Asian
2	White	0.41	3.81	Asian-White
2	Asian	0.57	3.81	Asian-White

Example data shows how each multiracial participant was represented twice—once for each monoracial parent group. Similarity scores reflect how closely their self-concept, on average, matched members of each group. Multiracial Identity Integration indicates how integrated their racial identities are.

Multiracial identity integration was measured using the modified BII scale. I regressed participants' similarity scores onto their identity integration scores, the parent monoracial category being compared, and the participants' multiracial combination. To account for dependencies between two similarity scores from the same participant, the model included a participant-level random intercept (μj). This random intercept captured individual-level variance, allowing us to accurately estimate how multiracial identity integration influenced participants' similarity scores within their specific parent racial categories. The equation for the model was as follows:

$$\text{Similarity score}_{ij} = \beta_0 + \beta_1(\text{identity integration}_{ij}) + \beta_2(parent\ monoracial\ category_{ij}) + \beta_3(\text{multiracial combination}_{ij}) + \mu_j + \epsilon_{ij}$$

β_0 is the fixed intercept, β_1, β_2, β_3 are fixed effects for the predictors, μ_j is the random intercept for each participant, and ϵ_{ij} is the residual error.

Transparency & Openness

All close-ended data central to the study's primary analyses, including survey responses, human-coded ratings, and demographic variables, have been made publicly available on the Open Science Framework and can be accessed at https://osf.io/rykc5/?view_only=944e248f5fae4bf5bbb54948aeb66641.[11]

Results

Content: What themes and categories emerge in multiracial individuals' self-concept descriptions?

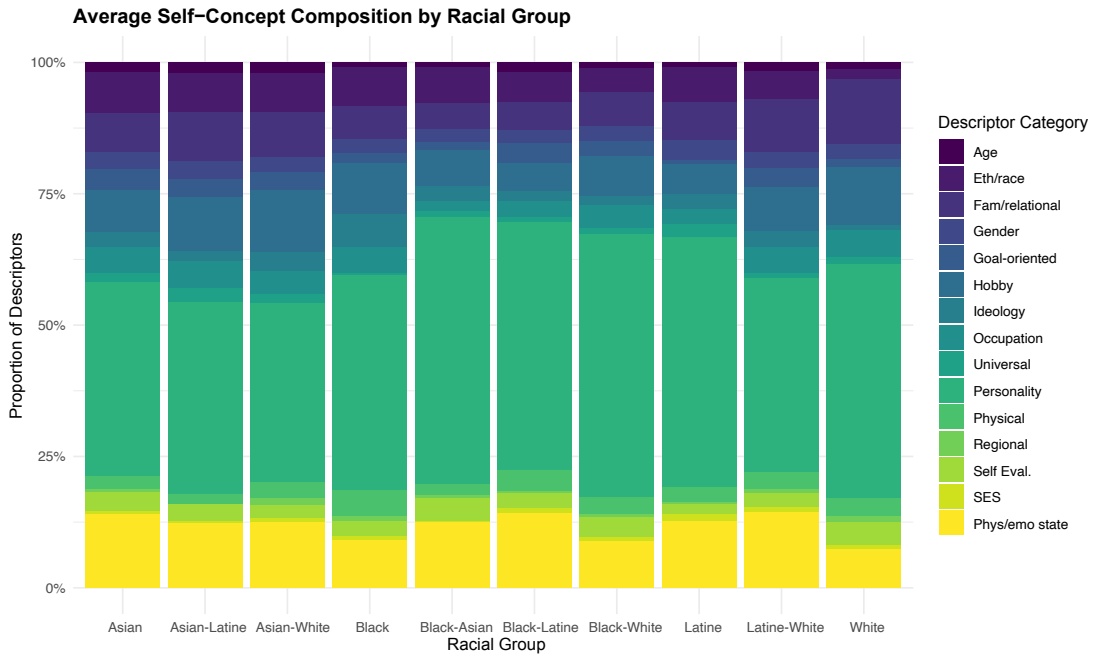
Overall, dual-minority multiracial participants emphasized personality traits, while majority-minority multiracial participants mentioned hobbies, and monoracial participants more often highlighted ideology and relational roles. Figure 4 depicts the average self-concept composition across racial groups by self-concept categories. The stacked bar chart reflects relative similarity in overall average composition; however, there are notable differences in how often certain categories are mentioned (e.g., personality traits and familial/relational roles). Comparing multiracial with monoracial participants, significant differences appeared across self-descriptors around family/relational roles, personality traits, ideology, and hobbies.

Using Holms-corrected *t*-tests, monoracial minority individuals mentioned ideology as part of their self-concept significantly more than dual-minority multiracial ($t(600) = 0.35$, $p = .034$, $d = 0.22$) and White individuals ($t(600) = 0.56$, $p = .034$, $d = 0.22$). Black/African American participants mentioned ideology significantly more than the three dual-minority groups,

[11]Due to the inclusion of identifying information, qualitative data are not publicly posted; however, these data may be made available upon request to the corresponding author and contingent on the completion of a data use agreement. All analysis scripts used in this study (and in supplemental materials) are also publicly available at the same repository link provided above.

majority-minority groups—Black-White, Latine-White—and White participants (see Appendix B). I also found that Black-Latine participants mentioned relational roles significantly less than White participants ($t(594) = -1.23$, $p = .035$, $d = 0.21$). Differences emerged between multiracial subpopulations as well. For example, Asian-White participants mentioned hobbies significantly more than Black-Latine participants ($t(594) = 1.20$, $p = .008$, $d = 0.21$). Lastly, dual-minority participants mentioned personality more as part of their self-concept than majority-minority multiracial individuals ($t(600) = 1.34$, $p = .016$, $d = 0.13$).

Figure 3.4. Self-concept content composition by racial group



Next, I estimated self-concept themes using structural topic modeling. I generated a 6-topic model, selecting the optimal number of topics based on established metrics: exclusivity, variational lower bound, residual, and semantic coherence (see Appendix B; Jacobson et al., 2024; Weston et al., 2023).[12] Exclusivity measures how uniquely words are associated with a

[12] For triangulation, I also analyzed thematic content using the text embeddings from LLM. I applied Gaussian Mixture Models, a clustering approach that groups people most similar across their text embeddings, to the similarity matrix (derived using GPT-4o) and found that the typologies of self-concept corroborated with those from STM. These findings are included in the Appendix.

single topic, with higher values indicating clearer topic distinctions. Variational lower bound reflects model convergence, with lower values (when all else is equal) suggesting better fit. Residuals, in this context, refer to their dispersion, though interpretations vary—some researchers look for local minima, while others consider values below 1 as sufficient for good fit. Finally, semantic coherence measures how frequently high-probability words within a topic co-occur, a metric that correlates well with human assessments of topic coherence, though its validity remains debated (Koltcov et al., 2019). While each of these metrics provides insight into model fit, none alone can fully determine an ideal, interpretable model. Once the optimal number of topics was selected, the topics were labeled using (1) highly associated words and (2) highly associated self-concepts within each topic. Since the top words for certain topics (e.g., topic 2 and 4) appear to be gender-oriented, I compared topic prevalence across gender.

Topic 1, Agentic Traits, include self-reflective attributes, such as "creative," "happy," "smart," "strong" (see Figure 5). The highly associated self-concepts are comprised of personality traits and states. For example, physical states, "tired" and "bored," are highly associated words with Topic 1. Table 4 presents a highly associated self-concept with trait descriptors: "empowering, inspiring, brave, competent." Topic 1 is significantly more prevalent for women (M = 0.21) than men (M = 0.16) and nonbinary individuals (M= 0.13), with a small effect size ($F(2, 601) = 3.32$, $p = .037$, $\eta^2 = 0.01$).

Topic 2, Social Roles, include family roles, such as "daughter," "sister," "wife," "mother" and hobbies and occupations, such as "reader," "student" and "artist." The highly associated self-concepts are predominantly nouns reflective of social roles, as shown in Table 4. There is a significantly higher topic prevalence for women (M = 0.31) than nonbinary

individuals (M = 0.18) and men (M = 0.11), with a large effect size($F(2, 601) = 44.58$, $p < .001$, $\eta^2 = 0.13$).

Topic 3, Emotional States, included words like “loves,” “life,” “trying.” These self-concepts are framed with “I am a person” and “I am someone,” as depicted in Figure 5 as top words. The highly associated self-concepts include “feeling,” “trying,” “praying,” “wanting.” These self-descriptors are future and goal-oriented, and reflective of a dynamic self. Topic 3 is significantly more prevalent for nonbinary individuals (M = 0.27) than men (M = 0.16) and women (M= 0.11), with a small effect size ($F(2, 601) = 8.99$, $p < .001$, $\eta^2 = 0.03$).

Topic 4, Physical Entity, includes physical attributes, gender-related words (e.g., man and male), holistic language, such as human, and a balance of positive personality traits and family roles.” The highly associated self-concepts include descriptors of physical attributes. “Man,” “son,” and “husband” were top words, such that topic 4 prevalence was significantly more prevalent for men (M = 0.31) than women (M = 0.12) and nonbinary individuals (M = 0.10), with a large effect size ($F(2, 601) = 49.72$, $p < .001$, $\eta^2 = 0.14$).

Topic 5, Multiple Heritage/Identities, revolve around describing one’s cultural heritage. The highly associated self-concepts are from participants who describe their ethnicity, race, and other social identities. In Table 3, the individual describes navigating multiple cultures. Another representative is framed as: “I am a guy who doesn't seem to fit to any race or nationality due to being mixed.” A highly associated self-concept focuses on multiple identities: “a genderfluid person, pansexual, biracial, white.” Similarly, of high prevalence in both minority monoracial and multiracial participants. Topic 5 is more prevalent for nonbinary individuals (M= 0.19) than men (M = 0.15) and women (M = 0.09), with a small effect size ($F(2, 601) = 6.20$, $p = .002$, $\eta^2 = 0.02$).

Topic 6, Social Traits, often highlights interpersonal qualities, communication styles, group dynamics, and overall social presence. The highly associated self-concepts involve traits that are interpersonally-relevant, such as “kind,” “honest,” “caring.” Topic 6 is significantly more prevalent for women (M =0.17) than nonbinary individuals (M = 0.13) and men (M = 0.11), with a small effect size ($F(2, 601) = 6.85$, $p = .001$, $\eta^2 = 0.02$).

Figure 3.5. The 14 highest probability words for the STM topics

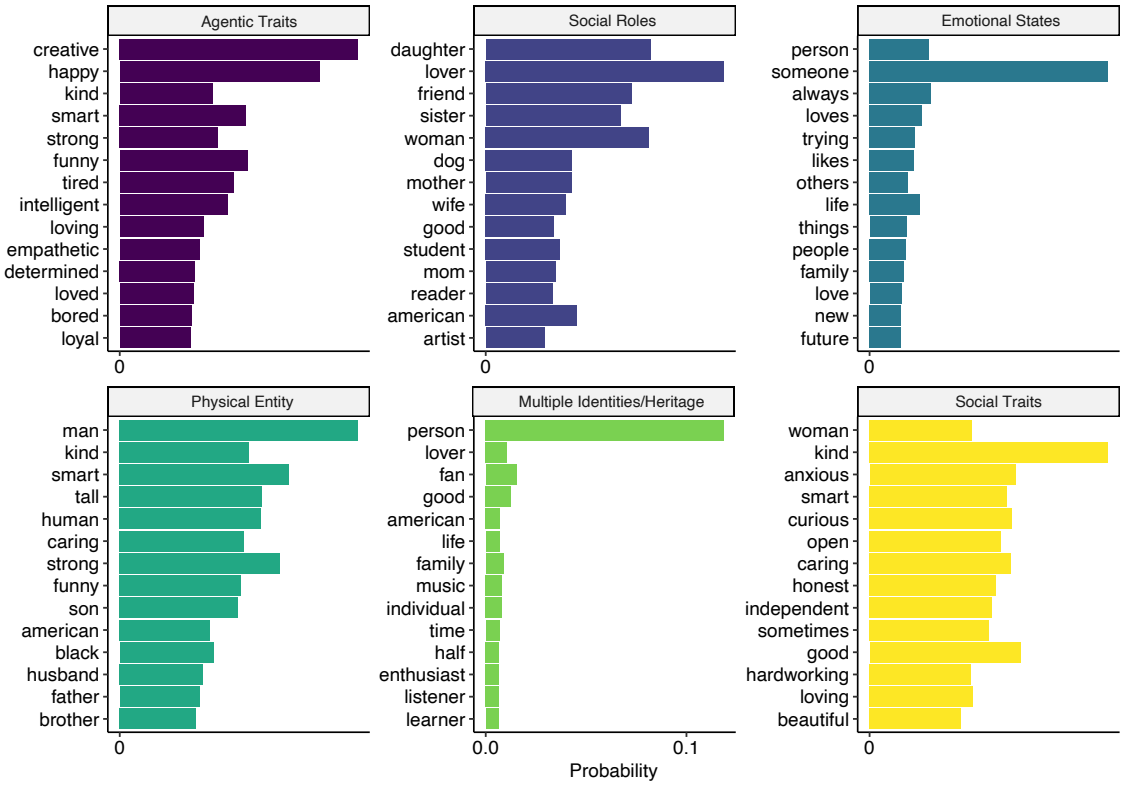


Table 3.4. Labeled STM topics in the 6-topic model with self-concept and prevalence represented as the mean (and median) gamma probability of documents.

Number	Label	Sample Self-Concept	Topic Prev.
1	Agentic Traits	I am creative, I am empowering, I am intelligent, I am beautiful, I am mesmerizing, I am quick-witted, I am easy-going, I am comedic, I am selfless, I am hopeful, I am grateful, I am inspiring, I am one in a million, I am cherished, I am competent, I am adventurous, I am flirtatious, I am charming, I am brave, I am amusing.	19 (.06)
2	Social Roles	I am a woman, I am a writer, I am a philosopher, I am a musician, I am an artist, I am a dog lover, I am a wife, I am a hard worker, I am a lover of plato, I am a lover of classic books, I am a hardcore fan of plain potato chips. I am a jewelry maker, I am a movie maker, I am a lover of the outdoors, I am a flamenco dancer, I am a crocheter, I am a painter…	.22 (.08)
3	Emotional States	I am a person who has been isolated from society, is loved by hardly anyone, and has had to try to make it in spite of that, which seems impossible. I am a person with asperger's syndrome, and the condition has deeply affected my life and my interpretation of life. I am a person who wants to succeed despite all odds, but success is seeming more and more impossible. I am an alien to the world. I am a person who has been outcasted due to asperger's syndrome, my race, my upbringing, and my physical beauty. I am struggling with life. I am confused about life…	.14 (.02)
4	Physical Entity	I am human, i am 22, i am a software developer, i am wearing clothes, i am sitting cross legged, i am an absolute unparralel genius, i am typing, i am full, i am logical, i am not religious, i am not very political, i am on a laptop, i am in a basement…i am in deep thought, i am apprehensive, i am not wearing socks unless i wear shoes	.19 (.05)

5	Multiple Heritage/Identities	I am a unique individual, defying categorization. I am a product of multiple heritages. I am a living testament to human interconnectedness. I am a complex identity shaped by various influences. I am a chameleon adapting to different environments. I am a puzzle with pieces from different backgrounds. I am a symphony of cultures harmonizing within me. I am a question mark exploring the meaning of identity. I am a blend of traditions and modern influences. I am a walking contradiction, embracing both similarities and differences. I am a mirror reflecting the complexities of society., I am a mosaic of experiences shaping my perspective. I am a global citizen with a local heart. I am a hybrid of languages, customs, and beliefs. I am a person caught between two worlds, yet finding my place. I am a symbol of hope for a more inclusive future…	.12 (.02)
6	Social Traits	I am kind, I am caring, I am loyal, I am trustworthy, I am efficient, I am loving, I am honest, I am proud, I am confident, I am curious, I am nondesicive, I am empathetic, I am driven, I am flexible, I am genuine, I am hardworking, I am independent, I am nurturing, I am open minded, I am patient.	.14 (.05)

Table 5 depicts differences in STM topic prevalence across racial groups. Prevalence in self-concepts with agentic traits varied by racial group ($F(9, 594) = 4.33$, $p < .001$, $\eta^2 = 0.06$). Individuals of Black-Latine (M = 0.27) and Black-Asian descent (M = 0.28) had the highest prevalence in agentic traits, the lowest prevalence for individuals who are Asian-White (M = 0.10) and Black/African American (M = 0.10). Prevalence in self-concept with social traits varied by racial group ($F(9, 594) = 3.61$, $p < .001$, $\eta^2 = 0.05$). Individuals who are Latine (M = 0.22) and Black-Latine (M = 0.20) had the highest prevalence in social traits, the lowest prevalence for individuals who are Asian-Latine (M = 0.08) and Asian-White (M = 0.09). These findings corroborate with human-coding: personality traits were most prevalent in dual-minority compared to majority-minority participants. Social roles varied in prevalence by racial group

($F(9, 594) = 2.07$, $p = .031$, $\eta^2 = 0.03$). Individuals of Asian-White (M = 0.29) and Asian-Latine descent (M = 0.30) had the highest prevalence in social roles, the lowest prevalence individuals who are Black/African American (M = 0.13). Self-concepts with navigating multiple identities varied by racial group ($F(9, 594) = 3.82$, $p < .001$, $\eta^2 = 0.05$). There was the highest prevalence in multiple identity self-concepts for Black/African American individuals (M = 0.30), the lowest prevalence for individuals who are of Latine (M = 0.04) and Black-Latine descent (M = 0.08). Similar to human-coding content categories, hobbies and relational roles were high prevalence in Asian-White and Asian-Latine participants, and ideology in Black/African American participants (see Appendix B).

Table 3.5. Self-concept categories, STM topic prevalence, DCS, and semantic coherence by racial group.

Mean STM topic proportions, mean similarity score within participant self-descriptors, and mean count for mention of ethnicity/race.

	Multiracial							Monoracial						
	Asian-White	Lat.-White	Black-White	Black-Latine	Asian-Latine	Asian-Black	**Total**	Black	Lat./Hispanic	Asian	White	**Total**	η^2	p-value
Content Categories														
Age	0.37	0.33	0.21	0.34	0.39	0.19	0.31	0.17	0.18	0.33	0.23	0.23	0.02	0.278
Ethnicity/race	1.39	0.93	0.82	1.13	1.41	1.28	1.12	1.33	1.21	1.43	0.35	1.08	0.03	**0.031**
Family/ relational	1.64	1.84	1.16	1	1.76	0.98	1.42	1.13	1.36	1.4	2.23	1.54	0.04	**0.001**
Gender	0.54	0.55	0.49	0.49	0.65	0.44	0.53	0.5	0.71	0.57	0.52	0.57	0.00	0.798
Goal-oriented	0.61	0.66	0.53	0.73	0.63	0.30	0.60	0.33	0.14	0.73	0.29	0.38	0.02	0.269
Hobby	2.23	1.53	1.34	1.04	1.96	1.30	1.57	1.73	1.04	1.5	1.97	1.57	0.03	**0.013**
Ideology	0.66	0.52	0.33	0.35	0.37	0.53	0.47	1.17	0.54	0.53	0.19	0.61	0.04	**0.001**
Occupation	0.83	0.90	0.76	0.58	0.96	0.40	0.76	0.87	0.54	0.90	0.90	0.81	0.02	0.177
Personality	6.38	6.69	8.93	9.07	6.91	9.70	7.8	7.4	8.82	0.83	8.06	7.76	0.06	**<0.001**
Physical Char.	0.59	0.61	0.59	0.75	0.35	0.40	0.58	0.87	0.54	0.43	0.61	0.61	0.02	0.258
Regional Id.	0.22	0.11	0.10	0.08	0.02	0.09	0.12	0.17	0.07	0.10	0.19	0.13	0.02	0.272
Self-evaluation	0.47	0.50	0.69	0.54	0.61	0.84	0.61	0.53	0.36	0.67	0.81	0.53	0.02	0.364
SES	0.17	0.17	0.11	0.19	0.09	0.05	0.14	0.13	0.21	0.13	0.13	0.15	0.01	0.710
Phys/emo state	2.32	2.62	1.59	2.73	2.30	2.37	2.31	1.63	2.36	2.57	1.32	1.96	0.03	0.051
STM Topics														
Agentic Traits	0.10	0.22	0.19	0.27	0.12	0.28	0.19	0.10	0.23	0.18	0.16	0.17	0.06	**<0.001**
Social Roles	0.29	0.23	0.20	0.18	0.30	0.17	0.23	0.13	0.21	0.22	0.21	0.19	0.03	**0.031**

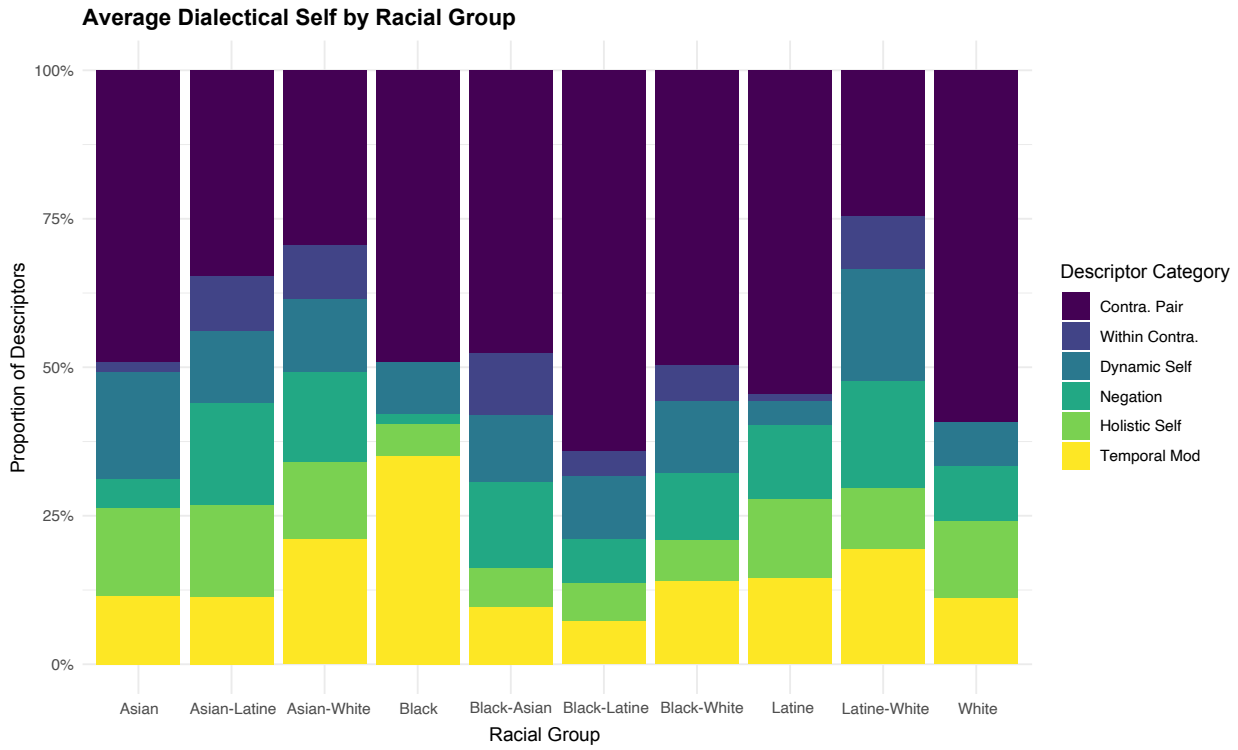
Emotional States	0.18	0.13	0.13	0.12	0.19	0.11	0.14	0.06	0.06	0.13	0.19	0.11	0.02	0.188
Physical Entity	0.19	0.17	0.19	0.16	0.20	0.20	0.18	0.28	0.24	0.19	0.24	0.24	0.01	0.591
Multiple Identities	0.16	0.14	0.09	0.08	0.11	0.09	0.12	0.30	0.04	0.13	0.11	0.15	0.05	**<0.001**
Social Traits	0.09	0.11	0.19	0.20	0.08	0.15	0.14	0.13	0.22	0.14	0.10	0.15	0.05	**<0.001**
Dialectical Coding Schemes														
Contra. pairs	0.75	0.49	1.62	2	1.13	1.37	1.18	0.93	1.89	1	1.03	1.20	0.09	**< 0.001**
Within-contra.	0.23	0.18	0.20	0.13	0.30	0.30	0.21	0	0.04	0.03	0	0.02	0.02	0.295
Negation	0.39	0.36	0.37	0.24	0.57	0.42	0.37	0.03	0.43	0.10	0.16	0.18	0.02	0.172
Holistic	0.33	0.21	0.23	0.20	0.50	0.19	0.26	0.10	0.46	0.30	0.23	0.27	0.01	0.624
Temporal	0.53	0.38	0.46	0.22	0.37	0.28	0.39	0.67	0.50	0.23	0.19	0.39	0.01	0.550
Dynamic	0.31	0.37	0.40	0.33	0.39	0.33	0.35	0.17	0.14	0.37	0.13	0.20	0.01	0.601
Sem. Coherence														
Avg.Cosine Sim. Score	0.44	0.45	0.47	0.46	0.44	0.46	0.45	0.47	0.48	0.47	0.47	0.47	0.04	**0.003**

Note. The p-values and effect sizes (η^2) for the ANOVAs are presented comparing across racial groups (Total columns not included). The p-values are bolded where the topics and features significantly differ across racial groups.

Complexity: Do multiracial individuals construct a self-concept characterized by higher semantic coherence and dialecticism?

I evaluated whether dialectical patterns universally appeared in self-concept or whether they varied by racial group. As shown in Figure 6, White and Black/African American participants do not mention any self-descriptors containing within contradiction (e.g., "I am happy yet depressed"). Also, temporal modification commonly appears most in self-concepts of Black/African American individuals. Only significant differences in contradictory self-descriptors emerged across race. Specifically, I found that Black-Latine individuals had the most contradictory self-descriptors, 2.00 (95% CI [1.65, 2.35]), followed by Latine, 1.89 ([1.28, 2.51]), Black-White participants, 1.62 ([1.30, 1.95]), and Black-Asian, 1.37 ([0.88, 1. 87]). Asian-White, 0.75 ([0.43, 1.07]) and Latine-White participants, 0.49 ([0.17, 0.80]) had the lowest contradictory self-concept. Post-hoc analysis suggested differences between multiracial participants, such that Black-Latine participants had a significantly more contradictory self-concept than Asian-White ($t(594) = 1.25$, $p < .001$, $d = 0.21$) and Latine-White participants ($t(594) = 1.51$, $p < .001$, $d = 0.26$). Similarly, Black-White participants had a more contradictory self-concept than Asian-White ($t(594) = 0.88$, $p = .007$, $d = 0.16$) and Latine-White participants ($t(594) = 1.13$, $p < .001$, $d = 0.20$). Latine monoracial participants were also more contradictory in self-concept than Asian-White ($t(594) = 1.15$, $p = .048$, $d = 0.13$) and Latine-White participants ($t(594) = 1.41$, $p = .003$, $d = 0.16$).

Figure 3.6. The dialectical self by racial group



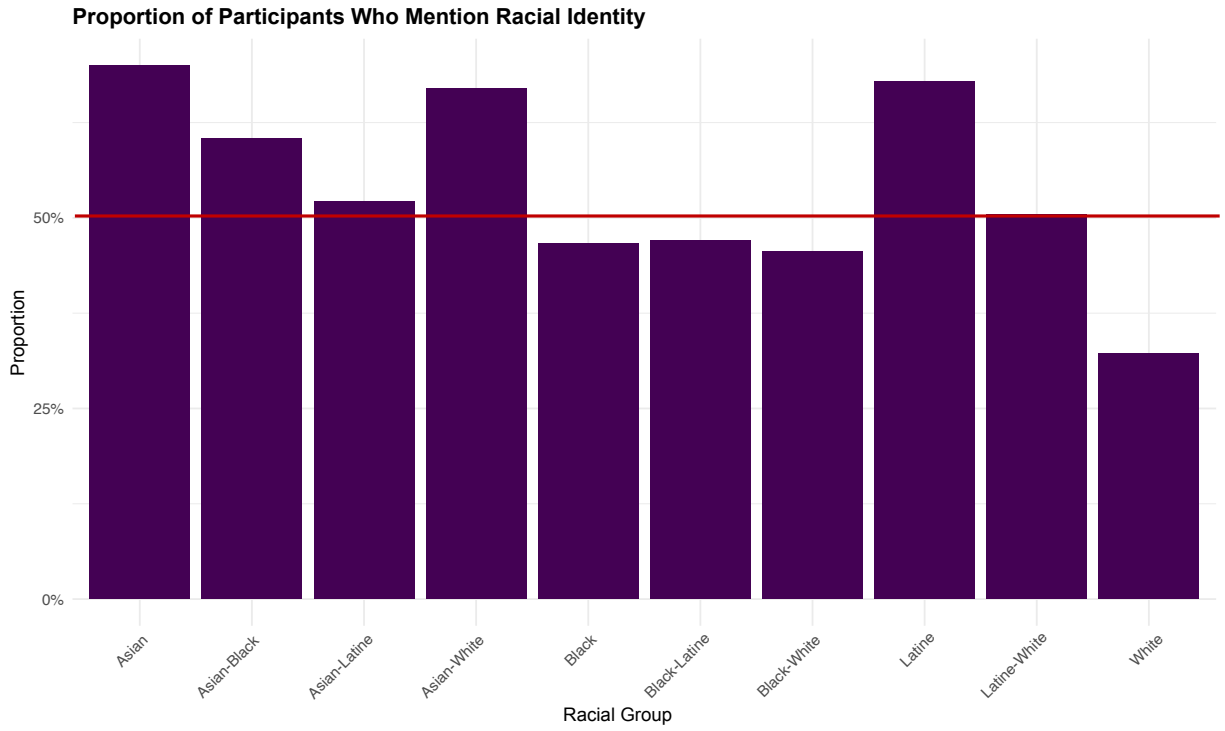
To determine whether there is greater cohesion in thematic self-concept, I used text embeddings to measure the average semantic similarity between the twenty self-descriptors within each participant. Participants of Asian-Latine and Asian-White descent have self-concepts with the lowest semantic coherence across self-descriptors (see Table 5). On the other hand, monoracial participants have the highest semantic coherence in self-concept descriptors. (Illustrative examples of both high- and low-coherence self-concepts are provided in the Supplemental Materials.)

Centrality: How does racial identity feature in self-concept?

I examined what proportion of individuals mention ethnicity/race (see Figure 7 and Table 5). Asian, Latine, Asian-White and Asian-Black participants were those most likely to mention ethnicity/race, with at least 50% of participants in those groups referring to their race. On the

other hand, fewer than 50% of the participants who are Black-White and White included ethnicity/race as part of their self-concept. Approximately half of Black-Latine participants mention ethnicity/race. Participants of only minority descent mentioned race-related self-descriptors more as part of their self-concept than individuals who are of majority-minority descent. Participants of Asian descent mentioned their racial identity as part of their self-concept the most. In contrast, White participants, except individuals of Asian descent, mentioned race as part of their self-concept the least.

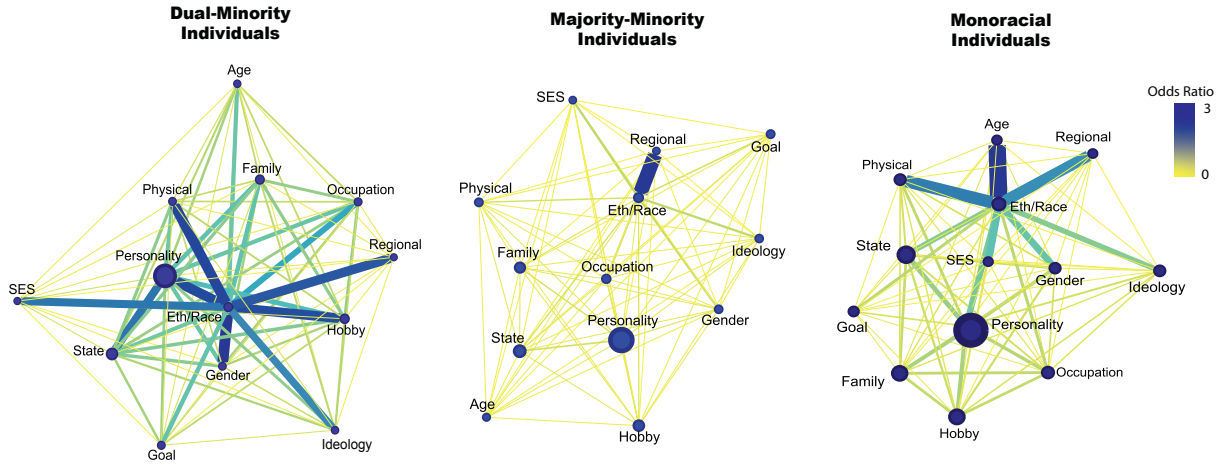
Figure 3.7. Proportion of participants who mention ethnicity/race-related self-descriptors.



Next, I examined whether people mentioned other self-concept categories—e.g., age and gender—more often with ethnicity/race; this would support conceptualizing the self across these categories. I constructed a network where each node represented a content category, and edges depicted the strength of the odd ratios between categories, how strongly the presence of one category predicted the presence of another. Thicker edges indicate stronger associations, while

node size reflected how frequently each category was mentioned overall (see Figure 8). For monoracial participants, ethnicity/race was strongly associated with regional identity, age, and physical attributes, as well as SES and gender. On the other hand, for majority-minority multiracial participants, mention of race was strongly associated with mention of regional identity. Interestingly, this differed from dual-minority participants, where mention of ethnicity/race co-occurred with, in addition to regional identity, physical attributes, hobbies, personality traits, SES, and gender. A wider variety of self-concept categories had a higher likelihood to co-occur with racial identity for dual-minority individuals compared to individuals who are part-white.

Figure 3.8. Network plots using logistic regression where the likelihood of one content category in self-concept is associated with the presence of ERI in self-concept.



Note. The thicker edge shows higher association between the categories, while the node size is how frequently that category appears in general.

Similarity: How similar are multiracial individuals' self-concepts to those of their parent racial groups, and does multiracial identity integration influence similarity?

Similarity Q1.

I wanted to determine whether racial groupings drove similarities in self-concept. I tested the hypotheses that participants were most similar to (1) those of their monoracial parent category (e.g., individuals of Black-Asian descent with others who are Black, Asian, and both), (2) only those of the same mixed-race, (3) those of a matching category type, such as dual minority, majority-minority, and majority, or (4) those of overlapping racial descent—Asian, Black, Latine, White. Hypotheses were tested by comparing the observed similarity matrix to a priori matrices representing each of the hypotheses described above. I compared the model fits across hypotheses and found that grouping peoples' self-concept across Asian and non-Asian individuals fit our data structure more than other a priori matrices (BIC = -376069.1; see Table 6). The data structure grouped across participants of the same racial combinations and their monoracial parent categories as the second best fit (BIC = -375991.9). For example, participants who are of Black-Asian descent are similar in self-concept to other individuals of this racial combination and those who are monoracial Black/African American and Asian. I ran correlations between the a priori matrices with the observed to show directional relationship. There was a significant, positive correlation for the a priori matrices of shared Asian descent ($r(602) = 0.03$, 95% CI = [.02, .03], $p < .001$) and monoracial parent categories ($r(602) = 0.02$, 95% CI = [.01, .02], $p < .001$). While dual-minority matrix was significantly correlated with the observed data (and third best fit with BIC), it was negatively associated with the observed matrix. Overall, these results suggest that self-concept similarities are more strongly influenced by overarching racial group distinctions, specifically between Asian and non-Asian identities, rather than strictly

by shared multiracial identity or specific racial combinations. Furthermore, the grouping by overlapping parent categories highlight the unique identity patterns within mixed heritage individuals (BIC = -375991.9).

Table 3.6. Best fitting model by racial group framework.

Ethnicity/Race	BIC	Difference
Asian Descent	-376069.1	-
Parent Category Match	-375991.9	77.2
Dual, majority-minority, majority	-375965.2	26.7
Same Racial Combination	-375955.1	10.1
Black Descent	-375954.3	0.8
Latine Descent	-375944.8	9.5
White Descent	-375944.0	0.8

Note. Racial group framework, sorted from smallest to largest. There is the difference between each model and the preceding best model.

Similarity Q2.

A multilevel model was fit to determine whether the more participants' racial categories were integrated, the more similar their self-concepts were to people of those parent categories. Multiracial identity integration was positively associated with how similar they were to people of their parent categories, $b = 0.01$ (95% CI [0.00, 0.02], $p = .009$). The main effect of parent ethnicity/race was also significant ($F(3, 481) = 253.57$, $p < .001$), indicating that multiracial individuals' average similarity varied across monoracial parent groups. Specifically, multiracial participants showed the highest similarity scores when matched with monoracial Latine ($M = 0.52$, 95% CI [0.52, 0.53]) and Asian ($M = 0.52$, [0.52, 0.53]) participants, followed by White ($M = 0.51$, [0.51, 0.52]), and the lowest similarity when matched with monoracial Black/African American participants ($M = 0.50$, [0.50, 0.51]).

Discussion

This study investigated how multiracial individuals construct their self-concept across key psychological dimensions. There was systematic variation in self-concept content by racial background: monoracial participants often emphasized ideology, whereas multiracial participants highlighted more individualized markers such as personality traits and hobbies, among dual-minority and majority-minority subgroups, respectively. Patterns of complexity also differed: Black-Latine participants exhibited the highest levels of internal contradiction, while Asian-White and Latine-White individuals showed the lowest semantic coherence. Notably, racial identity was most salient for participants of Asian descent, suggesting group-based differences in identity centrality. At the structural level, similarity analyses revealed broad clustering along an Asian versus non-Asian axis, yet multiracial individuals who identified strongly with both parent groups also resembled their monoracial counterparts.

When invited to define themselves in their own words, people draw from a wide array of identity domains (e.g., personality, social roles, values, emotions, physicality, and heritage) to build a portrait of self that is at once personal and socio-cultural. These results suggest that self-concepts are not simply lists of traits or roles, but structured systems with varying degrees of coherence, contradiction, and semantic overlap across individuals. These differences were patterned by race, gender, and multiracial configuration, yet they also point to broader psychological processes involved in identity construction (e.g., integration). By mapping both the content and structure of self-descriptions, this study extends identity frameworks that have traditionally focused on stable traits. The *general discussion* will highlight how patterns of self-concept and integration differ across monoracial and multiracial participants, and between ethnic/racial backgrounds. Building on racial and multiracial differences, the next section turns to gender as a critical axis through which self-concept organization and expression diverge.

Intersectionality beyond racial identity

Gender differences in self-concept content provides further insight into how intersecting social categories shape identity expression. These findings revealed distinct patterns across gender groups, indicating that the structure and salience of identity domains are also informed by gender norms. Social roles and personality traits were more prevalent in women's self-descriptions, consistent with longstanding evidence that women tend to employ more communal language when describing themselves (Cross & Madson, 1997; Kashima et al., 1995). This aligns with broader socialization theories that emphasize interdependence and relational orientation in women's self-construals (Guimond & Chatard, 2014; Impett & Peplau, 2006). However, contrary to prior research, agentic attributes (typically associated with assertiveness and autonomy) were more frequent among women than men in our sample. This divergence may reflect evolving gender norms or greater flexibility in how women articulate their identities in the contemporary sociocultural context.

Men's self-concepts, by contrast, were more likely to include concrete and factual descriptors, consistent with Newman et al.'s (2008) finding that men's language skews toward physical objects and less personal content (see also Eagly & Revelle, 2022). Our STM-derived topic *Physical Entity*—which captured literal and descriptive language (e.g., "human," "tall," "man")—was more prevalent in male participants, further supporting the association between men and a factual or object-based linguistic style.

In addition to these binary comparisons, our results offer novel insight into nonbinary individuals' identity construction. Emotional language and references to *Multiple Identities* appeared most frequently in their self-concept typologies. These patterns may reflect an intentional departure from rigid, essentialist identity categories. Consistent with prior research (Richards et al., 2016), nonbinary individuals described their gender in fluid terms, suggesting a

broader resistance to fixed social classifications that extends into other domains of self-definition.

From Conceptualizing the Self to Navigating Hybridization

The distinctions observed across multiracial subgroups support a shift away from essentialist notions of "the multiracial experience." Instead, self-concept emerges as a situated response to the affordances and constraints of specific cultural, racial, and social positions. These findings challenge the notion of coherence as a singular ideal, instead pointing to multiple valid forms of self-organization, ranging from ideologically anchored to highly individualized configurations.

This study highlights how multiracial individuals construct their self-concept, revealing that identity is about what they say and how they organize and express those meanings across contexts. Ultimately, these findings support framing the self as an iterative process, especially for those at the intersection of multiple ethnic/racial worlds.

Building on these insights, the next study shifts from cognitive representation to process, examining how individuals actively *navigate* their hybridized identities. Whereas Study 2 captured the structure and content of self-concept as expressed in language, Study 3 explores the strategies people use to reconcile, integrate, and communicate across cultural boundaries. This focus on navigation extends the theoretical model of the self as dynamic and negotiated, tracing how multiracial individuals translate internal complexity into everyday acts of meaning-making and belonging.

CHAPTER IV
Study 3:
Navigating Multiracial Identity across Environment, Cultures and Psychological Processes

Introduction

Multiracial individuals are an emergent and expanding population that increasingly reflect the complexity of ethnicity/race in contemporary society. As the number of people identifying with more than one racial background continues to rise, understanding how these individuals construct and sustain a coherent sense of identity across multiple cultural frameworks has become increasingly important. Navigation strategies represent the deliberate actions and reflective techniques individuals use to manage, assert, or reconcile their racial identities in diverse settings, emphasizing the role of personal agency in shaping one's racial self-concept (Tsai et al., 2024). These strategies involve both internal processes, such as integrating multiple heritages into a coherent sense of self, and external actions, such as choosing how to identify or behave in racially coded contexts. Through these strategies, multiracial individuals continuously construct meaning around who they are and how they are seen.

Unlike monoracial individuals, who may experience relative stability in racial categorization, multiracial people often encounter shifting racial ascriptions across settings, requiring ongoing negotiation of who they are and how they are perceived (Rockquemore & Brunsma, 2002; Root, 1996). These contextual shifts, across family, school, work, and community environments, demand adaptive strategies to maintain coherence and authenticity. For this reason, navigation strategies are not merely reactive coping mechanisms but *developmental tools* that allow multiracial individuals to exercise agency within social systems that constrain or invalidate their identities.

Navigation strategies, such as frame-switching, strategic self-presentation, and learning about heritage, are central to this task of multiracial identity development (e.g., Norman & Chen, 2019; Shih et al., 2007; Tsai et al., 2024). Applying these strategies allows multiracial individuals to assert control over their racial identity formation and presentation (Noriega, 2016 ; Ortiz, 2017; Tsai et al., 2024). These strategies are particularly salient in response to everyday racial microaggressions, such as identity denial (e.g., "You don't look [race]") and identity questioning (e.g., "What are you?") (Nadal et al., 2011; Tran et al., 2015), By employing different identity management approaches, multiracial individuals can cultivate a sense of authenticity while adapting to shifting racial expectations imposed by society.

The adoption of navigation strategies is important, but the parameters around use of these strategies has yet to be delineated. Moreover, prior work has relied on relatively small samples and insufficient sampling of more specific categories of multiracial people (e.g., dual-minority). Therefore, the extent to which strategies are universal or used primarily by individuals of specific racial backgrounds is unclear. In the present study, I will describe the use of navigation strategies by multiracial individuals in relation to personal and environmental factors, as well as effects on well-being. In this process, I conceptualize navigation strategies as profiles that reflect different modes of managing racial identity. I examine these profiles in relation to four interrelated domains that together capture the social, psychological, and experiential dimensions of identity:

(1) sociocultural factors, or the individual and community characteristics that shape opportunities and pressures for identity expression;

(2) identity processes, the cognitive and affective mechanisms through which individuals integrate and make sense of their racial backgrounds;

(3) well-being, as a key psychological outcome linked to how successfully individuals navigate identity-based challenges; and

(4) liminal life events, or culturally significant experiences that bring racial identity to the forefront of self-understanding.

Each of these domains highlights a different facet of identity navigation, from its social ecology to its emotional and developmental consequences. Addressing these questions requires contextualizing individual narrative accounts with a more systematic understanding of navigation strategy profiles, the constellations of behaviors and attitudes through which individuals negotiate multiracial identity in everyday life. By situating navigation strategies within this broader framework, the study aims to capture how multiracial individuals actively construct, maintain, and refine their racial identities within sociocultural norms.

Measuring Navigation Strategies

A critical first step in understanding multiracial identity is to identify the concrete strategies individuals use to navigate their racial backgrounds. While prior research has identified identity negotiation and coping strategies in response to racial invalidation (e.g., Shih & Sanchez, 2005; Sanchez et al., 2009), and developmental work has explored how exploration and cultural socialization shape identity (e.g., Gaither, 2015; Syed & Azmitia, 2008), these behaviors have rarely been quantified in large, diverse multiracial samples.

Understanding these strategies is essential because they reflect agency in the face of structural constraints. Some strategies serve adaptive functions, such as cultivating cultural fluency, while others may manage identity tension or marginalization. Identifying which strategies are common or distinctive across groups provides insight into how multiracial individuals maintain psychological stability. For instance, some research finds that perceived

discrimination is positively associated with stronger multiracial identification (Norman & Chen, 2020), suggesting strategies may differ based on an individual's specific racial background and experiences.

The first research question therefore asks: *What strategies do multiracial individuals use to navigate their racial identities, and how do these strategies co-occur as patterns of identity management?* Addressing this question establishes the empirical foundation for understanding how navigation operates within an individual.

Sociocultural Factors

Broader social and structural contexts shape how individuals learn about, express, and reconcile their racial identities. Individual and community-level factors (i.e., racial appearance, genealogical background, socioeconomic resources, and exposure to racial diversity) create opportunities and constraints that influence the strategies people adopt (Rockquemore et al., 2009; Shih et al., 2007; Townsend et al., 2012). Research suggests that multiracial identity work varies across gender: women are more likely than men to be aware of their multiracial identities, and women who are first generation multiracial were more likely to self-identify as multiracial (Xu et al., 2021). Navigation strategies may also differ by age, in that older people may have more navigation strategies to access.

Racial dynamics and socialization exist within a family: some emphasize learning about multiple racial backgrounds, while others deliberately de-emphasize race to promote colorblind belonging (Shih et al., 2007). Phenotype also shapes the way people experience multiracial identity, such that individuals who appear racially ambiguous or are frequently misclassified are more likely to identify as multiracial and may rely on strategies that assert authenticity or cultural knowledge (Ortiz, 2017). Socioeconomic context matters as well. Multiracial individuals

from higher-SES backgrounds are more likely to claim a multiracial identity, while those from lower-SES groups may strategically identify monoracially to mitigate stigma (Townsend et al., 2012). Finally, community-level characteristics, such as neighborhood diversity and local racial attitudes, can amplify or constrain the salience and expression of these strategies (Morning & Saperstein, 2018; Rogers et al., 2021).

These factors determine which strategies are available and adaptive in a given environment. This leads to a second research question: *How do sociocultural factors – at the individual, family, and community levels – predict the use of specific multiracial navigation strategies?*

Identity Processes

Identity navigation is as much a psychological process as a behavioral one, reflecting the internal work of making meaning from multiple racial backgrounds. These underlying processes determine whether individuals experience their identities as conflicting, complementary, or seamlessly blended, shaping how they actively construct a sense of self (Amiot et al., 2007; Syed & Azmitia, 2008). To understand *why* certain strategies are deployed, it is essential to link them to core identity processes like *exploration* (the active contemplation and investigation of one's heritage), *integration* (the perceived compatibility or conflict between identities), and *identity styles* (the characteristic ways one maintains their identity, from hybridizing cultures to alternating between them contextually) (Lilgendahl et al., 2018; Ward et al., 2018)

These processes provide the cognitive and motivational mechanisms for behavioral strategies. For instance, an individual high in identity exploration may more frequently use navigation strategies, while someone with an alternating identity style would naturally employ frame-switching. Linking strategies to these processes clarifies whether they reflect adaptive

identity development or are primarily coping responses to racialized pressures (Cheng & Lee, 2009).

I ask in the third research question: *How are navigation strategies associated with underlying psychological identity processes, specifically exploration, integration, and identity style, in shaping multiracial identity?*

Well-Being

The strategies multiracial individuals use to navigate their identities have important implications for psychological well-being. Actions such as asserting authenticity or exploring heritage can function as adaptive tools for maintaining coherence and belonging (Gaither, 2015; Syed & Azmitia, 2008). Yet, these same behaviors may also operate as coping responses to social exclusion or identity invalidation (Sanchez et al., 2009; Shih & Sanchez, 2005; Tran et al., 2016). Determining whether navigation strategies are protective or symptomatic of distress is therefore critical for understanding their psychological consequences.

Certain navigation strategies may facilitate higher well-being. With evidence that multiracial individuals often report greater exclusion (Sanchez & Bonam, 2009; Shih & Sanchez, 2005) and higher depressive symptoms (Miller et al., 2019) relative to monoracial peers, awareness of discrimination can catalyze feelings of solidarity among multiracial individuals (Giamo et al., 2012) and heighten the desire for accurate social recognition of their identities (Remedios & Chasteen, 2013). Therefore, identifying as "multiracial" may be associated with higher well-being. Individuals who feel compelled to prove their legitimacy by demonstrating racial knowledge may be reacting to such invalidation. Alternatively, this same behavior may function as armor, an assertive form of identity management that bolsters self-concept and protects against psychological harm (Shih et al., 2007).

This tension extends to the nature of the strategies themselves. Approaches grounded in curiosity and learning, such as exploration of heritage, are linked to healthier adjustment and resilience (Morrison, 1995). Whereas avoidance-oriented strategies, such as minimizing racial discussions or disengaging from racial identity altogether, may offer short-term relief while undermining longer-term coherence and belonging.

This leads to the fourth research question: *Are specific profiles of navigation strategy use associated with higher or lower levels of psychological well-being among multiracial individuals?*

Liminal Life Event

Beyond the psychological and sociocultural correlates of navigation, it is equally important to understand when and under what conditions identity navigation becomes most visible. Moments of transition, what anthropologists refer to as *liminal life events*, offer a theoretically powerful lens for observing navigation strategies in practice (van Gennep 1960). These are periods of social or developmental transition in which established identity boundaries are destabilized, prompting individuals to renegotiate how they understand and present themselves across racial contexts.

For multiracial individuals, such events often foreground issues of belonging, authenticity, and recognition. Culturally defined rituals, relational turning points, and other life transitions make racial identity especially salient, requiring individuals to decide which cultural frameworks to emphasize and how to reconcile competing expectations. In these moments, navigation strategies are not hypothetical or abstract, they are enacted in real time as individuals engage in the cognitive and emotional work of aligning internal identification with external perception.

As Bauer and McAdams (2004) note, life transitions are psychologically generative because they invite narrative reconstruction: people reinterpret past experiences and integrate new ones into their evolving sense of self. For multiracial individuals, these passages often magnify the developmental "work" of identity, integrating multiple heritages, managing dissonance, and transforming tension into coherence. Thus, examining narratives of liminal events provides insight into the mechanisms through which multiracial individuals sustain a coherent and agentic sense of self within dynamic social landscapes.

In this study, I analyze participants' narratives of identity-relevant life events as texts that reveal both the *content* and *emotional tone* of multiracial identity work. Structural Topic Modeling (STM) is used to identify recurring thematic patterns, what individuals emphasize when describing these events, and sentiment analysis assesses how these experiences are evaluated affectively, from affirming to painful or ambivalently mixed. Together, these approaches capture the psychological and emotional aspects of navigation as it unfolds within culturally significant life transitions.

Accordingly, the final research question asks: *How are navigation strategies enacted and expressed when multiracial individuals describe liminal life events that engage multiple racial identities?*

The Present Study

The present study investigates how multiracial individuals navigate their racial identities across social, psychological, and experiential domains. Conceptually, treating navigation strategies as profiles across item endorsement allows for a person-centered understanding of multiracial identity management. This approach reveals which strategies individuals use and how these strategies coalesce into distinct modes of navigating racial identity across contexts. Across

four interconnected domains, the project examines the architecture of multiracial identity navigation. The first domain situates navigation within its sociocultural ecology, exploring how structural and interpersonal contexts, such as racial appearance, community diversity, and socioeconomic background, shape which strategies are available or endorsed. The second domain turns inward to identity processes, linking navigation profiles to mechanisms of exploration, integration, and identity style that sustain coherence. The third domain focuses on well-being, testing whether certain patterns of navigation, such as higher endorsement, are linked to greater psychological health or a consequence of identity negotiation. Finally, the fourth domain examines how navigation strategies are enacted and narrated through liminal life events, moments of transition that bring multiple racial identities into contact and reveal the developmental process of meaning-making.

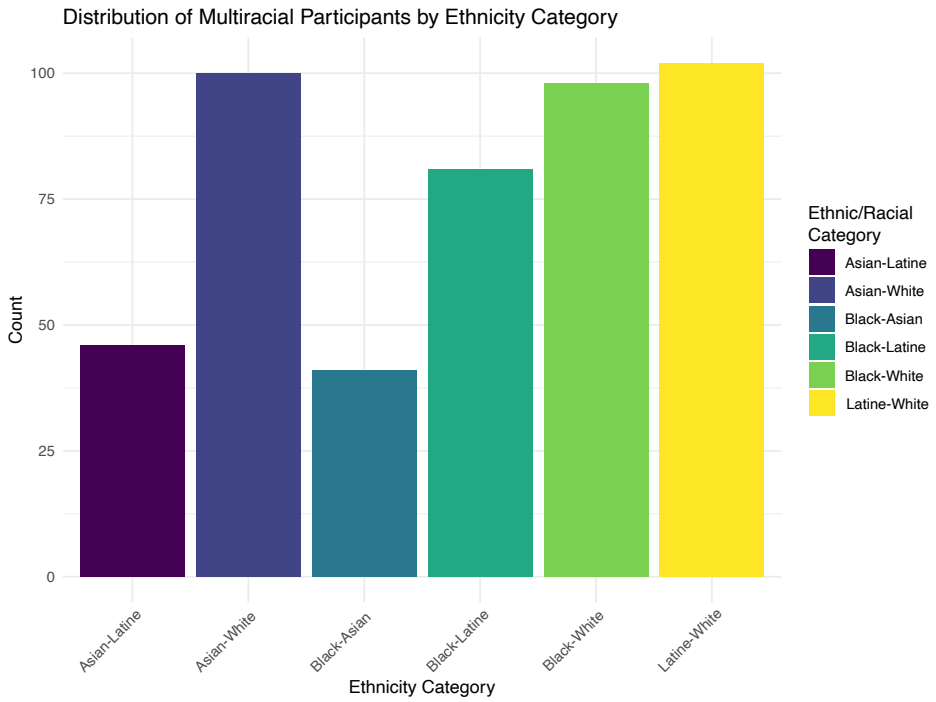
To address these aims, the study integrates qualitative and quantitative approaches. Open-ended narratives provide a grounded view of how individuals describe their lived strategies in their own words, while statistical modeling identifies broader patterns of strategy co-occurrence and their links to sociocultural context, identity processes, and well-being. Together, these domains construct a multilevel framework for understanding multiracial identity as both *agentic* and *contextually situated*. By integrating sociocultural context, internal process, subjective well-being, and narrative meaning, the study moves beyond static conceptions of racial identity toward a dynamic model of navigation, one that captures how multiracial individuals actively construct, maintain, and make sense of who they are across shifting racialized landscapes.

Methods

With the sample from Study 2, I exclude monoracial participant and participants who did not complete the navigation strategy items, resulting in a final sample of 468 participants (see

Figure 1). Ethnic/racial demographic is the following: 102 Latine-White, 100 Asian-White, 98 Black-White, 81 Black-Latine, 46 Asian-Latine, and 41 Asian-Black participants. There are 272 women (58%), 180 men (39%), and 16 nonbinary persons (3%).

Figure 4.1. Demographic breakdown of multiracial sample by ethnicity/race



Measured Variables

Navigation Strategies. Participants responded to the open-ended item: "How do you find balance or coherency between your different ethnic/racial identities?" This was complemented by our follow-up question, a close-ended item: "When trying to find balance or coherency in your racial identity, which of the following do you think you have done in the past?" Participants selected all that applied from the following list:

Table 4.1. Close-ended navigation strategies.

Navigation Strategy	Item(s) to Select
Self-definition	Identify as "mixed"
	Identify as "monoracial"
	Disidentify with race altogether
Selection	Pick and choose only certain customs that feel appropriate to practice
Learn	Make time to learn about your racial identities in your own free time
Shared Feathers	Prove yourself as a member by showing knowledge about racial community
Experiment	Experiment with your appearance
Solidarity	Share experiences with others who feel outgroup
Frame-Switch	Change your behavior or way of speaking to fit different social or cultural settings
Ambiguity Tolerance	Be open to ambiguity

Sociocultural Factors. Participants reported their self-identified ethnicity/race, gender, age, generational status, reflected race, generational loci. For reflected race, participants were asked, "What race/ethnicity do you believe you are perceived as by others?" Open-ended responses were categorized using regular expressions (regex), a systematic text-searching method in R that identifies words or phrases matching specified patterns (e.g., "mixed," "multiracial," "depends"). This approach allowed for transparent, replicable coding of responses into four categories reflecting perceived racialization: context-dependent (perceived race varies by context), different (perceived as a race different from either self-identified race), mixed (perceived as multiracial), and one (perceived as only one of their racial identities). For example, if a participant wrote that they're perceived as multiracial or mixed, they would be categorized as "mixed." If they listed both of their racial identities separately, they were categorized as both. Participants also completed the Collective Self-Esteem Scale (Luhtanen & Crocker, 1992), an 8-

item measure of participants' perceptions of how their racial group is viewed by others (societal attitudes). Several indicators captured participants' socioeconomic context: household size; educational attainment; perceived peer relative status, assessed with the item, "Compared to your peers, would you say that you were as well off?" (0 = No, 1 = Yes), with higher values reflecting greater perceived economic parity or advantage relative to peers; and food resource insecurity, measured with the item, "How often did you go hungry or not get a meal before the age of 18?" (1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Often, 5 = Always). Because predictors varied in scale and measurement, ordinal variables (e.g., education, generational status) were modeled as monotonic effects, and continuous variables (e.g., age and societal attitudes) were standardized to improve comparability and interpretability across predictors.

Identity processes.

Identity integration was measured using a modified Bicultural Identity Integration (MII) scale (BIIS-1; Benet-Martínez & Haritatos, 2005), in which the term "bicultural" was replaced with the word "multiracial". In some cases, the wording was adjusted to be more general (e.g., the item, "I find it easy to harmonize ___ and American cultures," was changed to "I find it easy to harmonize my racial identities." Participants reported agreement with items on a scale from 1 ("strongly disagree") to 5 ("strongly agree"). The scale demonstrated good internal consistency, ω_h[13] = 0.86 (M = 3.64, SD = 0.67, range = 1.29–5.00).

Hybridized identity styles was assessed using the Multicultural Identity Styles Scale (MISS), also completed on a scale from 1 ("strongly disagree") to 5 ("strongly agree"), with moderate internal consistency, ω_h = 0.48 (M = 3.55, SD = 0.66, range = 1.92 –5.00). While the

[13] I apply Omega Hierarchical for BII and MISS because BII has two dimensions—harmony/conflict and blended/compartmentalized—with identity integration as the general factor. MISS has two dimensions—hybrid and compartmentalized identity styles—with identity style as the general factor.

MISS is thought to be unidimensional, I found that the items loaded onto the two factors; accordingly, the general factor only accounted for roughly half of the reliable variance in the total score. Using the *unidim* statistic (Revelle & Condon, 2023) to evaluate unidimensionality, I determined that each subscale represents a distinct construct. The unidimensionality index considers the extent to which a set of items measures a single latent factor. A high index suggests that most of the shared variance among items is driven by one dominant dimension, indicating strong item homogeneity with respect to the construct. Specifically, the blended identity styles subscale exhibited a unidimensionality index of 0.98, and the alternating identity styles subscale exhibited a unidimensionality index of 0.94. The two subscales as treated as separate constructs in subsequent analyses. Items were from cultural to racial identities, for example, "some situations make it hard to be ____ and American at the same time" to "some situations make it hard to be multiple racial identities at the same time." MISS item 6 was removed since it didn't make sense upon modification: "I am Hispanic in an American way" would've been, for example, "I am Asian in a Latine way," such that the order of identity mattered.

Lastly, for *identity exploration,* I used the 5-item exploration subscale in the Multigroup Ethnic Identity Measure (MEIM) on a scale from 1 ("strongly disagree") to 4 ("strongly agree"). This has been modified to account for exploring two racial identities rather than one. The following items appear for each racial identity:

1. "I have spent time trying to find out more about my _____ identity, such as its history, traditions, and customs."
2. "I am active in organizations or social groups that include mostly members of the _____ community."

3. "I participant in cultural practices of my _____ identity, such as special food, music or customs."

The scale demonstrated good internal consistency with a unidimensionality index of 0.72 (M = 2.72, SD = 0.50, range = 1.00 – 4.00).

Life Event. Anthropologists define a rite of passage as the liminal crossing of borders into a culturally different realm, where rites can be those of initiation or detachment (Turner 1969; van Gennep 1960). I focus on rites of passage because they often draw on culturally defined scripts. Participants were asked to describe a life event where their racial identities were salient.

> Part I: "Think of an event or episode that has marked a transition into a new life phase and that has involved **two or more** of your ethnic/racial identities. This could be a memory of any kind of experience—positive, negative, or mixed— the important thing is that it relates to how you understand yourself as a multiracial person.
>
> Tell us about the event in as great detail as possible."
>
> Part II: "Reflecting on the memory—listed above—what were your thoughts and feelings during the event, and why was it significant to you? In addition, please comment on the long-term impact of the event on your sense of who you are (if any) and what this event means to you today."

In a preliminary study, 80 (36% dual-minority) multiracial participants recruited through Prolific (52% women; age range = 18-29) provided engaging, thoughtful responses to the narrative prompt.

Well-being. Well-being was measured using the 5-item Well-Being subscale from the SAPA Personality Inventory (SPI; Condon, 2018).

Analysis Plan

RQ1: What are navigation strategies endorsed by participants?

To identify and characterize patterns of how multiracial individuals navigate their ethnic/racial identities, I used two complementary approaches:

(1) Qualitative coding of open-ended responses to identify and describe navigation strategies, and

(2) Quantitative latent class analysis (LCA) of close-ended strategy endorsements to model distinct profiles of strategy use.

The open-ended responses provided a bottom-up, descriptive account of participants' identity navigation in their own words, while the LCA offered a data-driven, person-centered model of strategy co-occurrence. These approaches were later compared to examine whether individuals of certain latent profiles endorsed open-ended strategies and how both varied across ethnic/racial subgroups.[14]

Open-Ended Navigation Strategies

Participants described, in their own words, how they navigate their multiracial identity in everyday life. Following thematic analysis procedures (Braun & Clarke, 2006), responses were first coded inductively to capture the range of strategies participants articulated. Codes were iteratively refined into overarching themes once thematic saturation was reached, defined as the point at which no new strategy types emerged (Saunders et al., 2018). Coding proceeded in increments: an initial random subset of 20 responses was coded, followed by additional batches

[14] Initially, I planned to conduct a latent class analysis combining close- and open-ended navigation strategy indicators for an integrated profile of self-reported and self-described strategies. However, the resulting two-class solution was non-interpretable: one class reflected near-complete endorsement of all close-ended strategies and minimal open-ended strategies, while the other showed the reverse pattern. For transparency, results from this exploratory model are included in the Appendix. Instead, I look at the open-ended strategies separately.

of ten until saturation was achieved. Thematic saturation was reached at 30 narratives, with no additional themes emerging even onto 50 coded responses.

In the second, deductive phase, themes were situated within existing frameworks of multiracial identity development (e.g., Poston, 1990; Renn, 2004; Root, 1996; Rockquemore & Brunsma, 2002) and prior work on identity navigation strategies (e.g., Shih et al., 2007; Norman & Chen, 2019; Tsai et al., 2024). The final codebook documented each strategy's definition, decision rules, and representative excerpts.

Four trained research assistants coded the open-ended responses after a two-week training led by the first author. I recognize that analyses of multiracial identity are unavoidably shaped by the social positions, experiences, and interpretive lenses of the research team. Reflexivity and disclosure of positionality are essential for transparency and for situating how researcher perspectives may influence coding and interpretation. The team conducting this study comprised women: Filipina, Black/African American, White, and multiracial. To ensure consistency, approximately 100 simulated practice responses were generated for calibration. Each participant response was double-coded by two independent raters (234 responses per rater). Because the coding categories were categorical, interrater agreement was assessed using Cohen's κ (Cohen, 1960). Agreement was moderate across the two RA pairs ($\kappa_1 = .68$; $\kappa_2 = .57$). Lower agreement in the second pair likely reflected the several responses that plausibly fit more than one navigation strategy category. The first author served as a tie-breaker when discrepancies occurred between coders. Because several strategies can co-occur within a single narrative, categories were defined to allow overlap rather than forcing mutual exclusivity. Disagreements therefore concentrated at category boundaries and were resolved by adjudication (first author) using decision rules (see Appendix C). This approach prioritizes representational accuracy over

artificially inflating agreement. I re-read all responses for which more than two codes were initially assigned to determine whether multiple navigation strategies were appropriately represented. This process also highlighted the importance of including more ambiguous or overlapping narratives in future coder training sets to better capture the nuanced ways individuals describe their identity navigation experiences.

Close-Ended Navigation Strategy Profiles

Latent class analysis (LCA) was conducted to identify unobserved subgroups (latent classes) of participants based on patterns of navigation strategy endorsement. LCA is a person-centered, finite mixture modeling approach that estimates the probability that each participant belongs to a given latent class using maximum likelihood estimation (Collins & Lanza, 2009 ; McCutcheon, 1987). This method assumes that the population is composed of a finite number of mutually exclusive classes that differ in their conditional probabilities of endorsing each strategy item. In other words, LCA simultaneously estimates (1) class membership probabilities, or the proportion of the sample likely to fall into each class, and (2) item-response probabilities, which reflect the likelihood of endorsing a given strategy conditional on class membership. Together, these parameters provide a probabilistic representation of heterogeneity in navigation behaviors rather than relying on deterministic group assignment.

Models specifying two to six classes (K = 2–6) were estimated using 20 random starts and up to 10,000 iterations to ensure convergence stability and avoid local maxima, a known challenge in mixture modeling where multiple likelihood solutions may emerge (Nylund et al., 2007). Model fit was evaluated using several information criteria: the Akaike Information Criterion (AIC; (Akaike, 2003), which assesses model fit while penalizing for the number of estimated parameters; the Bayesian Information Criterion (BIC; Schwarz, 1978), which imposes

a stronger penalty for model complexity and tends to favor more parsimonious models; and the Sample-Size Adjusted BIC (SABIC; Sclove, 1987), which moderates BIC's conservatism by adjusting for sample size. Lower values on each criterion indicate better balance between fit and parsimony.

Classification precision was assessed using the Average Posterior Probability of Assignment (APPA; Nagin, 2005), representing the mean probability that individuals are correctly classified into their most likely class. APPA values approaching 1 indicate clear class separation, whereas values below .70 suggest poor discrimination between classes. Item-level conditional probabilities were examined to ensure that classes were meaningfully differentiated by distinct endorsement patterns rather than overall response intensity. Model selection emphasized both statistical fit and conceptual interpretability. The optimal solution was defined by lower AIC, BIC, and SABIC values, acceptable classification accuracy (APPA ≥ .80), and theoretically coherent profiles. These criteria jointly ensure that classes capture substantively distinct patterns of strategy endorsement rather than statistical artifacts of model complexity. With classification accuracy (APPA ≥ .80), participants will be assigned to their most likely latent class (i.e., hard class assignment) based on maximum posterior probability from the LCA. This approach ensures that each participant is placed in the class to which they most probabilistically belong, thereby facilitating clearer group comparisons.

The resulting class membership variable will then be used in subsequent analyses to examine how navigation strategies profiles (as represented by the latent classes) relate to other key outcomes, such as measures of sociocultural context, identity integration, and well-being. In doing so, these follow-up analyses will help clarify the psychological and social correlates of each strategy profile and the broader implications for multiracial identity development.

RQ2.1: Do navigation strategies vary across individual & community characteristics?

To examine whether individual and community-level sociocultural factors predicted variation in navigation strategy class membership, I estimated two Bayesian multinomial logistic regression models using participants' endorsed navigation strategies as the outcome variable. Bayesian modeling provides a flexible inferential framework that quantifies uncertainty through full posterior distributions, allowing direct probability statements about model parameters. This approach is particularly advantageous when dealing with moderate sample sizes and unbalanced groups, as it incorporates prior information to stabilize estimation without biasing results.

Each model predicted the log-odds of belonging to a navigation strategy profile. The general form of the multinomial logistic model for participant *i* in class *k* (k = 2, 3) can be expressed as:

$$\log\left(\frac{P(Y_i = k)}{P(Y_i = 1)}\right) = \beta_{0k} + \sum_{j=1}^{J} \beta_{jk} X_{ij} + \varepsilon_{ik},$$

where $Y_i$is the observed class membership, $P(Y_i = 1)$corresponds to the probability of belonging to a strategy profile and X_{ij}represents the sociocultural predictors. Separate models were specified for individual-level and community-level variables, each estimated via Hamiltonian Monte Carlo sampling implemented in the *brms* package (Bürkner, 2017) interfacing with *Stan*.

Individual-Level Model

The individual-level model included gender, ethnic/racial combination, phenotype alignment, multiracial parental background, education, and generational status as predictors of navigation strategy class membership. Education and generational status were treated as

monotonic ordinal effects, preserving the rank-order nature of these variables without assuming equal distances between categories. Age was standardized. The model was defined as:

$$\log\left(\frac{P(Y_i = k)}{P(Y_i = 1)}\right) = \beta_{0k} + \beta_{1k}(\text{Gender}_i) + \beta_{2k}(\text{Ethnicity}_i) + \beta_{3k}\left(\text{Age}_i\right) + \beta_{4k}(\text{Phenotype}_i) + \beta_{5k}(\text{GenerationalLoci}_i) + \beta_{6k}(\text{Education}_i) + \beta_{7k}(\text{GenerationalStatus}_i) + \varepsilon_{ik}.$$

I adopted weakly informative priors to provide mild regularization while minimizing prior influence on parameter estimation. Specifically, regression coefficients for each non-reference class were assigned Normal(0, 1) priors, constraining estimates toward reasonable log-odds values without excluding plausible effects (Gelman et al., 2008). Intercepts were modeled with Student-t(3, 0, 2.5) priors, allowing heavier tails to accommodate baseline heterogeneity in class prevalence. These priors improve parameter identifiability and guard against overfitting, particularly with moderate sample sizes and correlated predictors.

Community-Level Model

A second model evaluated community-level predictors, including food resource security, peers' relative socioeconomic status, societal attitudes, and household size. To account for differences in scale, societal attitudes and household size were standardized, and food resource access was modeled as a monotonic ordinal predictor. The model was defined as:

$$\log\left(\frac{P(Y_i = k)}{P(Y_i = 1)}\right) = \beta_{0k} + \beta_{1k}(\text{FoodResource}_i) + \beta_{2k}(\text{PeersRelative}_i) + \beta_{3k}(\text{SocietalAttitudes}_i) + \beta_{4k}(\text{HouseholdSize}_i) + \varepsilon_{ik}.$$

Priors were specified analogously to the individual-level model: Normal(0, 1) for regression coefficients and Student-t(3, 0, 2.5) for intercepts. All models were run with four MCMC chains, 4,000 iterations per chain (1,000 warm-up), using a fixed seed (1121) to ensure

reproducibility. Convergence was confirmed by inspection of trace plots and $\hat{R} < 1.01$ for all parameters, with effective sample sizes exceeding 1,000, indicating stable and well-mixed posterior distributions (Gelman & Rubin, 1992; Vehtari et al., 2021). The $\hat{R}$statistic (R-hat) assesses whether the Markov chains have converged to a common distribution: values close to 1 suggest that the model has stabilized and that the results are reliable. Effective sample size reflects how much independent information the chains provide; higher values indicate that the posterior estimates are based on a sufficiently large number of effectively independent draws, ensuring precision and stability in the reported estimates.

Posterior inference was based on posterior means, 95% credible intervals (CrIs), and the Probability of Direction (pd), which quantifies the proportion of a parameter's posterior distribution that lies on the same side of zero as its median, indicating how confidently the model estimates the direction of an effect (Makowski et al., 2019). In other words, if the 95% credible interval does not cross zero, the entire posterior distribution is either positive or negative, the effect can be considered directionally reliable. The pd statistic quantifies the directional certainty of each effect, with thresholds of pd ≥ .975 indicating strong evidence for an effect. Conceptually, pd plays a similar role to the frequentist p-value, with pd ≥ .975 roughly corresponding to p ≤ .05, though pd expresses certainty about the effect's direction rather than extremity under a null hypothesis.

A Bayesian framework was selected for its ability to (a) quantify uncertainty directly through posterior distributions; (b) regularize estimates for smaller, less frequent subgroups using weakly informative priors; and (c) provide coherent probabilistic evidence of effect direction rather than dichotomous significance testing. Collectively, these properties make

Bayesian multinomial regression a rigorous and interpretable method for examining sociocultural predictors of navigation strategy class membership.

RQ2.2: Do navigation strategies vary by how integrated, alternating, and explored Multiracial individuals' racial identities are?

To examine whether navigation strategy profiles differed in identity processes, I estimated a series of Bayesian Gaussian regression models with each standardized identity process variable as the outcome and navigation strategy class as the categorical predictor. This approach tests whether participants in low-, medium-, and high-endorsement navigation strategy profiles differed in how integrated, alternating, or explored their racial identities are. Identity process variables included (a) MII capturing perceived harmony or conflict between one's racial identities; (b) the MISS subscales, Alternating and Blended identity styles, indexing the extent to which individuals switch between or merge their cultural frames; and (c) the MEIM, indexing ethnic identity exploration and commitment. All identity process scores were standardized (z-scored) prior to analysis to facilitate comparison of effect magnitudes across constructs.

For each identity process, I estimated a separate model of the form:

$$\text{Identity Process}_{\text{i}} = \beta_1(\text{Class1}_i) + \beta_2(\text{Class2}_i) + \beta_3(\text{Class3}_i) + \varepsilon_i,$$

where $\text{Identity Process}_i$represents the standardized identity process score for participant *i*, and the model omits an intercept so that each class receives its own mean estimate ($\beta_1, \beta_2, \beta_3$) and corresponding posterior distribution. Using a zero-intercept means that the model estimates a separate mean for each class rather than treating one class as the reference group. So instead of expressing results as *differences* from a reference class, each coefficient ($\beta_1, \beta_2, \beta_3$) directly represents the average standardized score for that class. This makes the output easier to interpret descriptively. For example, I can state that "participants in the high-endorsement class reported

higher MII" rather than describing the difference relative to a reference group. This structure yields class-specific posterior means and 95% credible intervals that can be directly compared to assess credible differences in identity integration, identity style, and exploration across classes.

Models were implemented using the *brms* package (Bürkner, 2017) in R with a Gaussian likelihood and identity link. Regression coefficients were assigned Normal(0, 1) priors, which are weakly informative and provide mild regularization without constraining plausible mean differences (Gelman et al., 2008). All models were estimated with four Markov Chain Monte Carlo (MCMC) chains, 4,000 total iterations per chain (1,000 warm-up), and a fixed seed (14) for reproducibility (as described in RQ2.1).

A Bayesian framework was selected because it allows for direct probability statements about class-level differences and yields robust estimates even under moderate sample sizes and unequal class proportions. Unlike frequentist ANOVA-style comparisons, Bayesian regression provides full posterior distributions for class means and differences, facilitating uncertainty-aware inferences about how identity integration, alternation, and exploration vary across navigation strategy profiles.

RQ3: Are navigation strategies associated with well-being?

I tested whether strategies for navigating racial identities are associated with higher or lower levels of well-being among multiracial individuals. To examine the relationship between navigation strategy profiles and well-being, I applied the same Bayesian regression framework described above for class membership and identity processes, stratifying well-being scores by navigation strategy class. This approach allowed us to estimate class-specific means and quantify uncertainty in well-being differences across profiles while accounting for unequal class sizes and moderate sample variability.

The standardized well-being score served as the dependent variable, and class membership as the categorical predictor in an equivalent approach to the identity processes models, which estimated an independent mean for each class. The resulting model is expressed as:

$$\text{WellBeing}_i = \beta_1(\text{Class1}_i) + \beta_2(\text{Class2}_i) + \beta_3(\text{Class3}_i) + \varepsilon_i,$$

where WellBeing_i represents the standardized well-being score for participant *i*, and each β_k denotes the posterior mean well-being estimate for Class *k*. The residual term $\varepsilon_i$is assumed to follow a normal distribution with variance σ^2. All regression coefficients were assigned Normal(0, 1) priors for adequate flexibility for class-level mean differences (Gelman et al., 2008). Each model was estimated using four MCMC chains with 4,000 total iterations per chain (1,000 warm-up) and a fixed seed for reproducibility.

By modeling well-being stratified by navigation strategy class, I assessed whether individuals who engage more frequently in strategies of identity negotiation (e.g., learning about heritage, proving authenticity, frame-switching) report different levels of psychological well-being compared to those who engage less in such strategies. This framework provided a probabilistic and uncertainty-aware estimate of how patterns of racial identity navigation relate to subjective well-being among multiracial individuals.

RQ4: How are navigation strategies applied when describing a liminal life event that engages multiple racial identities?

I apply two approaches to analyzing the responses to the narrative prompts. First, I applied the bottom-up STM approach to determine *content*, whether experiences are associated with navigation strategy endorsement. Navigation strategy endorsement was included as a prevalence covariate into the STM model. This approach allows us to assess whether greater

endorsement of navigation strategies is associated with increased emphasis on specific topics, offering insight into how identity navigation shapes the way individuals articulate their lived experiences. I also included participants' ethnicity/race as a content covariate to examine whether the thematic composition of responses differed across ethnic/racial groups (see Appendix C for word clouds). Participants from different racial combinations may have emphasized distinct aspects of identity in their narratives. The same model selection procedure described in Studies 1 and 2 was used to determine the optimal number of topics and can be found in the appendix.

Second, in a top-down approach, valence is detected using valence analysis, a computational method that determines the emotional tone of text. Building on the study by Lilgendahl et al. (2018), which explored bicultural identity by coding narratives for positive, negative, or mixed valence, I analyzed narratives to assess whether life events involving both racial identities were experienced as positive, negative, or both. To accomplish this, I ran sentiment analysis on the narratives with NLP using GPT-3.5-mini LLM, which assigns sentiment scores ranging from -1 to 1 to each narrative. In this scoring system, scores near 1 indicate strong positivity, scores near -1 indicate strong negativity, and scores around 0 suggest a neutral tone. The use of Roberta-base was particularly advantageous because it is known for its robust performance, high accuracy, and ability to handle multiple languages, making it well-suited for a diverse dataset (Rathje et al., 2024). To examine whether topic prevalence was associated with sentiment, I ran separate linear regression models with sentiment scores as the outcome and each topic's prevalence as the predictor. Overall, these themes reveal how multiracial individuals construct meaning around liminal life events, demonstrating how they reconcile, negotiate, and internalize their racial identities.

Results

Descriptives

The five most frequent close-ended navigation strategies were people who chose to identify as "mixed", learn about their ethnic/cultural heritages in their free time, have openness to ambiguity, frame-switch, and experiment with their appearance (see Table 2). Still, forty-five multiracial individuals self-identified as "monoracial", predominantly those who are part-White. There were 39% of the sample that endorsed picking and choosing which customs felt appropriate to practice, while 36% of the sample demonstrated knowledge about their racial communities to *prove* themselves as a member. For the open-ended navigation strategies, participant responses were coded as one strategy from the codebook (with a few exceptions).[15] The five most frequent strategies were the following: Acceptance of self (Yin & Yang), de-emphasis of race, forge their own cultural bridge, lean into one over the other, absence of navigation strategies (imbalanced). The thirty-four participants who described imbalance and without strategy were predominantly part-White (80%).

Table 4.2. Frequency of endorsed close- and open-ended navigation strategy by count

Navigation Strategies			
Close-ended		Open-ended	
Strategy	Count	Strategy	Count
Identify as Mixed	432 (92.3%)	Acceptance (Yin/Yang)	137 (29.3%)
Learn in own Free Time	279 (59.6%)	De-emphasis of Race	69 (14.7%)
Open to Ambiguity	260 (55.6%)	Cultural Bridge	49 (10.5%)
Frame-Switch	195 (41.7%)	Lean into One	35 (7.5%)
Experiment w/ Appearance	186 (39.7%)	Imbalanced	34 (7.3%)
Pick & Choose Customs	181 (38.7%)	Practice Customs	33 (7.1%)
Marginalized Solidarity	173 (37.0%)	Active Learning	29 (6.2%)
Prove w/ Ethnic Knowledge	166 (35.5%)	Frame Switching	29 (6.2%)
Disidentify with Race	91 (19.4%)	Shared Values	25 (5.3%)
Identify as Monoracial	45 (9.6%)	Solidarity	20 (4.3%)

[15]Only a few participants mentioned more than one strategy of finding balance between their multiple racial identities.

Define their Identities	20 (4.3%)
Open-mindedness	20 (4.3%)
Normalcy of Mixedness	8 (1.7%)
Avoid Unwelcoming Spaces	3 (0.6%)

Note. Percentage of the total sample is in parenthesis.

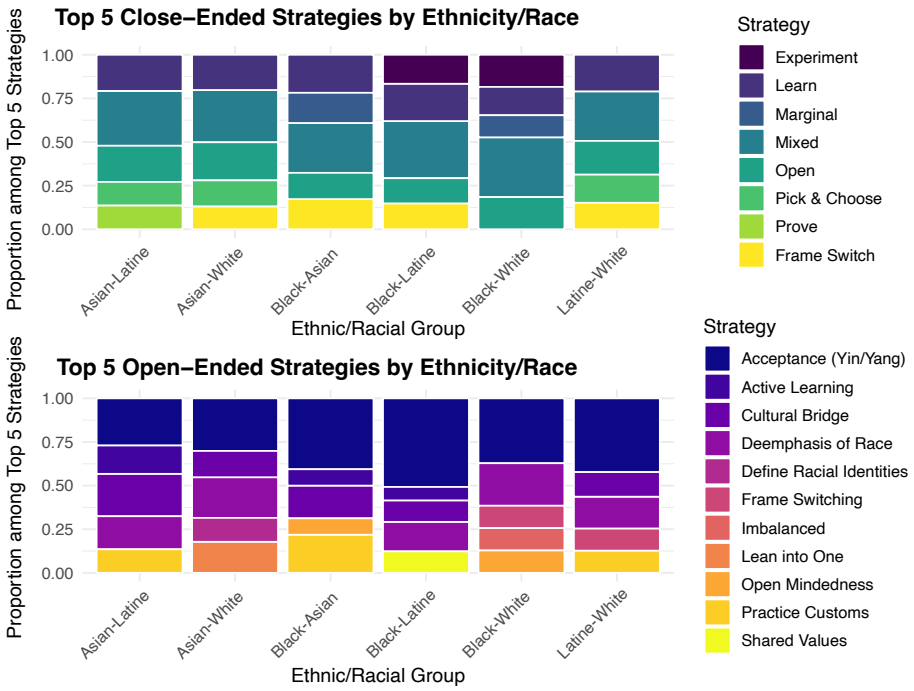
To illustrate group differences in the most common forms of navigation, Figure 2 displays the five most frequently endorsed close-ended navigation strategies within each ethnic/racial group. For each group, endorsement proportions were calculated as the mean binary endorsement across participants and then rescaled within group so that the five most common strategies summed to 1. This normalization highlights the *relative salience* of the most prominent strategies rather than their absolute endorsement rates. As shown in the figure, Black-Asian and Black-White individuals identify with marginalized others; Black-Latine and Black-White describe experimenting with their appearance; Asian-White, Black-Asian, Black-Latine, and Latine-White described frame-switching. Asian-Latine individuals describe proving their membership through knowledge of heritage. Lastly, Asian-Latine, Asian-White, and Latine-White participants pick and choose only certain customs that feel appropriate to practice. A commonly endorsed strategy across *all* backgrounds is identifying as "mixed" and being open to ambiguity.

Across six ethnic/racial combinations, participants varied in the open-ended strategies they described when reflecting on their experiences. As shown in Figure 2, the most frequently endorsed strategies across groups were self-acceptance (Yin/Yang), cultural bridge, and deemphasis of race.

Acceptance (Yin/Yang) was the most common strategy overall, reported by 21.7–40.7% of participants, and was particularly prevalent among Black-Latine (40.7%) and Black-Asian (31.7%) participants. Cultural bridging their racial identities appeared most often among Asian-Latine (19.6%) participants, while deemphasis of race was mentioned most frequently by Black-

White (19.4%) and Asian-White (17.0%) respondents. Asian-White individuals described defining what each racial identity means to oneself as a top strategy. Those of dual-minority descent described active learning, which didn't emerge as a top strategy for part-White participants. Black-Latine participant described recognizing the overlap in shared values, traditions, and histories between their cultural heritages. These patterns suggest that while themes of acceptance and integration were broadly shared, some strategy types (e.g., cultural bridge, deemphasis of race) may be more salient in specific ethnic/racial contexts.

Figure 4.2. Top 5 close- and open-ended navigation strategies by ethnicity/race



RQ1: Navigation Strategy Profiles

To identify the optimal number of latent classes that best captured distinct patterns of navigation strategy endorsement, latent class analyses (LCAs) were estimated with two to six classes (K = 2–6) using participants' close-ended navigation strategy items. Each model was estimated with 20 random starts and up to 10,000 iterations to ensure convergence stability.

Model fit was evaluated using the Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC), and sample-size adjusted BIC (SABIC), as well as the average posterior probability of assignment (APPA). Model fit indices were examined across all candidate solutions (see Table 2). Although the 2-class solution had the lowest BIC, the 3-class solution demonstrated comparable fit (ΔBIC = 21), improved AIC and SABIC, and acceptable classification accuracy (APPA = .82). Conceptually, the three-class model also offered greater granularity, capturing low, medium, and high endorsement profiles, compared to the coarser two-class (low vs. high) `solution. Accordingly, the three-class solution was retained as the optimal and most interpretable representation of navigation strategy profiles for subsequent analyses.

Table 4.3. Model fit indices for two- to six-class latent class solutions

Class	Loglik	AIC	BIC	SABIC	APPA	MinClassProp	MaxClassProp
2	-2545.52	5133.03	5220.15	5153.50	0.85	0.39	0.61
3	-2522.29	5108.58	5241.33	5139.77	0.82	0.18	0.56
4	-2500.80	5087.60	5265.98	5129.51	0.73	0.13	0.36
5	-2487.92	5083.84	5307.86	5136.48	0.68	0.13	0.25
6	-2473.16	5076.32	5345.97	5139.67	0.69	0.04	0.31

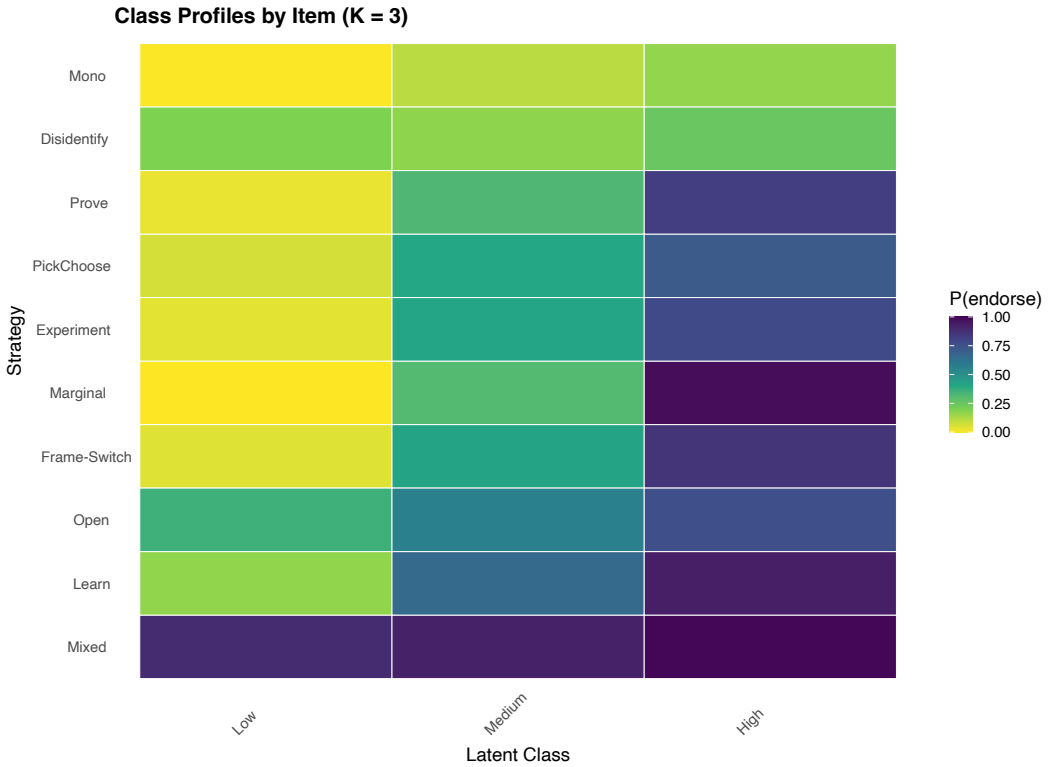
Note. Model fit statistics for latent class analysis (LCA) models estimates two- to six-class solutions based on closed-ended navigation strategy indicators. Lower AIC, BIC, and SABIC values indicate better model fit, with BIC typically emphasized for selecting the optimal number of classes. APPA refers to the Average Posterior Probability of Assignment, where higher values indicate clearer class separation. MinClassProp and MaxClassProp represent the smallest and largest estimated class proportions, respectively, reflecting the relative size of each class.

Figure 3 presents the endorsement of navigation strategies by latent class. In the low-endorsement class, participants generally reported minimal use of active navigation strategies. However, even within this group, there was consistent endorsement of identifying as “mixed”, being open to ambiguity, and occasionally disidentifying with race altogether. This suggests that while these individuals may not frequently engage in deliberate strategies to manage or assert their identity, they embrace certain core orientations toward flexibility and hybridity in self-definition.

Participants in the medium-endorsement class demonstrated higher endorsement across nearly all items, reflecting more active engagement with identity navigation. Members of this class tended to endorse learning about their heritage and seeking connection across group boundaries without strongly emphasizing social validation or outward displays of belonging. This pattern may represent a pragmatic or context-responsive navigation style, characterized by selective engagement depending on situational demands.

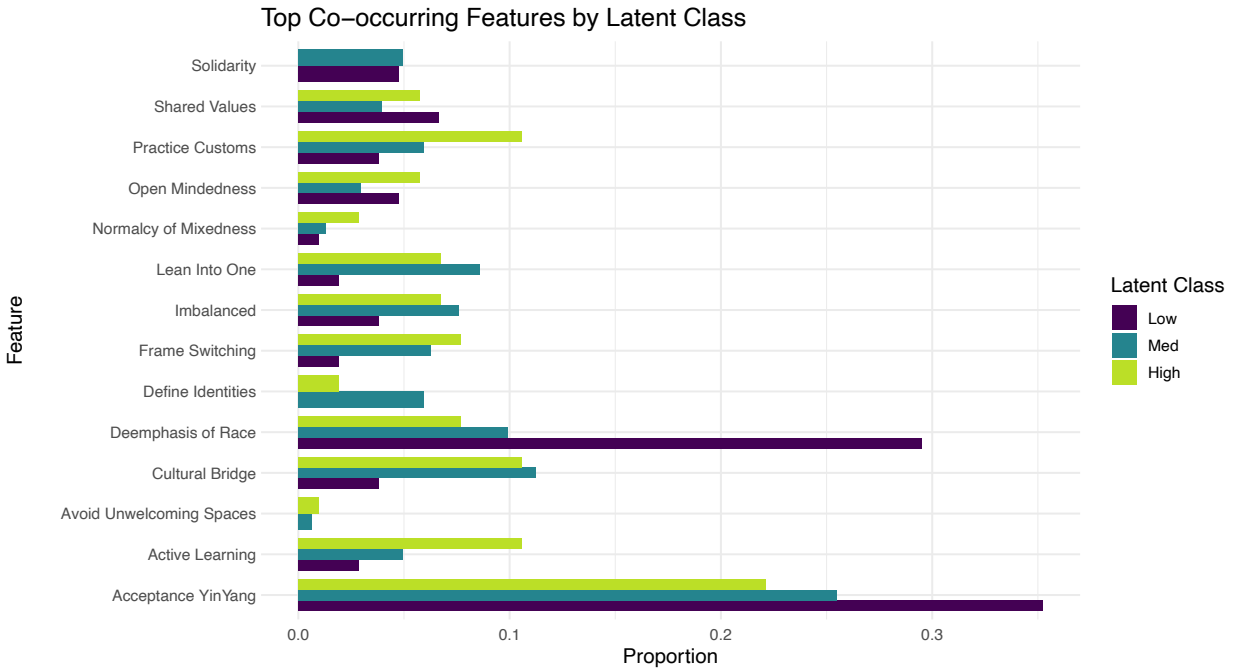
In contrast, the high-endorsement class reported strong agreement with nearly all strategies, particularly those involving relational and community-oriented efforts. Participants in this group were most likely to endorse identifying with other marginalized groups, learning about their racial heritages in their free time, and proving their membership to their racial communities through knowledge and action. Their profiles reflect an intentional and proactive approach to navigating multiracial identity, one that integrates both personal exploration and outward expressions of belonging and legitimacy.

Figure 4.3. Endorsement of close-ended navigation strategies by latent class



In Figure 4, I depict how open-ended navigation strategies vary by classes. Those who endorsed less navigation strategies described in their own words balancing their racial identities through self-acceptance (a sense of yin and yang). They also described a de-emphasis of race in their lives and recognized the shared values between their racial groups. This contrasts the active process of learning, practicing customs and forging a cultural bridge, which was described more by people who had medium and high endorsement for navigation strategies. Other strategies described by people who endorsed more strategies include leaning into one racial identity over the other, and frame-switching. The only strategy that emerged more for medium endorsement than low and higher is uniquely defining what each of their racial identities mean to them.

Figure 4.4. Co-occurrence of latent class for close-ended strategies with methods described by participants (in proportions).



RQ2.1: Individual and Community Predictors of Navigation Strategy Class Membership

A Bayesian multinomial regression was conducted to examine how individual and community-level characteristics predicted participants' likelihood of belonging to each navigation strategy profile. The model estimated posterior distributions for all predictors, with weakly informative priors and four MCMC chains that achieved convergence ($\hat{R} < 1.01$ for all parameters). The low-endorsement class served as the reference category, such that coefficients reflect the log odds of membership in the moderate- (Class 2) or high-endorsement (Class 3) profiles relative to the low-endorsement group. For interpretability, log-odds estimates were exponentiated to odds ratios (ORs). An OR greater than 1 indicates higher odds of belonging to a given class (relative to Class 1 as the reference), whereas an OR less than 1 indicates lower odds. Credible effects are now interpreted based on whether the 95% credible interval excludes 1, which represents the null value of no difference. For example, an OR of 2.50 would indicate that

participants with that characteristic are approximately two and a half times more likely to belong to that class than the reference group, while an OR of 0.50 would indicate half the odds of membership. Table 4 presents the posterior means, which values can be interpreted as the estimated odds ratio for each predictor, averaged across the posterior samples. Table 4 also reports the 95% CrIs, which convey the range within which the true value of the odds ratio falls with 95% posterior probability.

Individual-Level Predictors

At the individual level, age emerged as a robust predictor of navigation strategy class membership. Each one-unit increase in standardized age was associated with 29% lower odds of belonging to the moderate-endorsement class compared to the low-endorsement class (OR = 0.71, 95% CrI [0.51, 1.00]) and 46% lower odds of belonging to the high-endorsement class compared to the low-endorsement class (OR = 0.54, 95% CrI [0.35, 0.83]). The latter interval excluded 1, indicating a credibly negative association between age and high-endorsement membership. In contrast, the association between gender and class membership was weaker and uncertain. Compared to men, women were about 35% less likely to belong to the moderate-endorsement class (OR = 0.65, 95% CrI [0.34, 1.27]) and about 51% less likely to belong to the high-endorsement class (OR = 0.49, 95% CrI [0.21, 1.07]). Because both credible intervals included 1, these effects were not credibly different from zero, suggesting that gender was not a reliable predictor of navigation strategy profile membership. No specific racial/ethnic combination significantly predicted class membership; all 95% CrIs included 1, suggesting that differences among multiracial subgroups (e.g., Asian–White, Black–White, Latine–White) were not strongly associated with navigation profiles. Similarly, phenotype alignment and having a

multiracial parent did not yield clear evidence of differentiation across classes, as their posterior distributions overlapped substantially with zero.

Community-Level Predictors

Community variables, including peers' relative socioeconomic standing, societal attitudes, and household size, were not credible predictors of class membership. All odds ratio posterior means were near one with wide CrIs, suggesting limited influence of these contextual factors on individuals' navigation strategy patterns.

Overall, the Bayesian multinomial model indicated age as the most reliable predictor, with older people more likely to report lower endorsement of identity navigation strategies. Other individual and community factors showed considerable uncertainty. These results suggest that while multiracial individuals share broadly similar structural conditions, their navigation strategies may vary more strongly by experiences from age than by specific racial combination or contextual background.

Table 4.4. Predictors of navigation strategy class membership based on individual and community characteristics

	Predictor	**Posterior Mean**	**95% CrI (Low)**	**95% CrI (High)**
Individual				
Class 2	Female	0.65	0.34	1.27
	Asian-White	0.68	0.25	1.85
	Black-Asian	0.83	0.28	2.55
	Black-Latine	0.60	0.25	1.55
	Black-White	0.55	0.22	1.40
	Latine-White	1.36	0.53	3.57
	Phenotype Context-Dependent	1.16	0.40	3.36
	Different Phenotype	1.03	0.33	3.19
	Mixed Phenotype	0.58	0.20	1.74

	One Racial Phenotype	0.95	0.36	2.50
	Multiracial Parent	0.83	0.42	1.69
	Age	0.71	0.51	1.00
	Education	1.21	0.67	1.89
	Generational Status	0.70	0.35	1.29
Class 3				
	Female	0.49	0.21	1.07
	Asian-White	0.78	0.24	2.44
	Black-Asian	1.63	0.49	5.53
	Black-Latine	0.80	0.27	2.36
	Black-White	0.98	0.33	2.89
	Latine-White	2.03	0.68	6.04
	Phenotype Context-Dependent	1.19	0.38	3.85
	Different Phenotype	0.78	0.21	2.76
	Mixed Phenotype	0.53	0.15	1.83
	One Racial Phenotype	0.85	0.29	2.41
	Multiracial Parent	0.82	0.35	1.92
	Age	0.54	0.35	0.83
	Education	1.29	0.71	2.85
	Generational Status	0.81	0.40	1.88
Community				
Class 2	Peers Relative	1.10	0.70	1.75
	Societal Attitudes	1.00	0.80	1.26
	Household Size	0.98	0.78	1.22
Class 3				
	Peers Relative	0.72	0.44	1.15
	Societal Attitudes	0.91	0.72	1.15
	Household size	1.03	0.82	1.29

Note. Posterior means (odds ratios) and 95% credible intervals (CrIs) from Bayesian multinomial logistic regression models predicting latent class membership. Class 1 (Low Endorsement) served as the reference category for both comparisons (Class 2 vs. Class 1 and Class 3 vs. Class 1). Odds ratios greater than 1 indicate higher odds of belonging to the respective class relative to the reference class, whereas odds ratios less than 1 indicate lower odds. Individual-level predictors include gender, ethnic/racial combination, age, phenotype alignment, and generational loci; community-level predictors include peers' relative socioeconomic standing, perceived societal attitudes, and household size.

RQ2.2: Navigation Strategy Class Membership Predicts Identity Processes

To examine whether navigation strategy profiles differed in their identity processes, Bayesian regression models were estimated for each identity construct (Modified Bicultural Identity Integration (MII), Multicultural Identity Styles Scale (MISS–Alternating and MISS–Blended subscales), and the Multigroup Ethnic Identity Measure (MEIM)) with latent class membership entered as the grouping variable. In Figure 5, the posterior distributions are presented across latent classes. Table 5 summarizes the posterior means, standard errors, and 95% credible intervals (CrIs).

MII. Posterior estimates revealed meaningful differences in MII across navigation strategy classes. Participants in the low-endorsement class reported the highest levels of perceived identity integration (M = 0.25, 95% CrI [0.06, 0.45]), whereas those in the moderate-endorsement class reported MII scores near zero (M = −0.02, 95% CrI [−0.13, 0.10]), indicating average integration. The high-endorsement class exhibited the lowest MII (M = −0.24, 95% CrI [−0.45, −0.04]), with the CrI excluding zero, suggesting a reliably lower sense of cohesion between their racial and cultural identities. This pattern suggests that individuals who actively employ navigation strategies may experience lower identity integration between their cultural backgrounds.

MISS–Alter and MISS–Blend. For the Alternating Identity Style subscale, class differences emerged. The high-endorsement class showed higher scores (M = 0.27, 95% CrI [0.06, 0.47]), indicating a stronger tendency to shift between cultural frames depending on context. The low-endorsement class, by contrast, scored reliably lower (M = −0.32, 95% CrI [−0.51, −0.13]), suggesting less flexibility in alternating between identities. The moderate-endorsement class remained near the average (M = 0.03, 95% CrI [−0.08, 0.15]). Together, these results indicate that greater strategic engagement is associated with more dynamic and situational

identity negotiation. For the Blended Identity Style, differences across classes were small and uncertain; all 95% CIs included zero. Posterior means suggested a slight positive trend for the high-endorsement class (M = 0.10, 95% CrI [−0.11, 0.31]), but the overall evidence indicated no reliable differences in the extent to which participants viewed their cultural identities as blended or synthesized across classes.

MEIM. Stronger differences were observed for ethnic identity exploration and commitment. The high-endorsement class reported substantially higher MEIM scores (M = 0.37, 95% CrI [0.17, 0.57]), while the low-endorsement class reported reliably lower scores (M = −0.41, 95% CrI [−0.60, −0.22]). The moderate-endorsement class again fell near the overall mean (M = 0.03, 95% CrI [−0.08, 0.15]). This suggests that individuals who engage most extensively in navigation strategies were also most actively involved in exploring and affirming their ethnic-racial identity.

Figure 4.5. Posterior distribution across identity process constructs

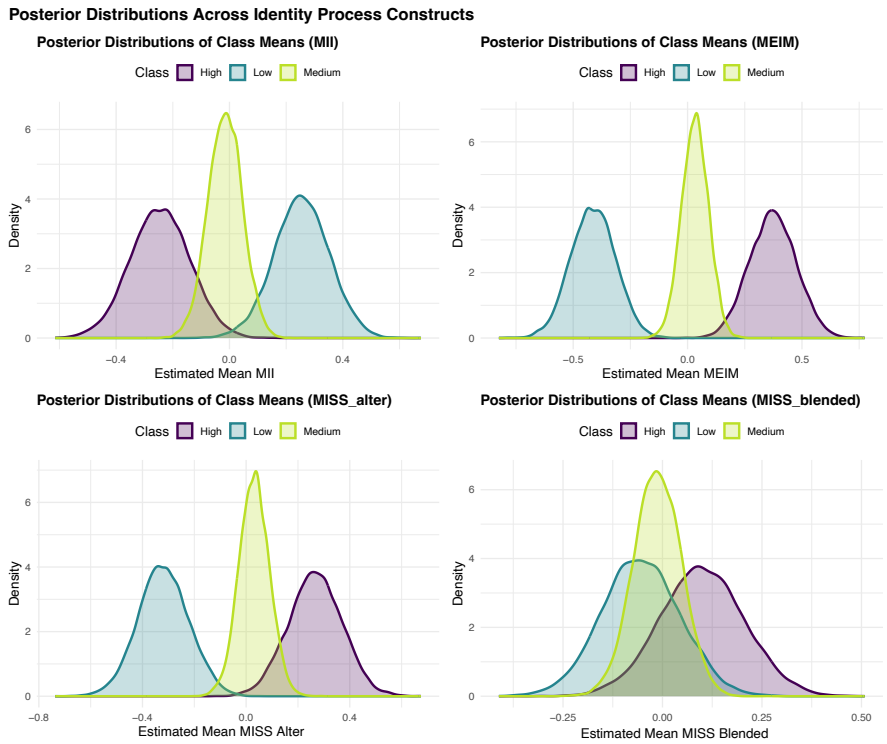


Table 4.5. Posterior estimates of identity process means stratified by latent class

Identity Process	Parameter	Estimate	Std. Error	95% CrI (Low)	95% CrI (High)
MII					
	Class 1	0.25	0.10	0.06	0.45
	Class 2	-0.02	0.06	-0.13	0.10
	Class 3	-0.24	0.11	-0.45	-0.04
	Sigma	0.99	0.03	0.93	1.06
	lprior	-4.16	0.04	-4.24	-4.10
MISS (Alter)					
	Class 1	-0.32	0.10	-0.51	-0.13
	Class 2	0.03	0.06	-0.08	0.15
	Class 3	0.27	0.10	0.06	0.47
	Sigma	0.99	0.03	0.92	1.05
	lprior	-4.18	0.04	-4.28	-4.11
MISS (Blend)					
	Class 1	-0.05	0.10	-0.25	0.14
	Class 2	-0.01	0.06	-0.13	0.10
	Class 3	0.10	0.11	-0.11	0.31
	Sigma	1.00	0.03	0.94	1.07
	lprior	-4.10	0.02	-4.15	-4.08
MEIM					
	Class 1	-0.41	0.10	-0.60	-0.22
	Class 2	0.03	0.06	-0.08	0.15
	Class 3	0.37	0.10	0.17	0.57
	Sigma	0.97	0.03	0.91	1.04
	lprior	-4.25	0.06	-4.37	-4.15

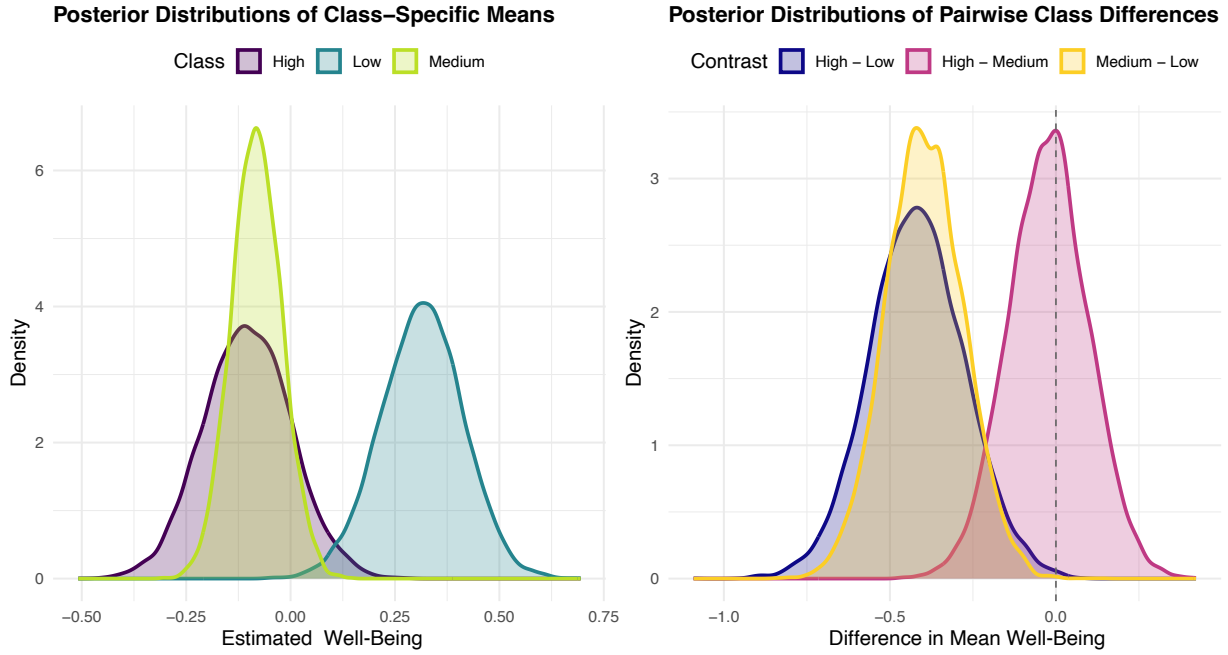
Note. Posterior means, standard errors, and 95% credible intervals (CrrIs) are reported for each latent class across four identity processes: Modified Bicultural Identity Integration (MII), Multicultural Identity Styles Scale—Alternating Identity Styles and Blended Identity subscales (MISS–Alter, MISS–Blend), and the Multigroup Ethnic Identity Measure (MEIM). Models were estimated separately using Bayesian regression with latent class as a grouping variable. σ represents the residual standard deviation (within-class variance), *lprior* denotes the log prior probability based on specified priors.

RQ3: Navigation Strategy Class Membership as a Predictor for Well-Being

To test whether navigation strategy profiles were associated with differences in well-being, a Bayesian linear regression was estimated with well-being as the outcome and latent class membership as the grouping variable. Posterior distributions were estimated with weakly informative priors, and all MCMC chains achieved convergence ($\hat{R} < 1.01$). Figure 6a displays the posterior distributions for class-specific means, Figure 6b displays the pairwise class differences, and Table 6 summarizes the posterior estimates and 95% credible intervals (CrIs).

Posterior means indicated clear differences in well-being across navigation strategy classes. Participants in the low-endorsement class exhibited the highest estimated well-being (M = 0.31, 95% CrI [0.12, 0.50]), with the credible interval excluding zero, indicating a reliably positive deviation from the sample mean. In contrast, both the moderate-endorsement (M = −0.08, 95% CrI [−0.20, 0.04]) and high-endorsement (M = −0.10, 95% CrI [−0.31, 0.10]) classes showed lower and statistically uncertain well-being estimates, with CrIs overlapping zero. Posterior contrasts between classes (Figure 6, right panel) showed that low-endorsement individuals were credibly higher in well-being than both moderate and high-endorsement individuals, as most of the posterior mass for these differences lay above zero. However, the difference between the moderate- and high-endorsement classes was negligible, with substantial overlap of their posterior distributions.

Figure 4.6a & 4.6b. Posterior distributions of class-specific means and pairwise class differences of strategy endorsement



Note. In Figure 6a, posterior distributions reflect estimated mean well-being scores for each navigation strategy endorsement class (low, moderate, and high). Shaded regions represent 95%

highest posterior density (HPD) intervals. Higher values denote greater well-being. Figure 6b displays the estimated differences in mean well-being between endorsement classes (moderate – low, high – low, and high – moderate). Differences between moderate and high classes were uncertain and near zero.

Table 4.6. Posterior estimates of well-being stratified by latent class

Parameter	Estimate	Std. Error	95% CrI (Low)	95% CrI (High)
Class 1	0.31	0.10	0.12	0.50
Class 2	-0.08	0.06	-0.20	0.04
Class 3	-0.10	0.10	-0.31	0.10
Sigma	0.99	0.03	0.93	1.06
lprior	-4.15	0.04	-4.23	-4.10

Note. Posterior means, standard errors, and 95% credible intervals (CrIs) are reported for each latent class across well-being. σ represents the residual standard deviation (within-class variance), *lprior* denotes the log prior probability based on specified priors.

Navigation Strategy Endorsement as a Continuous Variable

The latent class approach identified distinct navigation strategy profiles. However, the pattern of results across identity processes and well-being suggests that endorsement of navigation strategies varies along a continuum rather than forming discrete subgroups. Class membership showed a monotonic, graded association with both identity integration and well-being, consistent with a linear trend rather than categorical distinctions. To more accurately capture this dimensional structure, follow-up analyses models navigation strategy endorsement as a continuous variable. This also provided a more statistically rigorous and interpretable test, preserving individual variability, improving precision, and avoiding information loss from classification uncertainty. I re-analyzed across the three domains:

1. Sociocultural factors (e.g., gender, racial heritage, phenotype alignment, and community context),
2. Identity processes (MII, MISS, MEIM), and
3. Well-being

As depicted in Table 7, the continuous-variable analyses largely corroborated the patterns observed in the latent class models, reinforcing a consistent directional relationship between navigation strategy engagement, identity processes, and well-being. Higher endorsement of navigation strategies was associated with lower multiracial identity integration (MII; $\beta = -0.09$, 95% CrI [−0.13, −0.05]) and greater alternating identity use (MISS–Alter; $\beta = 0.09$, 95% CrI [0.05, 0.13]). Endorsement also positively predicted ethnic identity involvement (MEIM; $\beta = 0.11$, 95% CrI [0.07, 0.15]) and negatively predicted well-being ($\beta = -0.07$, 95% CrI [−0.11, −0.03]). Contextual variables, including perceived societal attitudes, peers' relative socioeconomic standing, and household size, were not credible predictors of strategy endorsement, suggesting that these navigation patterns reflect individual-level identity dynamics.

These converging results suggest that the patterns identified across latent classes reflect an underlying linear continuum of strategy engagement, where increasing endorsement corresponds to heightened identity negotiation and salience, and reduced feelings of identity integration and well-being.

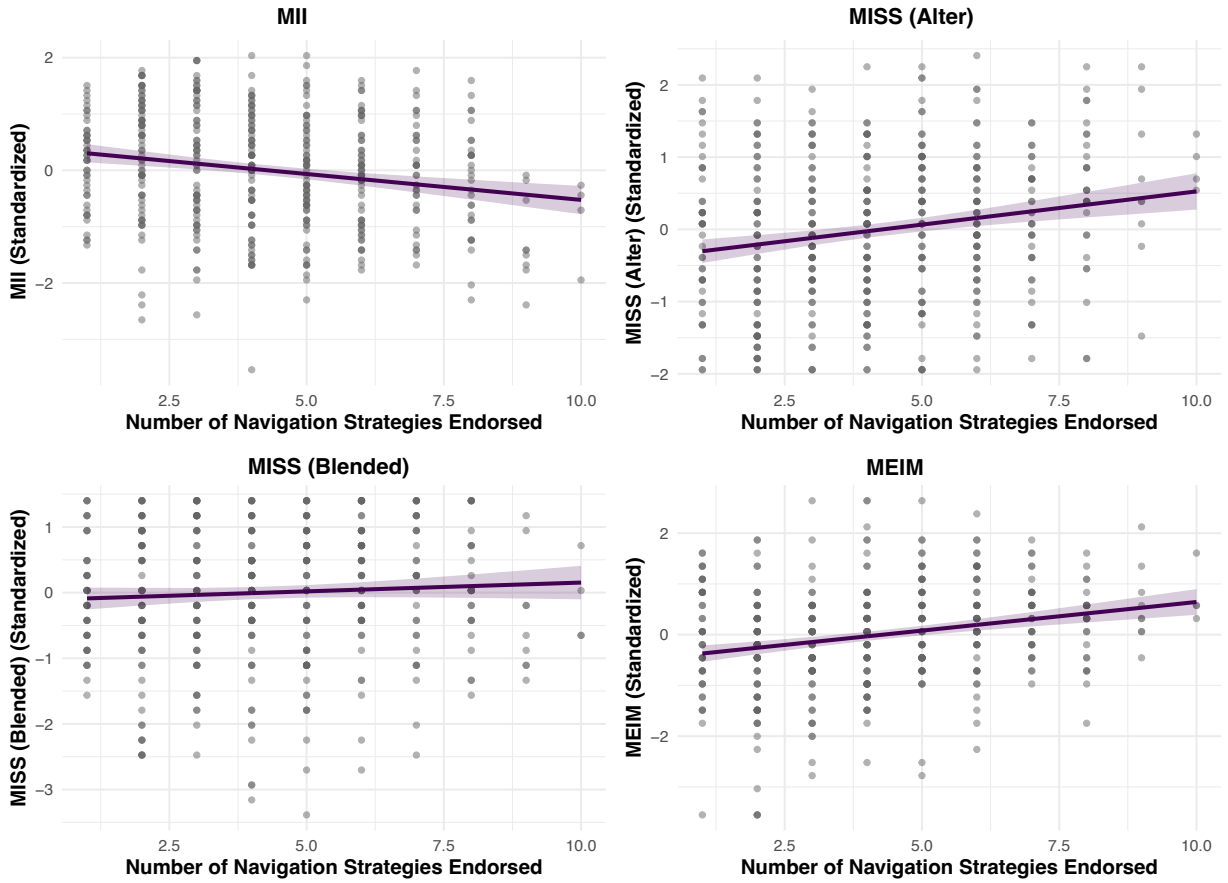
Table 4.7. Posterior estimates from Bayesian linear regressions predicting strategy endorsement and associated outcomes

Construct	Parameter	Estimate	Std. Error	95% CrI (Low)	95% CrI (High)
Individual Char.					
	Female	-0.49	0.29	-1.06	0.08
	Asian-White	-0.13	0.46	-1.03	0.79
	Black-Asian	0.11	0.51	-0.89	1.12
	Black-Latine	-0.16	0.45	-1.03	0.73
	Black-White	-0.07	0.46	-0.97	0.85
	Latine-White	0.56	0.44	-0.31	1.42
	Pheno. Context-Dependent	0.16	0.50	-0.85	1.13
	Different Pheno.	-0.21	0.55	-1.30	0.86
	Mixed Pheno.	-0.71	0.54	-1.78	0.34
	One Racial Pheno.	-0.28	0.47	-1.21	0.62
	Age	-0.33*	0.14	-0.61	-0.04
	Multiracial Parent	0.06	0.31	-0.54	0.66

	Education	0.20	0.15	-0.07	0.52
	Gen. Status	-0.17	0.29	-0.68	0.46
Community Char.					
	Peers Relative	-0.14	0.20	-0.54	0.26
	Societal Attitudes	-0.01	0.10	-0.21	0.19
	Household Size	0.08	0.10	-0.12	0.28
	Food Resource	0.20	0.11	-0.03	0.42
MII					
	NS Endorsement	-0.09*	0.02	-0.13	-0.05
MISS (Alter)					
	NS Endorsement	0.09*	0.02	0.05	0.13
MISS (Blend)					
	NS Endorsement	0.03	0.02	-0.02	0.07
MEIM					
	NS Endorsement	0.11*	0.02	0.07	0.15
Well-Being					
	NS Endorsement	-0.07*	0.02	-0.11	-0.03

Note. Posterior means, standard errors, and 95% credible intervals (CrIs) are reported for each predictor. Individual- and community-level characteristics were entered as predictors of navigation strategy endorsement. Separate Bayesian linear regressions were then estimated with endorsement as a continuous predictor of identity processes (MII, Alternating and Blended Identity Styles, MEIM) and well-being. Negative coefficients indicate lower outcome values with higher endorsement, whereas positive coefficients indicate higher outcome values. σ represents the residual standard deviation (within-model variance). All models used weakly informative priors and achieved convergence ($\hat{R} < 1.01$). Estimates marked with an asterisk indicate Bayesian 95% credible intervals that fall entirely in one direction (all positive or all negative).

Figure 4.7. Posterior distribution for navigation strategy endorsement across identity processes



RQ4: How are navigation strategies applied when describing a liminal life event that engages multiple racial identities?

I selected a six-topic model based on multiple diagnostic metrics (see Appendix C for details on model selection and fit indices). Topic 1 is labeled "Synchronized Ritual", such that the five most representative narratives describe being fully immersed in a family event, ranging from a funeral, Christmas gathering, baptism, and a wedding. Every narrative describes coexistence of cultural practices through compare and contrasts. These are rites of passage in the traditional sense; however, each event is culturally juxtaposed by family members. There are a few who mentioned feeling closer to one of their racial identities over the other because of

socialization through family (e.g., “this made me lose touch with my French culture and become more attached to my Indian culture”; “I felt more comfortable and closer to my Vietnamese family, than I did to my American (German) family”). There are moments where both identities are joined together: “both of my grandmothers gifted me with special items, my Chinese grandma gave me a beautiful ring that belonged to her that had jade in the setting. When I wear my jade it makes me feel my Chinese side strongly. My Irish grandma gave me a handmade Irish lace fabric that was equally as significant and special to me” and “the birth of my niece brought both sides of my family together.” There is language of recognizing socialization of cultural elements in their own family lineage. The most frequent and exclusive words involve family members and gatherings (see Table 8).

Topic 2 is labeled “Claiming Heritage through Practice,” with the most representative narratives describing instances of actively learning about their cultural backgrounds through traditional dances, festivals, and customs. All the narratives took place in a new physical space (e.g., going to university, new city, summer volunteering in mother’s hometown), which forced the individual to reconcile their identities on their own terms. Participants described graduation, entering campus and joining organizations, attending a festival as rites of passage. The frequent and exclusive words (e.g., integrating, roots, navigating, strengthened, excitement, complexities) fit the notion of experiencing for the first time. For instance, one individual described reconciliation: “This experience led me to embrace my multiracial identity and work to mend the cultural divides in my community.”

Topic 3 is “Off-White” to represent the experiences described where Multiracial people are both white passing yet not quite “white.” These participants describe transformative events where they learned the salience of their appearance (e.g., “I have light skin and light eyes, so I

was used to people making fun of me and calling me white, but in a sea of white people, I clearly stood out"; "I'm not black enough for black people and too black for white people…one of them said I passed the paper bag test"; and "I'm still not "one of them"). These are instances where phenotype is foregrounded and alienates the individual from all their racial identities. The top words are "group(s)", "color", and "surprised," which captures group dynamics and colorism.

Topic 4 is labeled "Shame Related to own Culture and History," which revolves strongly around feeling shame. The most representative narrative is a mixed Mexican individual who felt shame towards their own relationship with being Mexican. During this transformative life event, he went dancing in his close friend's hometown and described bringing shame to his friend for acting too White and not speaking Spanish. The experience changed his perception of his own relationship to his culture, a theme that echoes the other four representative narratives for this topic. An individual who is mixed Japanese-American describes shame when first learning about the history of Japan colonization of other nations, a sentiment doubled by her knowledge of American history. She says: "It's the moment I realized I do not have a culture that I can feel pride in." Another participant of Black-descent describes self-reflection and a changed perspective after reading the *Autobiography of Malcolm X*: "Learning about how things were in the past gave me a better understanding of the civil rights movement and what they were fighting for."

Topic 5 is labeled "Ongoing Negotiation," which represents narratives that describe conflict and unresolved conceptualization of the multiracial self. The most representative narrative is a Korean-Guatemalan individual who describes distancing from their racial backgrounds and claiming a neutral American identity after years of having a torn identity. An experience that emerged as highly associated with this topic is internalized racism: "To this day,

I feel a sense of shame for being Black, and I frequently wish I were just fully-White." One top narrative, however, also reflects a more hopeful ongoing negotiation, with the possibility of connecting with one racial background more in the future: "I did not grow up with the ability to have many Hispanic friends and this might also push me towards wanting to connect with the Hispanic community more."

Topic 6 is labeled "Racialized Belonging" because these life events all involve the confrontation with racialized belonging in everyday life transitions, moments when multiracial individuals are forced to *locate themselves* within racial hierarchies, communities, or systems of privilege and exclusion. Three of the top five narratives describe beneficial aspects of being mixed and belonging (e.g., "Being half Japanese and half white put me in a position where I could speak to the children in Japanese, teach them English, and have positive interactions like this one"; "I had friends from every part of the USA and the world like brothers to me now.") While the top narrative describes intense joy in belonging, there is also pain in the other experiences where people are alienated. Each narrator enters a destabilizing social context (the Japanese teacher on a remote island, a tenant displaced by an apartment fire, a child navigating schoolyard racism, or a parent relocating their family) and through these transitions, they become acutely aware of how their mixed identity mediates access to belonging. There is emotionally-charged language in the representative narratives, such as "very hurt", "very frustrated", "joyous", and "sad."

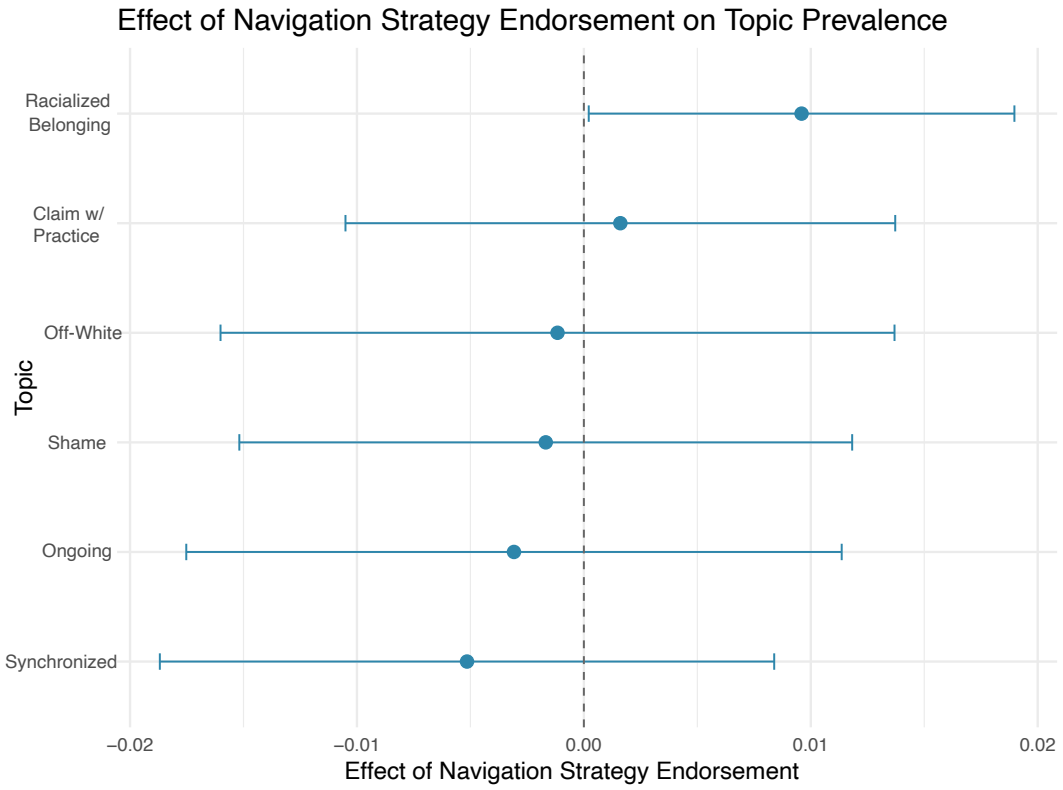
Table 4.8. Labeled STM topics in the 6-topic model with excerpt of most representative narratives, frequent and exclusive (FREX) words and prevalence represented as the mean (and median) gamma probability of documents.

Number	Label	Rite of Passage	FREX Words	Topic Prev.
1	Synchronized Rituals	"My black family celebrated the birth of my niece, by dedicating her to a spiritual Christian service in an all-black church, while she wore the traditional Puerto Rican gown from my grandmother (abuela)."	cousins, filipino, birthday, excited, party, family	.19 (.02)
2	Claiming Heritage through Practice	"I was actively involved in the Mexican Student Association, where I celebrated Dia de los Muertos and participated in traditional dances.. I engaged with the Chinese Student Association, learning about Lunar New Year festivities."	integrating, aspects, roots, navigating, strengthened, excitement	.11 (.01)
3	Off-White	"Both black, white and Hispanic kids didn't know how to view me… I was told that I was lighter than a paper bag so I was an acceptable color."	groups, considered, asians, group, color, surprised	.24 (.04)

4	Shame Related to own Culture and History	“I am proud of myself for continuing to learn, but I am not proud of the history of the ethnicities that make up part of my DNA.”	lack, disgusted, acting, heavily, slave, reading	.18 (.02)
5	Ongoing Negotiation	“It left me feeling relieved that the divide between Afro-Latinos and other Latinos was growing smaller, especially in younger generations. But I was also conflicted about who was supposed to fill the Afro-Latino, particularly Haitian, space in this specific dance setting.”	office, schools, struggle, younger, factors, confront	.21 (.03)
6	Racialized Belonging	“Learning about Japanese culture through a child's eyes is important to me ... I visited Japan when I was little to see my grandparents and I felt very nostalgic playing and interacting with my students.”	recess, three, attack, stress, valued, placed	.07 (.01)

Once the STM topics were labeled, I looked at whether the kinds of life transitions varied by endorsement of navigation strategies. Figure 8 depicts the linear regression estimate between the topic prevalence and strategy endorsement. The only significant association is between racialized belonging and high endorsement of navigation strategies. Participants who described experiences centered on belonging or lack thereof across racial contexts reported higher endorsement of navigation strategies. The prevalence of the remaining topics did not significantly differ by levels of strategy endorsement. (See Appendix C for how mentioning certain rites of passage varied by multiracial identity processes).

Figure 4.8. The types of rites of passage varies by navigation strategy endorsement.

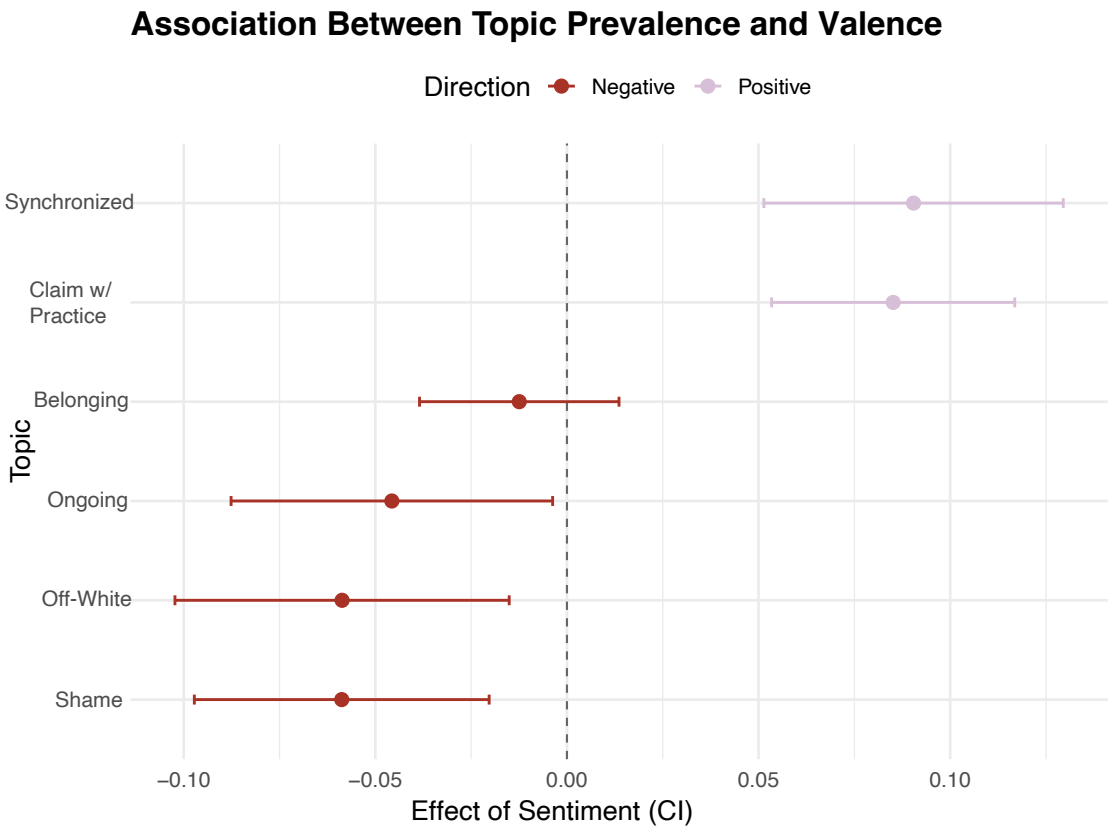


For examining the association between topic prevalence and narrative sentiment, I used GPT-3.5 to assign a valence score for each narrative. In Figure 9, regression coefficients (β) represent the direction and magnitude of association between each topic's prevalence and

narrative valence. Positive coefficients indicate that the topic was expressed in more positively valenced narratives, whereas negative coefficients indicate more negatively valenced or emotionally charged narratives. Regression analyses examining associations between topic prevalence and sentiment (OpenAI valence scores) revealed distinct emotional signatures across topics (see Figure 9). Narratives dominated by Topic 1 and Topic 2 were significantly and positively associated with valence ($\beta = 0.09$, $p < .001$; $\beta = 0.09$, $p < .001$), suggesting that these themes captured joy, connection, and cultural pride. In contrast, Topics 3–5 were negatively valenced (Topic 3: $\beta = -0.06$, $p = .009$; Topic 4: $\beta = -0.06$, $p = .003$; Topic 5: $\beta = -0.05$, $p = .033$), narratives that more often emphasized struggle, conflict, or loss**.** Topic 6 was not significantly related to sentiment ($\beta = -0.01$, $p = .35$), suggesting a more heterogeneous composition.

Narratives that describe synchronized rituals and claiming their heritage through practice were significantly associated with higher sentiment, reflecting themes of connection, pride, and cultural unity, while narratives around shame and marginalization were negatively associated with sentiment. Liminal events with mention of belonging showed no reliable association; this is likely because there are both narratives that describe the joy of belonging and the feeling of exclusion. Sentiment analysis revealed meaningful variation in the emotional tone of narratives describing liminal life events that engaged multiple racial identities. Overall, valence scores ranged from strongly negative to strongly positive, with an average tone slightly above the neutral midpoint ($M = 0.11$, $SD = 0.71$), indicating that participants generally described these events with mild positivity.

Figure 4.9. Association between rite of passage and valence of the narrative



These results align with the qualitative interpretation of topics: narratives that described synchronized cultural rituals, shared practice, and the joy of connection (Topics 1–2) were expressed with greater positivity, whereas those depicting racial conflict and marginalization (Topics 3–5) reflected more negative emotional tone.

Discussion

This study provides one of the first examinations of how multiracial individuals navigate their racial identities across psychological, sociocultural, and narrative domains. Across methods, navigation strategies form a continuum of engagement rather than discrete types, reflecting variation in how actively individuals manage and make meaning of their racial identities. Those who engaged more intensely in navigation reported greater identity exploration but lower integration and well-being, suggesting that active identity work can be both adaptive and taxing. Qualitative analyses revealed that those who endorsed many strategies often described deliberate learning, cultural bridging, and authenticity-seeking, while low-endorsers emphasized acceptance and reduced racial salience, patterns that mark developmental differences in how coherence is achieved. Finally, narrative analyses of liminal life events illustrated how navigation becomes most visible in transitional moments that test belonging and legitimacy, highlighting the symbolic dimensions of identity work. These findings portray navigation as a process that enables agency and growth even as it exposes the psychological costs of managing belonging in a racialized world. In the sections that follow, I discuss the navigation strategies, their sociocultural and psychological correlates, and how they manifest narratively within lived experience.

The Nature of Navigation Strategies

Latent class and continuous analyses converged to show that navigation strategies exist along a graded spectrum of engagement, from minimal to highly agentic, capturing the *intensity* of identity work rather than discrete types of individuals. This dimensional pattern underscores that multiracial identity management reflects how salient and effortful racial meaning-making becomes across contexts. In this sense, navigation strategies mark the degree to which identity

work is actively engaged rather than a categorical distinction between “resolved” and “unresolved” multiracial selves.

Notably, individuals who reported low endorsement of navigation strategies often described approaches rooted in self-acceptance, openness to ambiguity, and a de-emphasis of race. Contrary to the assumption that less engagement reflects avoidance or disengagement, these patterns suggest a form of *passive integration*, a stance in which racial identity is not a site of tension but an accepted aspect of the self. This finding highlights that identity coherence can emerge through acceptance and release.

This pattern aligns with dynamic models of self and culture (Heine, 2012; Markus & Kitayama, 2010) in which identity is maintained through adaptive recalibration rather than closure. In practical terms, the dimensional pattern indicates that “high engagement” does not necessarily imply dysfunction; it may instead reflect an active, contextually sensitive form of identity maintenance. This insight provides a foundation for interpreting the qualitative and psychological findings that follow: how differing levels of engagement manifest in open-ended narratives, how sociocultural factors contour these patterns, and how navigation relates to well-being and emotional experience.

Insights from Open-Ended Strategies

The open-ended responses offer an interpretive layer to the quantitative findings. Participants’ descriptions ranged from active engagement (e.g., deliberate learning, cultural bridging, authenticity-seeking) to more reflective or accepting orientations (e.g., self-acceptance, openness to ambiguity, and de-emphasis of race).

Those who endorsed fewer navigation strategies often articulated an integrative worldview grounded in balance and acceptance, what several described through metaphors of

"yin and yang." Rather than signaling avoidance, this stance reflected an ease with ambiguity and an internalized coherence that required less active negotiation. In contrast, those who endorsed a broader array of strategies described more deliberate and effortful processes of navigating racial boundaries: learning about heritage, asserting authenticity, and adapting behavior across social contexts. These accounts highlight that high engagement in navigation reflects active negotiation of identity boundaries, whereas low engagement may mark a settled integration in which identity tension has diminished.

Medium-endorsers occupied an interpretive middle ground, describing identity in terms of definition and discovery. Their narratives often centered on clarifying what each racial identity meant to them, suggesting a developmental stage in which reflection is accompanied by moderate engagement. Across this continuum, open-ended responses make visible the different "temporalities" of identity navigation: moments of struggle, reconciliation, and eventual acceptance. In the *General Discussion,* I delve more into how this relates to identity development frameworks.

Sociocultural Dimensions

Broader social and structural factors, such as racial appearance, family socialization, socioeconomic resources, and community composition, are often assumed to shape how multiracial individuals learn, express, and reconcile their identities (Rockquemore et al., 2009; Shih et al., 2007; Townsend et al., 2012). Yet, in the present study, these sociocultural variables were not strong or consistent predictors of navigation strategy use. Strikingly, the patterns of low, medium, and high endorsement were remarkably uniform across markers such as education, generational locus, generational status, and even reflected race. This suggests that, despite differing social circumstances, multiracial individuals engage in identity work along a shared

continuum of intensity. The absence of strong structural differentiation points to a key theoretical insight: navigation may be less a product of where individuals are demographically and more a reflection of how they make meaning of their identities across contexts.

One possible explanation is that structural variables capture exposure, whereas navigation captures response. Many contextual features can elicit a response in navigation and does not necessarily determine whether or how individuals take up the work of negotiation. From a cultural psychological perspective, identity navigation from subjective encounters with tension, moments that demand integration and coherence (Heine, 2012; Markus & Kitayama, 2010). The uniformity across structural groups therefore suggests that these psychological dynamics are near-universal features of multiracial experience: regardless of background, individuals must engage in some form of reflection, adaptation, or acceptance to sustain belonging across racialized contexts. These findings also don't speak to the individual strategy applied and how that is context-dependent.

At the same time, age emerged as the only meaningful differentiator for strategy endorsement. Older multiracial individuals endorsed fewer navigation strategies than younger participants, perhaps suggesting a shift toward a more passive stance in relation to racial identity. Rather than continuing to actively manage or negotiate their identities, older individuals may approach racial identity with greater acceptance or detachment, engaging less in deliberate strategies to navigate belonging. This pattern is consistent with qualitative findings showing that multiracial adults often describe "settling into" an identity position or disengaging from the work of negotiation as they age (Jackson, 2012). Lower endorsement among older participants may also reflect not disengagement but internalization and routinization of identity navigation, where strategies become implicit, embodied, or less consciously mobilized.

This pattern reinforces the idea that navigation is a shared psychological process embedded in variable social worlds. These findings help explain why identity processes and psychological correlates, such as integration, exploration, and well-being, were more predictive of navigation intensity than structural variables. This interpretation bridges toward the next section, which examines psychological profiles of navigation.

Psychological Correlates: Integration, Exploration, and Well-Being

The psychological profile of navigation reveals its dual nature as both adaptive and taxing. Higher engagement in navigation strategies was linked to lower identity integration and well-being but greater exploration and alternating identity styles. Navigation thus appears to intensify when coherence is under strain, signaling not resolution but active work toward it.

This pattern aligns with dynamic models of identity development, in which integration is fluid and renegotiated across time and context (Schwartz et al., 2011). For multiracial individuals, navigation may operate as a form of cognitive accommodation: an effortful adjustment that reconciles tensions between internal identification and external perception. Such engagement can deepen understanding and flexibility, yet its effortful nature may temporarily heighten psychological strain, particularly in contexts where belonging is uncertain (Sanchez et al., 2009; Gaither, 2015).

The association between navigation and exploration reinforces this developmental interpretation. Individuals who navigate more intensely are not necessarily fragmented but are more actively searching for meaning and coherence. Alternating identity styles, often seen as signs of instability, may instead indicate adaptive flexibility, allowing individuals to shift frames as contexts demand (Hong et al., 2000 ; Luna et al., 2008). However, this flexibility carries costs: the vigilance and self-awareness it requires can blur boundaries and exhaust emotional resources

(Shih & Sanchez, 2005). The tension between adaptability and fatigue encapsulates the paradox of multiracial identity work, growth through strain, coherence through contradiction.

Yet the pattern among low-engagement participants complicates this developmental framing. Contrary to the assumption that less engagement reflects avoidance or disengagement, these individuals reported high identity integration and well-being alongside lower exploration, a profile that reflects what Phinney (1992) described as the *achieved* status. In Phinney's model, achievement follows active exploration. For many multiracial participants, coherence seemed to arise through acceptance and release, and low identity exploration. This "passive integration" reflects a stance in which racial identity is no longer a site of tension but an accepted aspect of the self. The finding suggests that, for multiracial individuals, identity resolution may sometimes emerge through sustained tolerance for ambiguity and a stable sense of wholeness that renders further exploration unnecessary (see *General Discussion* for further details). Navigation strategies become most tangible in lived experience, during liminal life events that describe belonging, legitimacy, and self-definition.

Navigation as Meaning-Making in Liminal Life Events

The narratives of liminal life events reveal how multiracial individuals live out identity navigation as a process of transformation. These moments (graduations, family rituals, relocations, and first encounters with racialization) function as contemporary *rites of passage* that bring the question of "Who am I becoming?" into focus. In classical terms, van Gennep (1960) described rites of passage as transitions through three phases: separation, liminality, and incorporation. For multiracial individuals, these phases exist not only through prescribed cultural ceremonies but through the lived negotiation of belonging: transitions that mark shifts in awareness as someone who is truly mixed, and a new perspective is earned.

Across narratives, different types of transitions emerged, each reflecting a distinct movement in identity development. In one group of stories, corresponding to topics such as "Synchronized Rituals" or "Claiming Heritage through Practice," participants described traditional milestones hybridized and redefined by the individual (weddings, funerals, holidays) as symbolic affirmations of cultural belonging. These events served as incorporative rites, reaffirming connection and legitimacy within family or community. They mark a transition into recognition, a sense of being seen as whole within a shared cultural space. In this case, individuals drew upon navigation through practicing customs, cultural bridging, recognizing shared values and traditions, and at times, leaning into one racial identity over the other.

A second cluster of narratives took place during graduation, learning spaces, school organizations emphasized the developmental dimension of transition (e.g., "Claiming Heritage through Practice" and "Ongoing Negotiation*")*. School and learning environments often became sites where participants encountered diversity, confronted stereotypes, or redefined themselves in public view. Here, the transition was cognitive and perspectival, a move from unexamined identity toward reflexive awareness. These stories resemble rites of competence, thresholds through which individuals acquire the knowledge or voice to articulate their multiracial standpoint.

The third type of narrative captured experiences of realization or rupture, as in "Off-White," "Shame," and "Ongoing Negotiation." These accounts centered on awakening to racialized realities, moments of exclusion, misrecognition, or structural awareness. For many, these were turning points that crystallized a sense of marginality and also agency: an understanding of themselves as both insiders and outsiders to racial systems. This *multiracial rite of passage* marks a transition to consciousness, from passive belonging to active navigation. It is

in these stories that navigation most visibly becomes moral and political, a stance toward one's place in society.

Only one kind of liminal event was significantly associated with navigation strategy endorsement. People who applied more navigation strategies also more often described a kind of experience revolving around belonging (*"*Racialized Belonging*").* This pattern suggests that active navigators are particularly attuned to the social dynamics of inclusion, interpreting transitions through the lens of belonging itself. The absence of systematic differences across other narrative themes, whether affirming or painful, reinforces that the *types* of transitions prompting reflection are widely shared. Multiracial individuals, regardless of how actively they manage their identities, encounter rites of passage that invite them to locate themselves within families, communities, and cultural lineages. The lack of differentiation by emotional valence for "Racialized Belonging" indicates that navigation is reactive to adversity or celebration and reflects an ongoing orientation toward meaning-making.

Across these variations, the underlying transformation is consistent: liminal life events push individuals toward new modes of coherence. In societies without shared rituals for hybridity, these everyday experiences become the stages on which multiracial individuals author their own initiations.

Study 3 Takeaways

Finally, the present findings reveal multiracial identity navigation in lived experience. Narratives of liminal life events showed that identity work becomes most visible when individuals move through social and developmental transitions that test belonging and legitimacy. The ethnicity- and race-related experiences examined in Study 1 illustrated the external contexts that give rise to identity work, while Study 2 showed how individuals manage

and organize those identities into coherent self-concepts. Study 3 extends this framework by illustrating how these processes are narrated and reinterpreted through lived events. This integration highlights that navigation operates across levels of experience, situational, psychological, and narrative, sustaining coherence within a racialized social world. The implications of this framework for understanding identity development are discussed in the following chapter.

CHAPTER V

GENERAL DISCUSSION

One goal of my dissertation was to demonstrate that marginalization is only one of the many narratives experienced by multiracial people, who, across diverse ethnic/racial backgrounds embody more multifaceted forms of identity than the singular notion of a "multiracial experience" implies. They are indeed hybridized and uniquely defined on their own terms. In the subsequent sections, I address the (1) possibilities of computational methods for measuring identity, (2) the multiracial identity content that emerged (experiential, self-representational, and negotiated), (3) features of multiple identities within and across social categories (intersectionality, dialecticism), and (4) broadly, how patterns in multiracial self-concept and navigation strategies inform existing multiracial identity development frameworks.

Applying Natural Language Processing to Measure Identity

Across the three studies, this dissertation applied NLP to examine how multiracial individuals construct, articulate, and navigate their identities. Each study used computational tools to extend, rather than replace, theory-driven and human-coded approaches. The strong overlap between previously identified themes and the STM-derived topics suggests topic modeling serves as an efficient method for quantifying narratives that builds from established literature. Study 1 used Structural Topic Modeling (STM) to uncover the content of race- and ethnicity-related narratives. The narrative identity content corroborated prior work examining narratives about ethnicity which identified broad content themes: (1) awareness of difference, (2) awareness of underrepresentation, (3) feelings of prejudice, and (4) feelings of connection (Syed & Azmitia, 2008, 2010). The first two themes – awareness of difference and awareness of underrepresentation – appear to be replicated in Study 1. Meanwhile, STM topics were more

specific examples of feelings of prejudice and feelings of connection. Crucially, the distinction between topics allowed us to identify subtle differences between participants.

For Study 2, I applied STM to analyze the self-concept content. The STM revealed six domains that map onto established self-concept categories, encompassing traits, emotions, ideological commitments, physical attributes, and social roles (e.g., Markus & Wurf, 1987; McAdams, 1996; Oyserman et al., 2002; Turner et al., 1987). These categories reflect foundational psychological dimensions through which individuals construct self-understanding: how they see themselves, how they feel, what they believe, what they are, and the social roles they inhabit. These patterns suggest a shared architecture of self-concept that cuts across racial categories.

Beyond these domains, a distinct topic—Multiple Identities/Heritage—emerged, primarily among multiracial and monoracial non-White participants. I will delve more into what this topic supports in the subsequent section (see Multiracial Identity Content). What it methodologically shows is that STM can uncover emergent, data-driven dimensions of identity that extend beyond predefined theoretical categories, revealing culturally embedded self-concept themes that traditional coding frameworks may overlook.

Study 3 extended this approach by applying STM to narratives about rites of passage involving participants' multiple racial identities. Unlike Studies 1 and 2, which built from established frameworks of narrative identity and self-concept, this question invited participants to describe moments of transformation in their own terms. The prompt was intentionally open, and many participants responded with remarkable thoughtfulness, sharing experiences that felt truly liminal to them where multiple racial identities were present: cultural transitions marked by rituals (holidays, weddings, funerals) and redefinition of the self, relative to their racial identities.

Here, STM offered a way to find patterns in these transformative events, and to see, at scale, what kinds of experiences participants identified as transitional in the making of a multiracial self. Topics clustered around themes of ethnic/cultural celebration and marginalization, which echoed the identity-awareness content captured in Study 1. Yet Study 3 emphasized identity as a process of revision, as participants reinterpreted their racial identities through transformation.

STM analysis was able to capture how individuals navigated ethnicity/race at a structural level, including racism and white supremacy. In Study 1, participants incorporated ecological topics, such as *Gained perspective of structural inequality* and *Whiteness,* into their narratives. In Study 3, the topic, *Shame Related to own Culture and History,* also involves reflection on the structural dynamics around ethnicity. Past research indicates awareness of structural factors is a touchstone for identity development (McLean et al., 2018; Renn 2008) and recent narrative research found structural factors, such as white supremacy and racial justice, appear in personal experiences described by emerging adults (Moffitt et al., 2022; Rogers et al., 2021). By tracing these shared linguistic signatures across hundreds of voices, STM centered identity as an ongoing negotiation between culture and lived experience, at the individual, interpersonal, and societal level.

Still, STM has its limitations. Although it can measure identity content and its relationship with other constructs across a large corpus of documents (e.g., motivation themes in Study 1 and navigation strategy endorsement in Study 3), with Study 2, I relied on other computational text analyses to measure semantic similarity between participants in self-concept (text embeddings) and dialectical subthemes (dictionary approach). Even then, human-coding was better suited for measuring between-statement contradictions within a self-concept. For

example, GPT3.5 classified “beautiful” and “fat” as contradictory, even after adjusting queries, which made me hesitant to trust automated judgments of contradiction in self-descriptions. For research questions that are top-down (deductive) with a codebook that has strong theoretical grounding and nuance, human-coding remains the most appropriate tool. While NLP offers scalability and replicability, human coding remains valuable for identifying culturally embedded, psychologically complex meaning with open-ended data.

Multiracial Identity Content

The linguistic patterns observed among multiracial participants align with theoretical work suggesting that multiracial identity development involves negotiation across multiple social and cultural contexts (Rockquemore & Brunsma, 2002; Shih & Sanchez, 2005). These narratives reveal how multiracial individuals make sense of their identities across everyday experiences, life transitions, and the ways they cognitively represent the self.

Monoracial-Multiracial Continuum

Patterns by Ethnic/Cultural Background

In Study 1, one topic—*Experience of Cultural Shock*—explicitly centered family and heritage (e.g., “Christmas,” “rituals,” “holiday,” “ethnic”). Cultural exposure may often occur for multiracial individuals through family settings, especially in the context of initial exposure. Broader vocabulary used by multiracial participants (e.g., “frat,” “jokes,” “friends,” “school,” “restaurant,” “street”) indicates that identity meaning-making unfolds through peer, public, and everyday settings where belonging and difference are negotiated. Notably, these situational contexts parallel those of Black/African American participants, who also described identity within communal and institutional spheres (e.g., “schoolmates”, “teacher,” “bus”, “store”, “church”). By contrast, Asian and Latine participants referenced family members and

communication (e.g., “mom,” “aunt,” “talk,” “speak”), reflecting experiences of race and culture on the cusp of bicultural negotiation, where identity is constructed through relational connection and emotional attunement (Yip et al., 2021; Syed & Azmitia, 2008).

As shown in Study 2, the most robust clustering in self-concept emerged along an Asian versus non-Asian axis, indicating that both monoracial and multiracial participants of Asian descent were more semantically similar to each other than to participants of shared White, Black or Latine descent. Such clustering may arise from a more convergent set of identity scripts available to Asian and mixed-Asian participants. In self-concept content, there *were* patterns among part-Asian participants. Asian-Latine and Asian-White participants had the lowest thematic coherence in self-concept and mentioned hobbies and relational roles significantly more than other racial combinations. These findings are consistent with theoretical models of cultural self-construal that emphasize relationality, duty, and social embeddedness in East and Southeast Asian contexts (Heine, 2001; Markus & Kitayama, 1991). A more prescribed set of cultural expectations and identity scripts may provide common reference points for how individuals from Asian backgrounds describe themselves, contributing to the observed semantic alignment. Participants of Asian descent, both monoracial and multiracial, were also among the most likely to include ethnicity/race in their self-descriptions. Prior work suggests that Asian Americans are frequently positioned as perpetual foreigners, regardless of generational status or cultural fluency (Cheryan & Monin, 2005; Huynh et al., 2011). This pattern may reflect the salience of visible markers or the role of cultural heritage in shaping self-understanding among Asian participants. In the next section, I delve into how similarity in self-concept to their monoracial counterparts may also depend on how multiracial individuals relate multiple aspects of their background.

Multiracial Identity Integration and Co-Descent

Individuals with higher multiracial identity integration were significantly more similar in self-concept to members of their monoracial groups. Similarly, findings on biculturalism demonstrated that bicultural Latine-Americans with higher levels of cultural blendedness perceived greater personality overlap with both their heritage and host cultural groups (Miramontez et al., 2008). Consistent with bicultural identity integration theory (Benet-Martínez & Haritatos, 2005), individuals with integrated identities can fluidly draw from multiple cultural sources to construct a coherent self.

Importantly, the second-best fitting model in the a priori comparisons was one in which participants were grouped by both their specific multiracial combination and their monoracial parent groups. This suggests that semantic similarity a product of both shared monoracial ancestry and shared mixed-race configuration. Individuals who shared a racial combination (e.g., Asian-Black) and its component groups (Asian, Black) showed more coherent overlap in self-concept, especially when they identified with both sides of their heritage (Miramontez et al., 2008). These findings support a co-descent framework (Iverson et al., 2022), in which multiracial identity is shaped through the synthesis of inherited racial meanings. Identity integration, in this view, is not merely a subjective internal psychological state, but an externally driven structuring process that renders multiracial individuals more similar to both of their parent groups.

Multiple Identities/Heritage

Building on patterns of monoracial-multiracial continuum, in Study 2, a topic uniquely emerged for both minority monoracial and multiracial participants—*Multiple Identities/Heritage*. This topic encompassed self-descriptors referencing bicultural or multiracial backgrounds, dual

ancestry, or intersecting identity affiliations, often framed as integral to one's understanding of self (e.g., "I am multicultural, to an African father and an American mother; I am mixed and multi-ethnic, of hybrid identity and blended heritage; interculturally connected, I see myself as a person who hails from multiple ethnic groups with different cultures"). Rather than a peripheral feature, cultural hybridity functions as a central organizing schema, especially for individuals whose identities span multiple racial, ethnic, or cultural worlds. These findings support prior research showing that cultural blending and dual identification can play a formative role in shaping identity content and coherence for certain individuals (Miramontez et al., 2008; Benet-Martínez & Haritatos, 2005). The presence of plural identity themes may be a common feature of racially minoritized populations more broadly.

Interestingly, more than half of part-Black and monoracial Black and White participants did not mention race at all relative to self-concept, a striking omission given the centrality of ethnicity/race in multiracial discourse. This is validated by Study 3, where the second most common open-ended strategy described by participants is de-emphasis of race. One possible interpretation is that some mixed-race individuals, particularly those who are Black, choose to minimize or decenter racial identity to navigate ambiguity or avoid imposed categorization (Rockquemore & Brunsma, 2002). The strategic omission of race may function as a form of boundary negotiation or identity protection, especially when neither monoracial community fully recognizes them.

Multiracial Differs from Monoracial Self-Concept

The distinction between monoracial and multiracial identity is partly shaped by the specific ethnic and racial backgrounds that influence how multiracial identity takes form. There are, however, universal patterns that emerge in multiracial individuals of diverse ethnic/racial

combinations in self-concept. The liminal life events described by multiracial individuals are hybridized experiences, in that they navigate their multiple sociocultural norms at once. Multiracial members also had higher internal contradictions and lower semantic coherence in their self-concept than monoracial members. Monoracial individuals, particularly Black participants, frequently referenced ideology as central to their self-concept, often articulating shared values, political beliefs, and collective resistance to marginalization (Collins et al., 2021). This observation aligns with longstanding theoretical models of racial identity as a sociopolitical anchor among historically oppressed groups. In contrast, multiracial participants, especially those of dual-minority descent, rarely described themselves through ideological or group-based frameworks. Instead, dual-minority individuals tended to describe themselves through individualized, internally generated descriptors, foregrounding agentic traits like creativity, humor, or resilience. Majority-minority multiracial participants, such as Asian-White or Latine-White individuals, more frequently highlighted hobbies and interests, pointing to a self that is anchored in personal pursuits and lifestyle (see the next section for within-multiracial comparisons). In the absence of a singular, cohesive in-group ideology, multiracial individuals may construct self-concepts around internally consistent and affirming dimensions of self. The diversity of their backgrounds can lead to a more individualized approach to identity, where personal interests, hobbies, and personality traits become the stable, affirming markers of who they are. Such strategies may reflect an adaptive means of securing identity coherence and self-continuity in contexts where ethno-cultural anchoring is less readily available.

Although differences were found across studies between multiracial and monoracial individuals, their overlap suggests that multiracial identity development can also be understood

as an extension of shared monoracial identity processes, where multicultural themes of belonging are lived at the intersections of multiple racial identities.

Multiracial Identity: Dual-Minority and Part-White

Multiracial individuals varied in how their identities position them within systems of privilege and marginalization. Dual-minority individuals described navigating identity formation within shared contexts, whereas part-White (majority–minority) individuals occupy more liminal positions that afford both access to and distance from whiteness. Comparing these groups highlights how sociocultural norms and position shape identity negotiation and development of racialized selves.

Dialectical Complexity in Dual-Minority Participants

Dual-minority participants who mentioned racial identity demonstrated the broadest pattern of co-occurrence between race and other self-concept categories, including personality traits, hobbies, SES, and gender. This suggests that ethnicity/race may serve as a central organizing framework through which other aspects of self are interpreted and expressed. Within self-concept, ethnicity/race is integrally woven into diverse dimensions of personal identity. The tendency for dual-minority participants to exhibit this broader intersectional coherence suggests that identity is shaped by the relational meanings among categories. This interconnected self-concept structure challenges the notion that complexity in multiracial self-concept reflects fragmentation; instead, it may reflect the presence of a coherent, contextually situated identity system that integrates ethnicity/race with other lived dimensions of self.

In addition to an interconnected self-concept structure, dual-minority multiracial participants, particularly those of Black-Latine and Black-Asian descent, exhibited the highest prevalence of internal contradiction. These self-concepts often reflected opposing emotional

states (e.g., “I am happy” and “I am depressed”) or self-evaluations that conveyed complexity. By contrast, participants of Asian-White and Asian-Latine descent showed the lowest rates of contradiction, suggesting that navigating multiple cultural frameworks does not necessarily produce greater internal conflict. Instead, these patterns imply that not all multiracial configurations lead to the same contradictory structure in identity. Members of these racial combinations also mentioned agentic and social traits more than other combinations. These findings align with prior work on dialectical self-concepts, which frames contradiction as a marker of complex identity integration (Spencer-Rodgers et al., 2004; Binning et al., 2009). For dual-minority multiracial individuals, contradiction may not simply reflect confusion or instability but an adaptive response to navigating overlapping systems of marginalization, competing expectations, and conflicting identity scripts. In Study 3, Acceptance (Yin/Yang) as an open-ended strategy was highly prevalent among Black-Latine and Black-Asian participants. The presence of such internal oppositions suggests a capacity to hold tension and integrate multiple experiences of self.

Collective Integration and Solidarity

Extending this understanding, Study 3 examined how participants described navigation strategies and rites of passage. Differences between dual-minority and part-White individuals suggest that navigation is not simply a matter of personal preference but reflects the social affordances and constraints available to individuals occupying different positions in the U.S. racial order (Rockquemore & Brunsma, 2002; Root, 1996).

Dual-minority participants (e.g., Black-Latine or Black-Asian individuals) described externally oriented forms of navigation that emphasized connection and solidarity. The prominence of cultural bridging and shared values suggests that navigation may serve as a

restorative process, a means of reconciling histories of marginalization through identifying continuities between heritages. Black-Latine participants' emphasis on shared cultural values and traditions, for example, can be understood as a form of collective integration that transforms racial multiplicity into solidarity. This pattern aligns with prior research showing that awareness of structural discrimination can foster a sense of linked fate and intergroup empathy among racially marginalized individuals (Giamo et al., 2012; Sanchez & Bonam, 2009). Forging a culture bridge is an open-ended navigation strategy unique to Asian-Latine individuals. In contrast to the shared values, traditions, and customs strategy, *Cultural Bridge* reflects implied distance between their racial identities that the individual bridges themselves. Even within dual-minority individuals, there is a difference in the top strategies applied. For these participants, navigation functions less as boundary management and more to affirm kinship and commonality across difference.

Selective Engagement and Proximity to Whiteness

Part-White participants more often reported de-emphasis of race or selectively engaging with cultural practices, highlighting the way proximity to Whiteness may afford greater interpretive flexibility and, potentially, more distance between racial identities (Morning & Saperstein, 2018 ; Song, 2025). Prior work suggests that proximity to Whiteness can facilitate identity disengagement, a shift toward universalistic or individualized conceptions of self that buffer against the discomfort of racial scrutiny (Shih & Sanchez, 2005; Townsend et al., 2012). The relative prevalence of "selective practice" strategies (e.g., choosing which customs to adopt) among Asian-White and Latine-White participants supports this interpretation, suggesting that some forms of navigation are less about reconciling tension than defining within available cultural repertoires.

In Study 2, differences emerged based on combination of ethnic/racial backgrounds. The network analysis revealed that ethnicity/race was structured in its associations with other identity dimensions in group-specific ways. Part-White multiracial participants most often mentioned ethnicity/race in tandem with regional identity. From an ecological developmental perspective, macro-level sociocultural forces, such as national origin, geographic context, and migration history, can influence how individuals assemble their self-concepts (Rockquemore et al., 2009). For participants whose racial identity may not map neatly onto existing social categories, regional or cultural geography may offer a more coherent anchor for identity.

Conclusion. The content of multiracial identity varied by racial configuration: dual-minority individuals positioned race as an integrative and collective dimension of self, emphasizing solidarity and shared experience, while part-White individuals more often de-emphasized race, navigating identity through selective engagement and interpretive flexibility. Dual-minority individuals also exhibited the most internally complex and contradictory self-concepts, often blending agentic and social descriptors, patterns that suggest the sophistication of integrating multiple marginalized cultural frameworks with individuality. These contradictions, rather than indicating confusion, reflect a capacity to hold and synthesize competing narratives of self, underscoring the psychological adaptability that characterizes multiracial identity. Notably, even when race was not explicitly primed, as in Study 2, it still emerged as a structuring element of self-concept, underscoring its prevalence in identity representation. At the same time, Study 3 demonstrated that de-emphasis of race can coexist with psychological integration, particularly among older multiracial participants. These findings suggest that ethnicity/race remains relevant to self-definition for multiracial individuals, yet the centrality of ethnicity/race may shift

adaptively across the life course as individuals achieve greater coherence in how they understand their multiple identities.

Holding Multiplicity: Intersectional and Dialectical Features of Identity

Intersectionality. This dissertation foregrounds ethnicity and race as the primary lens, yet Studies 2 and 3 reveal that multiracial identity content is intertwined with other social identities, notably gender and age. Variation across these categories highlights how the meaning and structure of identity are contextually shaped rather than uniform across individuals. Gender differences in self-concept suggest that the language of self-definition is filtered through socially patterned norms for expressing self-concept. By systematically including gender-diverse participants in future multiracial identity research would highlight the potential for greater nuance in how gender and other identity categories co-construct the self.

In Study 3, the intersectional lens extended to age, revealing that endorsement of navigation strategies decreased with age. Older participants reported less active identity exploration and lower endorsement of alternating strategies, patterns that coincided with higher identity integration. These findings suggest a developmental trajectory in which identity work becomes less externally enacted and more internally consolidated over time. While younger individuals may actively experiment with situational expressions of race and culture, older participants appear to achieve greater coherence by integrating their multiple identities into a stable, reflective sense of self (see subsequent section Revisiting Identity Development Frameworks*)*. In both cases, intersectional contexts provide critical insight into the processes through which multiracial individuals negotiate complexity across life stages and social positions.

Dialecticism. While intersectionality captures how multiple social categories intersect to shape identity, an equally important feature of multiple identities emerges *within* a single category—race. For multiracial individuals, holding two or more racial identities can require accommodating opposing yet coexisting perspectives about self and group membership. In Study 3, this dialectical orientation surfaced most clearly in the *Yin and Yang* navigation strategy (the most frequently endorsed approach) as well as the *Open to Ambiguity* category, endorsed by over half of participants (55.6%). Tolerance for ambiguity is a common and adaptive orientation among multiracial individuals, reflecting a comfort with contradiction and balance as their resolution. Similarly, in Study 2, dialectical features appeared within self-concept structure itself: participants frequently paired contradictory self-views (e.g., "I am confident" and "I am insecure"). Dialecticism may be a feature of dual-minority multiracial identity, one that allows individuals to hold complexity, reconcile internal diversity, and maintain coherence within a racial self. This ethos of balance and flexibility echoes cross-cultural models of selfhood emphasizing harmony and relational attunement (Markus & Kitayama, 2010) and reflects the psychological adaptability theorized to emerge from navigating multiple cultural worlds (Gaither, 2015; Cheng et al., 2008). This dialectical stance challenges the deficit framing applied to multiracial identity: rather than evidence of confusion, ambiguity tolerance appears to be an indicator of psychological maturity and integration.

Revisiting Identity Development Frameworks

This dissertation builds on multiracial identity (e.g., Gullickson & Morning, 2011; Iverson et al., 2022; Poston, 1990; Renn 2004; Rockquemore & Brunsma, 2002; Root 1990), ERI (e.g., Cross et al., 2002; Phinney 1989; Umaña-Taylor et al., 2014), cultural (e.g., Benet-Martínez & Haritatos 2005; Peng & Nisbett, 1999), and, more broadly, identity development

frameworks (Erikson, 1968; Oyserman & Destin, 2010). The next sections articulate four axes through which this work advances theory, linking multiracial identity to general models of self, culture, and development.

1. Reframing what "integration" means in multiracial identity.

As Cross et al. (2017) stated, personal and social identity development are deeply intertwined. This premise anchors the present findings: the process of becoming oneself cannot be separated from the sociocultural contexts that confer meaning to that self. Classic models such as Poston's (1990) stage theory cast multiracial identity development as a linear progression, from enmeshment/denial to achieved integration, implying a developmental endpoint of resolution. This structure parallels Erikson's (1967) lifespan model, where identity formation involves successive tasks of differentiation and synthesis. Evidence from this dissertation partly supports this linear logic: older participants reported lower endorsement of navigation strategies, lower exploration, and higher integration and well-being. These patterns suggest that, with time, active negotiation may recede as coherence becomes internalized, a developmental trajectory consistent with Eriksonian consolidation.

The specific forms of navigation, however, tell a more complex story. The strategies individuals use, such as cultural bridging, de-emphasizing race, or embracing ambiguity, vary by racial and cultural background, revealing distinct pathways to coherence; this is even though the number of strategies endorsed did not differ by race, SES, or phenotype, suggesting a universal developmental progression across multiracial groups. While the trajectory of identity integration follows a broadly shared pattern of decreasing negotiation and increasing coherence over time, the expression of that integration reflects culturally specific idioms of resolution. Low endorsement of navigation strategies encompassed many such routes, identifying as "mixed,"

emphasizing acceptance, or situating the self beyond racial boundaries, each representing a viable form of coherence within different sociocultural frames.

Integration thus does not follow uniform social pathways; it emerges through diverse combinations of experiences and contexts. Here, Renn's (2004) nonlinear and ecological framework offers a crucial linkage: identity is shaped by shifting contexts and can take multiple adaptive forms. In this view, low navigation endorsement among older participants may reflect the routinization of identity work, coherence achieved through familiarity rather than completion.

These findings reconcile Poston's sequential emphasis with Renn's ecological pluralism. Identity development unfolds both across time and across contexts (as multiple integrated outcomes coexist). Integration, then, is a lifespan process of equilibrium, maintained through ongoing calibration between self and social world.

2. *There are a psychological cost and developmental function of identity work.*

In the present study, active identity work is both adaptive and taxing. This tension reflects what Oyserman and Destin (2010) describe as *identity-based motivation*, the effort individuals expend to maintain congruence between who they are and what their social context expects. Effort, in this sense, signals meaning: the more identity feels consequential, the more cognitive and emotional labor it demands.

This pattern echoes Cross's (2002, 2017) account of cyclical identity development, in which people repeatedly re-engage and internalize aspects of self as contexts shift. Younger participants' higher navigation intensity indicates ongoing recalibration. Lower engagement among older adults suggests a movement toward internalization or "recycling," where coherence is maintained with less conscious effort. The psychological cost (strain) and developmental gain

(integration) observed are what Cross and Vandiver frame as the dual nature of identity maturation in a racialized context. Navigation intensity captures the psychological cost of striving for coherence. This reframes ambiguity from a symptom of instability to a developmental condition of identity in plural worlds.

3. *Situating identity development within culture, not just psychology.*

The form that racial identity coherence takes is culturally patterned. This perspective reframes multiracial identity theory within a cultural psychology of selfhood: identity development unfolds through shared meaning systems. The ways people make sense of tension and harmony reflect the semiotic resources available to them, cultural scripts, such as homegoings, holiday rituals, language use, religious participation, and family traditions, that define what it means for them to feel "whole."

These findings extend Gullickson and Morning's (2011) account of racial ancestry dominance. In the current study, self-alignment intertwined with cultural socialization and the norms through which individuals experience belonging within their racial and ethnic communities. Such experiences suggest that ancestry dominance is both a reflection of social hierarchy and a product of differential cultural immersion, which identities are lived and practiced most fully in daily life.

Renn's (2004) ecological model rightly emphasized that identity develops through *settings*: the places, people, and communities that structure lived experience. The present findings demonstrate that cultural meanings provide environments for identity to unfold and supply the interpretive grammar through which coherence itself is constructed. In this sense, culture functions as medium and outcome of identity development, shaping how individuals understand integration as a symbolic act.

4. *Multiracial identity is a model for understanding selves in pluralistic societies.*

Multiracial identity offers a powerful model for understanding how individuals construct coherence in plural cultural worlds. The features of being multiracial individual (e.g., fluid boundaries and continual negotiation across contexts) mirror the broader human task of sustaining a coherent sense of self amid cultural and ideological multiplicity. Multiracial identity then is a prototype of navigating complexity, contradiction, and context-dependence as enduring conditions of life in diverse societies.

This framework intersects naturally with *multiethnic* and *multicultural* identity processes. While multiracial identity refers to navigating socially defined racial categories, multiethnic identity involves reconciling heritages rooted in distinct cultural or national traditions, and multicultural identity captures participation across cultural systems. Despite these distinctions, all three negotiate meaning across overlapping value systems, social expectations, and symbols of belonging. The overlap between minoritized monoracial and multiracial self-concept content in this study highlights that the experience of holding multiple cultural reference points (racial, ethnic, or social) is increasingly a shared human condition.

The present findings expand on bicultural identity frameworks (Benet-Martínez & Haritatos, 2005; Peng & Lee, 2001), which emphasize that integration involves reconciling conflicting schemas through cognitive flexibility and dialectical reasoning. Multiracial participants displayed this same capacity, though their negotiation extended beyond belief systems to include visible and social markers of membership, race and phenotype. Dual-minority individuals, in particular, illustrated adaptive forms of holism, contradiction, and balance, revealing how living within intersecting systems of marginalization fosters flexible, context-

sensitive integration. Their strategies demonstrate that coherence is achieved by resolving difference through inhabiting it.

Earlier models of multiracial identity development (Poston, 1990; Root, 1996; Rockquemore & Brunsma, 2002) framed growth as a process of *choice*—selecting, integrating, or transcending racial categories to achieve resolution. Poston's stage model portrayed integration as the developmental endpoint of conflict; Root offered multiple adaptive "resolutions" to otherness; and Rockquemore & Brunsma defined typologies (singular, border, protean, transcendent) that linked identity outcomes to social context. These frameworks emphasized *what* multiracial individuals ultimately become but less often *how* coherence is maintained across time and context.

The present findings extend these frameworks on two fronts. First, patterns of coherence, contradiction, and semantic overlap in self-concept reveal that multiracial selves are organized systems of meaning, capable of sustaining contradiction and unity simultaneously. Second, integration must be continually managed through the organization of self-knowledge. Study 3 complements this structural view by demonstrating that identity work, the navigation of meaning across racial and cultural boundaries, is actively adaptive. Navigation intensity functions as a psychological mechanism for maintaining equilibrium, linking developmental processes of exploration and integration to lived experience.

Multiracial identities thus exemplify a broader developmental paradigm. Contradiction can exist with coherence, plural socialization fosters dialectical reasoning, and living between worlds cultivates the flexibility needed to thrive within them, which eventually can become more passively internalized. This evolution likely reflects both demographic change and the growing

normalization of the ability to construct meaning, belonging, and coherence across multiple sociocultural frames.

Conclusion

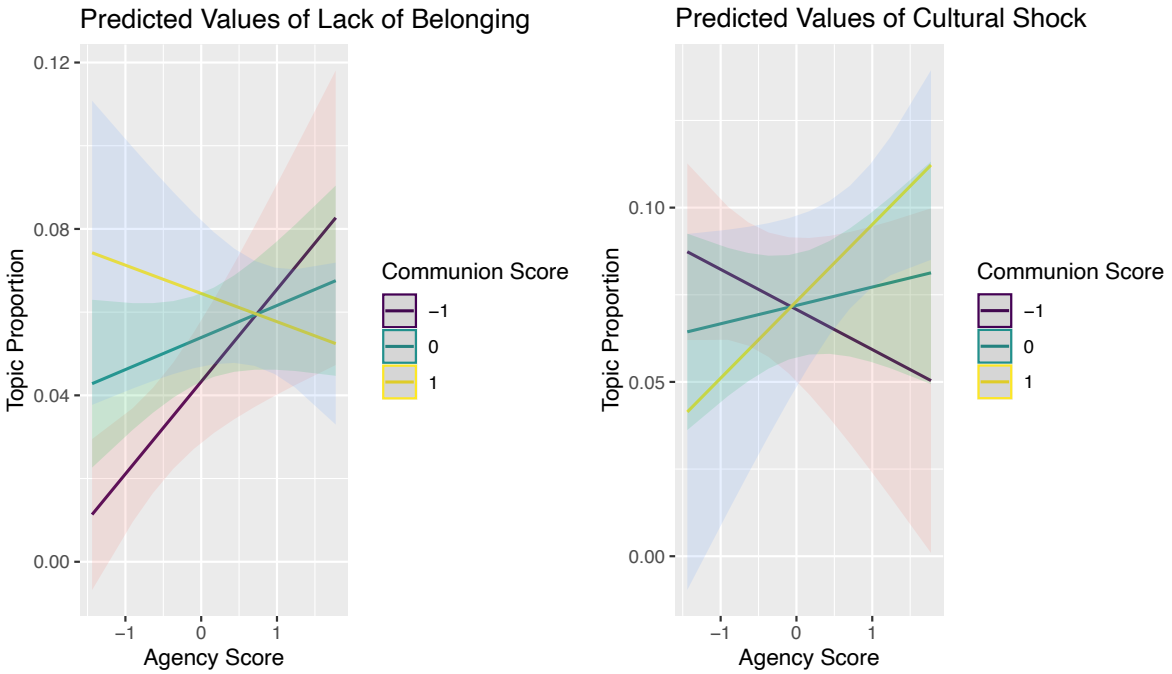
This dissertation shows that multiracial individuals occupy a distinct cultural position: unlike monoracial biculturals who navigate heritage and host cultures tied to geographic space, multiracial people encounter cultural norms through family lineages, yet their everyday experiences around ethnicity and race remain strikingly similar. At the same time, their self-concepts reveal meaningful variation: some do not reference race at all, instead grounding identity in individuating traits, while others embrace contradictions that reflect the multitudes they hold beyond any collective identity. Multiracial individuals, particularly those who are older, sustain race as a quiet but present anchor in the self, holding a yin-and-yang orientation in which racial identities coexist passively rather than contend, highlighting the diverse strategies through which multiracial people navigate who they are. These patterns raise important questions about developmental pathways and whether this de-centering mentality towards race is a feature of maturation or of a specific multiracial cohort. While this study supports the latter, future work can integrate longitudinal and cohort-comparative designs to clarify whether these patterns reflect age, sociocultural context, or both. Ultimately, a cultural-psychological lens is insightful for understanding multiracial identity, presenting both the universal threads and the combination-specific dynamics that shape how multiracial people come to know themselves.

APPENDIX A

Figure A.1. A correlation matrix between the 14 topics within the 14-topic model

In Appendix A Figure 1, there is little to no association for the correlations between topics within the 14-topic model (range of correlations from -.13 to .10). On the x- and y-axis, topics are in ascending order of prevalence. For example, topics *Awareness of cultural difference* and *Awareness of underrepresentation* have a negative association of 0.13.

Figure A.2a & A.2b. Agency and communion scores predicting topic proportions for *Lack of belonging in cultural space* and *Experience of cultural shock*.



The interaction was significant in predicting the prevalence of *Lack of belonging in cultural space* (unstandardized, b = -0.01, p = .010). Supplemental Figure A.2a shows the effect of agency on topic prevalence at different levels of communion. Words associated with lack of belonging in cultural space increased as narratives became more agentic for low communion narratives. In other words, when communion was low, there was a positive relationship between lack of belonging prevalence and agency. There was a significant interaction effect for *Experience of cultural shock* (b = 0.02, p = .040). In Supplemental Figure A.2b, words associated with cultural shocks were the most prevalent in narratives high in communion and agency. On the other hand, prevalence in culture shock had a negative relationship with agency when communion was low.

Table A.1. Labeled STM topics in the 5-topic model with exemplary narratives and prevalence represented as the mean (and median) gamma probability of documents. Please note that sample narratives have not been edited by the authors.

Number	Label	Sample Narrative Excerpt	Topic Prevalence
1	Intergroup interaction	Most of my schools in the past were mainly African American. So, being in a diverse school was different. I felt a mixture of emotions: nervousness, fear, excitement, happiness. I had never interacted with other ethnicities before, never had a friend outside my ethnicity, because my surroundings were mainly African American.	.28 (.27)
2	Marginalization	I was working on research project with a group of my caucasian classmate. Although we got along just fine, but I didn't feel that I was fitting in.	.26 (.25)
3	Awareness of systemic inequality	Of course the education system was lacking, focus was primarily on reading/writing in school whilst trying to keep up with the other counties and their education systems. In Michelle's situation, there was nothing restraining her education, a lot more was expected and could be afforded by the amass of funding her school had.	.17 (.16)
4	Awareness of cultural difference	I haven't seen my Aunt for a long time so we greeted each other by giving a kiss on each cheek. The next day at school my friend ask me why I kissed my aunt that way and I told her thats what we do when we greet someone. At that point I realized that the way my family does things like dance, eat, and greet people was different then everyone else.	.16 (.15)
5	Intragroup interaction/tension	This was a discussion over the incident where a Hmong hunter shot/killed a few White hunter over a racial altercation. This was the biggest thing that's happened in the Hmong community at the time as there was a lot of racial tensions between the Hmong community and the White community.	.13 (.11)

Table A.2. Self-identified ethnic group membership by pan-ethnic group.

Pan-ethnic category	Frequency
Other	
Middle Eastern	9
Unspecified	36
Haya Tribe	1
Hawaiian, Pacific Islander, Polynesian	1
Hawaiian, Portuguese, African American	1
Arab	1
Indian	1
Persian Jewish	1
Chaldean	1
Afghan	1
Chicana	3
Iranian	3
Ethiopian, Eritrean	2
East Indian	1
Moroccan	1
Somali	1
Oromo	1
Hmong	1
East African	1
Guyanese, Indian	1
Kenyan	1
Egyptian	3
African	1
Nigerian	2
Asian	
Asian	49
Chinese	126
Filipino	36
Cantonese	5
Pakistani	3
Taiwanese	10
Korean	18
Pacific Islander	12
Vietnamese	15
Japanese	16

Pan-ethnic category	Frequency
Thai	1
Parsi	1
Sinhalese	3
Tamil	2
Cambodian	7
Hmong	26
Bangladeshi	1
Singapore, Indonesia, Malaysia	1
Indian	23
Black / African American	
African-American	35
Creole	1
Oromo	1
Puerto Rican	1
Ethiopian / Eritrean	2
Ethiopian	2
Eritrean	2
Nigerian	1
Kenyan	2
Somali	3
Sudanese	1
Black / African, Cherokee Indian	1
Latine	
Latina/o/x	14
Hispanic	8
Mexican	46
Dominican	1
Puerto Rican	1
Chicana/o/x	12
Peruvian	1
Cuban, Egyptian	1
South American	1
Salvadorian	5
Central American	5
Latino, Filipino	1
Spanish, Persian	1
Latin-American, Arab-American	1

Pan-ethnic category	Frequency
Colombian	1
Native American	
American Indian	15
Navajo Tribe	2
Ojibwe / Souix	1
Multiracial	
Mixed	19
Filipino, White	7
Black, White, Creole	1
Black, White, Indian American	1
Black, Native American	2
Native American, White	15
Native Hawaiian, White	1
Latinx, White	8
Latinx, Filipino	1
Middle Eastern, White	1
Chinese, Mexican, Salvadorian	1
Black, South Asian	1
White	
White / Caucasian	330
European American	12
German	47
Italian	12
Jewish	17
Irish	39
Hispanic / Latino	13
Portuguese	1
Arab American	3
Egyptian American	1
Scandinavian	5
Czech	1
Czech, Polish	1
Russian	4
Russian, Ukrainian	1
Swiss	1
Croatian	3
Dutch	1

APPENDIX B

Table B.1. Ethnoracial combinations by pan-ethnic group

Pan-ethnic group	Ethnoracial combinations	Frequency
Asian-Latine	Asian, Hispanic	7
	Asian American, Latine	6
	Filipino, Mexican	6
	Asian, Hispanic, White	3
	Filipino, Hispanic	2
	Asian, Venezuelan, White	1
	Cambodian, Dominican	1
	Chinese, Hispanic, West Indian	1
	Indian, Mexican	1
	American, Brazilian, Japanese and Chinese	1
	Asian, Hawaiian, Mexican	1
	Asian, Hispanic, Indian	1
	Asian, Hispanic, Pacific islander	1
	Asian, Mexican	1
	Chilean, Pakistani	1
	Chinese, Filipino, Puerto Rican	1
	Chinese, Latina	1
	Filipino, Hawaiian, Mexican	1
	Hispanic, Korean	1
	Hispanic, South Asian	1
	Hispanic, South Korean	1
	Indigenous, and Pacific Islander, Latinx, Southeast Asian	1
	Japanese, Latino, White	1
	Japanese, Mexican	1
	Latine, Southeast Asian, White	1
	Mexican, Samoan	1
	Mexican, Vietnamese	1
Asian-White	Asian, White	52
	Filipino, White	13
	Indian, White	9
	Japanese, White	6
	Asian American, Chinese	2
	Southeast Asian, White	2
	Asian, Hawaiian, Portuguese	1
	Asian, Hispanic, White	1
	Asian, Native American, Polynesian, Hawaiian, White	1
	Asian, White, Hispanic	1
	Asian, White, Middle eastern	1
	Bengali, White	1
	Vietnamese, White	1
	Chinese, English, Scottish, Irish	1
	Chinese, Irish	1
	East Asian, White	1
	Filipino, German	1

	German, Vietnamese	1
	Hawaiian, White	1
	Indian, Irish/Scottish	1
	Indonesian, White	1
	Italian, Japanese	1
	Japanese, Native American, White	1
	Samoan, White	1
	Sicilian, Vietnamese	1
Black-Asian	Asian American, Black/African American	15
	Filipino, Black/African American	10
	Korean, Black/African American	4
	Japanese, Black/African American	3
	Pacific islander, Black/African American	2
	Southeast Asian, Black/African American	2
	Asian American, Black/African American, Caribbean American	1
	Asian, Black/African American, Mexican	1
	Chinese, Black/African American	1
	East Asian, Black/African American	1
	Filipino, Black/African American, Hawaiian	1
	Thai, Black/African American	1
	Black, South Asian	1
Black-Latine	Black/African American, Hispanic	17
	Black/African American, Latine	16
	Black/African American, Mexican	13
	Black/African American, Puerto Rican	10
	Black/African American, Latine, White	6
	Black/African American, American native/indigenous	4
	Black/African American, Dominican	3
	Black/African American, Nicaraguan	2
	Black/African American, Panamanian	2
	African, Hispanic	2
	Black/African American, Honduran	1
	Black/African American, Latine, Cuban, Mexican	1
	African American and Guarani (indigenous), Latino (Puerto Rican/Paraguayan), White (Italian)	1
	African, Belizean, Hispanic	1
	Black, Hispanic, Dominican, Indigenous (Taino), Latino	1
	Caribbean, Costa Rican	1
	Chile, Cuban	1
	Dominican, Haitian	1
	Haitian, Salvadoran	1
	Jamaican, Salvadoran	1
Black-White	Black/African American, White	84
	Black/African American, White, Native American	5
	Black/African American, Latine, White	2
	Black/African American, White, Asian	2
	African, White	1
	Asian, Black, White	1

	Black American, Hispanic, White	1
	Black, Kiowa, White	1
	Black, Middle eastern, White	1
	Black, Russian	1
	Ethiopian, White	1
	Irish, West Indian	1
Latine-White	Hispanic, White	35
	Latine, White	26
	Mexican, White	13
	Latine/Hispanic, White, Black/African American	6
	Puerto Rican, White	5
	Mexican, White, Native American	4
	Latine/Hispanic, White, Native American	2
	Black/African American, Hispanic	1
	Chilean, Italian, Irish, White	1
	Cuban, German	1
	Hispanic, Jewish	1
	Latine, White, Asian	1
	Mexican, European, African	1
	Mexican, Greek	1
	Cuban-Mexican, White	1
	Guatemalan, White	1
	Hispanic, Irish	1
	Hispanic, Italian	1
	Honduran, White	1
	Irish, Mexican	1
	Irish, Mexican, Spanish	1
	Lithuanian, Puerto Rican	1
	Puerto Rican, White, Afro-Latino	1

Figure B.1. Generational status and loci within the multiracial sample

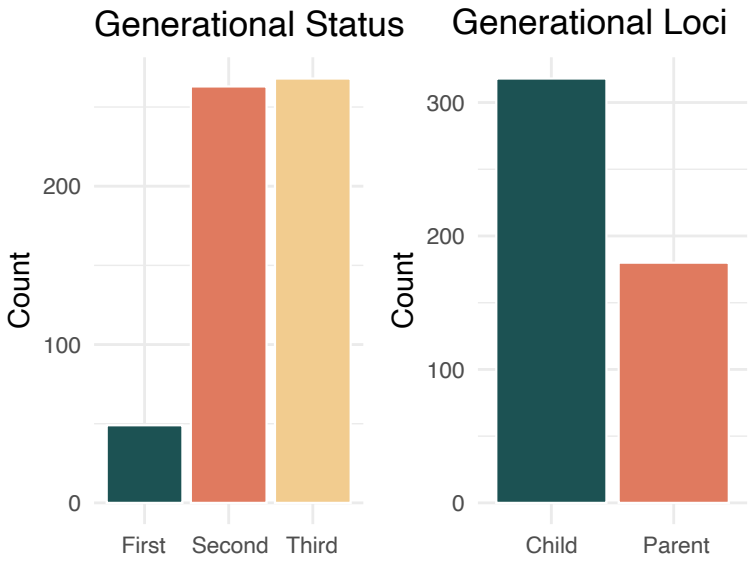


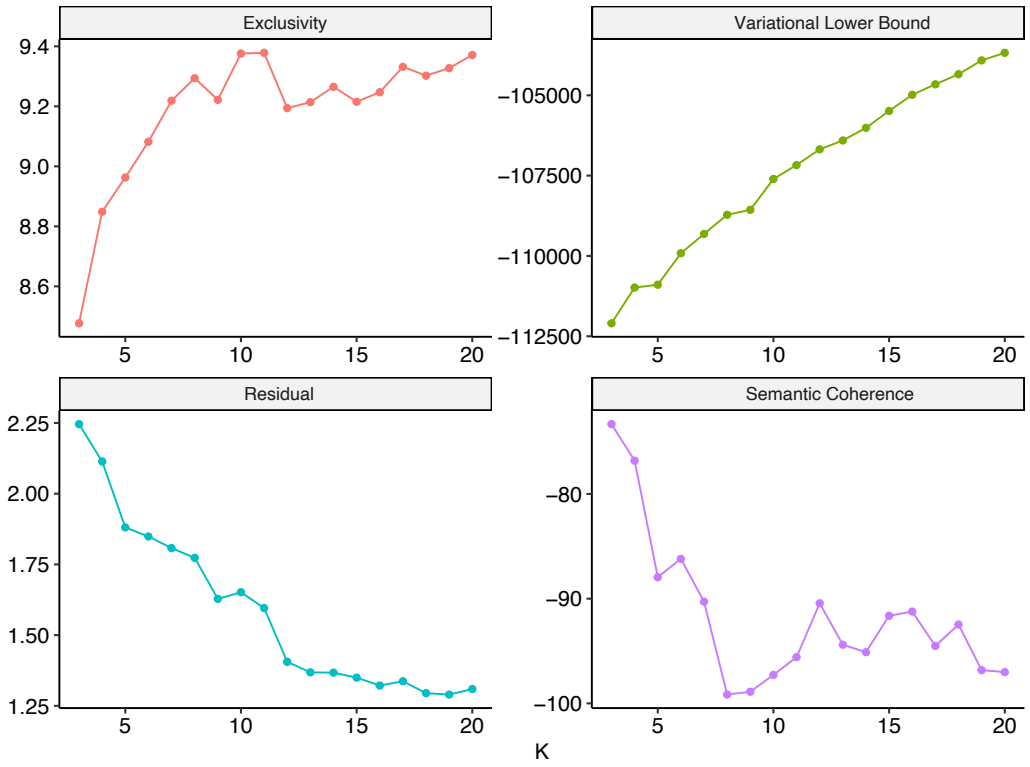
Figure B.2. Query for GPT to classify identity content using Python code

```
I have a list of words in a csv format. Each of these words can be classified in
multiple of the following categories:

[Gender, age, occupation, personality trait/attribute, physical state, emotional state,
family, self-ascribed hobby, oceanic descriptor, ethnicity/race/nationality, regional
identity, social class, goal-striving/aspirational, physical characteristics,
ideology/attitudes, self-evaluation]

Here are a few examples:
woman -> gender
twenty-one years old-> age
hardworking-> personality trait/attribute
active -> personality trait/attribute
korean-> ethnicity/ace/nationality
funny-> personality trait
liberal -> ideology/attitudes
```

Table B.2. Self-concept categories: Significant mean contrasts from post-hoc analysis Holm's correction across ethnicity/race.

Theme	Contrast	Estimate	SE	df	T ratio	p-value	Cohens D
Personality	Asian-White – Black-Asian	-3.32	0.83	594	-3.98	0.003	-0.02
Personality	Asian-White – Black-Latine	-2.69	0.67	594	-4	0.003	-0.02
Personality	Asian-White – Black-White	-2.55	0.64	594	-3.97	0.003	-0.02
Personality	Black-Asian – Latine-White	3.01	0.83	594	3.63	0.013	-0.02
Personality	Black-Latine – Latine-White	2.38	0.67	594	3.57	0.016	-0.02
Personality	Black-White – Latine-White	2.24	0.64	594	3.52	0.019	-0.02
Family	Black-Latine – Latine-White	-0.84	0.25	594	-3.35	0.038	-0.21
Family	Black-Latine – White	-1.23	0.36	594	-3.38	0.035	-0.21
Hobby	Asian-White – Black-Latine	1.2	0.32	594	3.78	0.008	-0.21
Ideology	Asian-Latine – Black	-0.8	0.22	594	-3.57	0.016	0.17
Ideology	Black – Black-Latine	0.81	0.2	594	4.03	0.003	0.17
Ideology	Black – Black-White	0.84	0.2	594	4.25	0.001	0.17
Ideology	Black – Latine-White	0.64	0.2	594	3.28	0.045	0.17
Ideology	Black – White	0.97	0.24	594	4	0.003	0.17

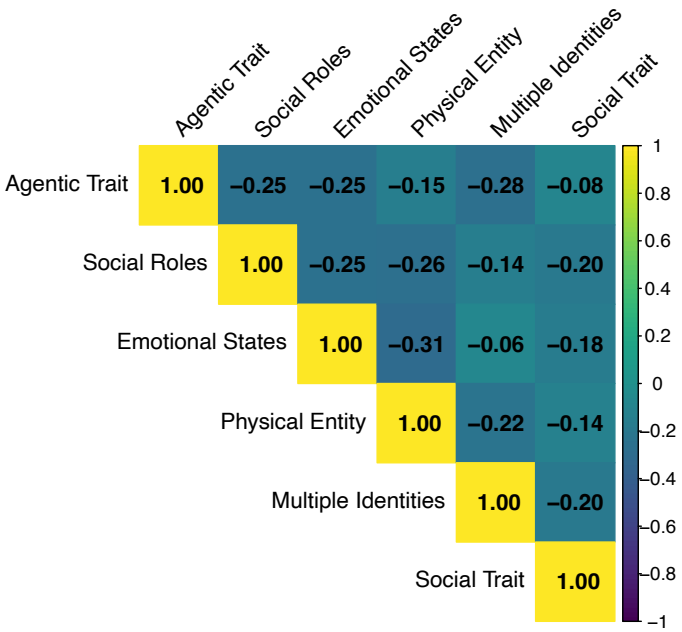
Figure B.3. Determine the optimal number of topics in STM using four metrices: Exclusivity, variational lower bound, residual, and semantic coherence



For STM, I apply four metrics to decide the optimal number of topics. I apply (1) exclusivity, (2) variational lower bound, (3) residual, and (4) semantic coherence. Exclusivity measures how uniquely words are associated with a single topic, with higher values indicating clearer topic distinctions. Variational lower bound reflects model convergence, with lower values (when all else is equal) suggesting better fit. Residuals, in this context, refer to their dispersion, though interpretations vary—some researchers look for local minima, while others consider values below 1 as sufficient for good fit. Finally, semantic coherence measures how frequently high-probability words within a topic co-occur, a metric that correlates well with human assessments of topic coherence, though its validity remains debated (Koltcov et al., 2019). While each of these metrics provides insight into model fit, none alone can fully determine an ideal,

interpretable model. In Figure B.3, altogether there is high exclusivity, moderately low residual, low variational lower bound, and high semantic coherence at six topics for the STM.

Figure B.4. Correlation matrix between topics within the 6-topic STM model



In Figure B.4, on the right-hand side there is the correlation range from -1 to 1. The STM topics were negatively correlated with each other. Since none of the STM topics were positively correlated, there were no overlapping topics to collapse.

Table B.3. Frequent and exclusive words by cluster for Spectral, AHC, GMM, and EGA from the LLM text embeddings

Clustering Approach	Cluster Number	Frequent Words	Exclusive Words
AHC	1	person, good, smart, kind, lover, woman, funny, creative, strong, friend	happy, student, friendly, anxious, quiet, worker, curious, wife, black, tall
	2	woman, lover, reader, adventurous, lgbtq+, life, lover, movies, new, parent.	lgbtq+, 27, ambiverted, apples, born., brown-haired, challenges, continuously, dog-owner, heritage
	3	always, at, loves, is, trying, love, on, it, more, by	day., myself., heritage, takes, believe, l, since, these, become., cultures
EGA	1	daughter, friend, sister, lover, woman, good, person, kind, wife, mother	creator, manager, neighbor, niece, auntie, disney, gardener, homemaker, aged, believer

	2	smart, kind, strong, funny, caring, intelligent, creative, good, determined, honest	genuine, affectionate, tough, dynamic, knowledgable, abundant, cunning, desirable, discipline, eager,
	3	always, good, at, loves, person., is, very, person, likes, on	life., future., cultural, day., i,Äôm, individual., myself., people., worker., it
	4	person, american, lover, good, smart, kind, woman, student, at, college	broke, science, swimmer, 1, heterosexual, accountant, allergic, mothers, side, 30s,
	5	tired, creative, smart, funny, trying, person, curious, good, anxious, at	exhausted, thirsty, ambiguous, lethargic, obsessive, soft, stress, upset, /, abuse
GMM	1	smart, funny, kind, strong, creative, caring, good, intelligent, loving, happy	diligent, genuine, polite, selfish, skilled, assertive, charming, cooperative, tolerant, accountable,
	2	person, lover, good, woman, friend, daughter, american, sister, at, smart	sister, chinese, photographer, conservative, heritage, korean, sibling, tired., color,
	3	always, is, at, loves, trying, very, good, love, on, person	day, myself, people., way, world., playing, political, takes, believe, education
	4	kind, friend, man, daughter, strong, woman, caring, funny, intelligent, smart	accepting, goat, simmer, tech-savvy, african, 19, 5'8", 7, 8, accomodating
Spectral	1	lover, woman, friend, daughter, person, sister, american, good, kind, smart	aunt, neighbor, niece, 1, 36, disney, gardener, homemaker, immigrant, lesbian
	2	smart, kind, funny, strong, creative, caring, good, loving, intelligent, happy	wise, ready, talented, awesome, diligent, genuine, polite, selfish, skilled, assertive
	3	at, always, good, person, very, on, trying, loves, is, likes	don't, values, because, family., so, his, think, try, concerned, doesn't

In the supplemental materials, I triangulate the topic modeling by clustering the text embeddings derived from LLM. I grouped people whose descriptors were overall more similar across different clustering techniques and found converging patterns. While LLM is beneficial in its capacity to derive semantic meaning from words and position them together, a limitation is the potential for biases found across the large corpora (e.g., Caliskan et al., 2017; Charlesworth et al., 2021). STM, on the other hand, is derived from the data itself (not an external source), and

not subject to biases the content LLMs are trained on. However, STM is a probabilistic model and does not account for complex linguistic relationships, like word order or nuanced semantics, as LLM does. Hence, I found it beneficial to include our clusters from LLM to complement the STM topics. First, I considered multiple clustering approaches for comparison to determine whether there is thematic overlap in content, narrowed down to one model, and compared with STM topics. I compared Spectral, Agglomerative Hierarchical, Exploratory Graph Analysis (EGA), and Gaussian Mixture Model (GMM) clustering for their ability to capture high dimensional structures (text embeddings) and interpretability to produce semantically coherent clusters. Unlike k-means, which assumes spherical clusters, these approaches can uncover asymmetrically sized clusters because they rely on more flexible grouping criteria—such as graph connectivity (in Spectral and EGA) or probabilistic density (in GMM)—rather than enforcing uniform cluster shapes. To determine optimal number of clusters, for spectral and agglomerative hierarchical clustering, I applied the elbow method, which looks for the point ("elbow") where adding more clusters stops making a big improvement, and the silhouette method, which assesses how well each data point fits within its assigned cluster versus how strongly it differs from other clusters (Luxburg, 2007 ; Rousseeuw, 1987). With EGA, the optimal number of clusters are determined by identifying distinct communities within the network structure (Golino & Epskamp, 2017), whereas for GMM, the optimal number of clusters is determined by model selection criteria such as the Bayesian Information Criterion (BIC) (Fraley & Raftery, 1998 ; McLachlan & Peel, 2000).

As shown in Table B.3, there was overlap in cluster content, such that across approaches self-concepts clustered around (1) social roles, (2) personality traits and states, (3) self-exploration and growth, and (4) multiple identity and heritage. In Table B.3, I include the high

frequency and exclusive words within each cluster. Conceptually, the clusters were interpretable from GMM and spectral, while AHC was difficult to interpret. For example, both clusters 1 and 2 contained overlapping attributes (e.g.,"lover"). Spectral and GMM clustering highly overlapped conceptually. GMM cluster 1 and spectral cluster 2 overlapped; GMM cluster 2 and spectral cluster 1 overlapped; finally, GMM cluster 3 and spectral cluster 3 overlapped.

Table B.4. Internal validity measures across cluster approaches.

Method	Silhouette	Dunn	kNN Connectivity	MST Connectivity
AHC	0.18	0.22	292	60
Spectral	0.02	0.05	1096	221
EGA	-0.01	0.06	1353	251
GMM	-0.03	0.05	1126	222

For narrowing down to a clustering approach, I applied internal validity measures: silhouette score (now that the number of clusters are determined), Dunn index, kNN connectivity, and MST connectivity (Brock et al., 2008). Silhouette score evaluates how well individual data points fit into their assigned clusters by comparing within-cluster cohesion to between-cluster separation, with higher values indicating well-defined clusters (Rousseeuw, 1987). Dunn index is calculated as the ratio of the smallest inter-cluster distance (minimum distance between points in different clusters) to the largest intra-cluster distance (the maximum distance between points within the same cluster). A higher Dunn Index indicates well-separated and compact clusters, suggesting a better clustering solution; kNN connectivity measures how often neighboring points belong to the same cluster, with lower values indicating stronger clustering structure; and MST connectivity evaluates global structure by analyzing cluster cohesion through a minimum spanning tree, where lower values suggest more cohesive clusters. The four validity measures assess different aspects of clustering quality. These evaluation metrics suggest that AHC outperforms the other methods in terms of cluster compactness and

separation, as indicated by the highest Silhouette (0.0989) and Dunn index (0.2642). GMM and spectral are comparable for Dunn index and lower connectivity scores (kNN & MST). EGA clustering exhibit weaker performance, with negative Silhouette scores and lower Dunn indices. Overall, while AHC provides the most distinct and compact clusters, GMM and spectral contain interpretable clusters with the similar connectivity and separation.

I chose Gaussian Mixture Model (GMM) as the clustering approach due to its internal validity, semantic interpretability, and probabilistic nature. GMM captures complex, elliptical distributions, making it well suited for nuanced data structures. Its probabilistic assignments also allow for soft clustering as well, reflecting the fluidity of identities rather than forcing rigid categories. Additionally, GMM uncovers latent structures by modeling data as a mixture of distributions, similar as a probabilistic framework like Structural Topic Modeling (STM).

Figure B.5. Word clouds for the four GMM clusters derived from the text embeddings.

Table B.5. Highest-probability self-concepts for each GMM cluster.

Cluster	Self-Concept 1	Self-Concept 2	Self-Concept 3
Traits & States	Strong, kind, loving, energetic, giving, hesitant, bold, blessed, limitless, kind, caring, calm, eager, relentless, open, innovative, vague, talkative, friendly, shy	Thoughtful, Forgiving, Empathetic, Open minded, selfless, kind, respectful, sweet, beautiful, self aware, generous, strong, capable, giving, friendly, genuine, Open,	Kind, Funny, Strong, Determined, Stubborn, Creative, Anxious, Loud, Active, Resilient, Shy, Introverted, Adventurous, Willing to learn, Unsure, Brave,

		emotionally intelligent, intelligent, wise	Responsible, Informed, Caring, Smart
Social Roles	A man, biracial, an LGBTQ+ ally, a progressive Democrat, an American, a dog-owner, a former teacher, a lover of video games, a reader, a writer, a brother	A woman, a dog mom, a podcast lover, 27, a Taurus, a lover of movies, a daughter, sister and aunt, someone who laughs at everything, a filipino and black woman, a reader, a swimmer	A nursing student, a caregiver, an aunt, a sister, a daughter, a friend, an animal lover, a christian, a woman, black, american, puerto rican, brown
Self-Improvement & Growth	Striving to be better than what I settle for., a good listener., fully capable of learning anything and unlearning what isn't serving me., goal oriented., struggling with creating a vision for my life., working on building better relationships with those around me., concerned more about the collective., of high esteem to those who know me., aware of my short comings., someone who blends in nearly anywhere, but fits in nowhere., curious., gentle by choice., thoughtful., often saddened by situations and surroundings I find myself in., a problem solver., trying even when there's not much left in my tank., scared of what's next., excited to see who I become., at odds with the parts of me that are not good., dreaming for a brighter future for everyone I know.	Just some weirdo on the Internet with too much free time., kind of a part of the furry fandom but kinda not., biracial cause my mother was from the Philippines while my dad was American., someone that does enjoy to cook occasionally and try new foods., someone that feels out of place where I live, looking for more meaning in my own life., finding something that helps define who I am more solidly., a naturally curious person who looks into just about anything, sometimes difficult to define even by myself., someone who is concerned with the wellbeing of his friends and loved ones, someone who has trouble keeping things going for myself, unsure about my own future, someone who has a lot of trouble with thinking about myself.	A recent graduate trying to survive in a turbulent economy, among the first in my generation to graduate from college, seeking to get a career in video game design, wanting for the economy to get better, wanting to move out the house away from my family, trying to seek a job., wondering where will I be in the next 5 years, not as self-driven as I thought., wonders whats the next step going forward, wondering if I'm actually prepared as an adult, by myself in life, surrounded by people, but feeling still alone, struggling to find ways to keep living, persevering despite having no will to live, wondering what the future would even hold, myself, the decision I made coming from me, and never been swayed by anyone else., despite everything else, I still live, wanting to live, resilient, even if the world would burn one day, unwavering.
Balanced Content	Father, Son, Brother, Friend, Firefighter,	A woman, a Black American, bold, beautiful,	A brother, a parent, a husband, a son, happy,

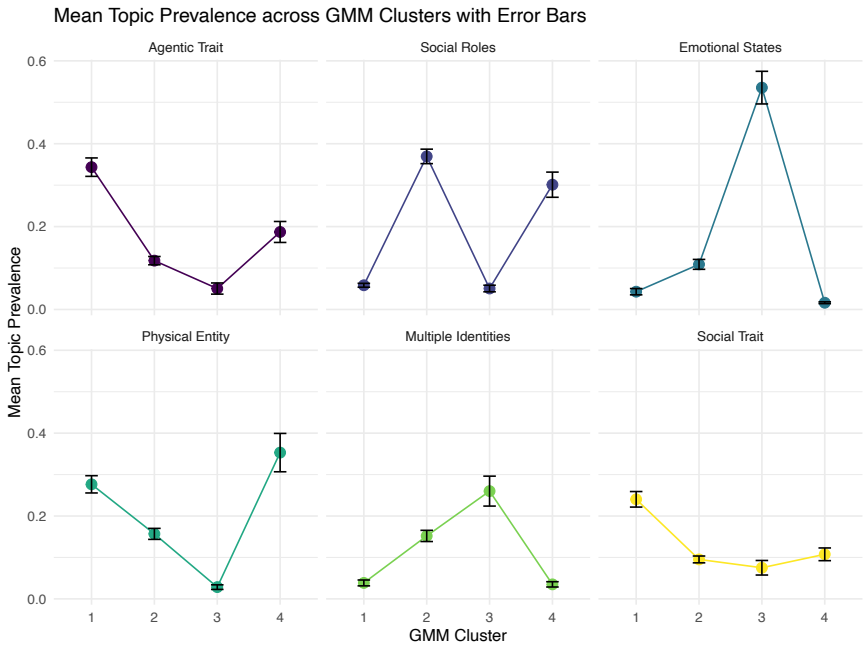
	Paramedic, Caring, Giving, Selfless, Worker, Multicultural, Catholic, Strong, Kind, Persevering, Athletic, Understanding, Passionate, Stubborn, Family	courageous, a doctor, energetic, married, inspired, logical, happy, active, accepting, excited, a leader, a conqueror, an overcomer, very ambitious, a mother	excited, knowledgeable, loving, caring, passionate, loyal, kind, tech-savvy, a sports fan, a movie buff, a coffee addict, strong, curious, fun, a foodie

I applied Gaussian Mixture Model (GMM), a clustering approach that groups people most similar to one another across text embeddings mapped onto semantic space. I labeled the clusters across (1) top words and (2) highly associated self-concept. I labeled cluster 1 as "Personality Traits and States" such that the participants' self-beliefs were predominantly comprised of personality traits and physical, emotional states. The top words depicted in Figure B.3 word cloud include "smart", "strong", "creative", "caring", "funny." The highly associated self-concepts were almost exclusively personality traits and physical/emotional states. Cluster 2 is "Roles and Relationships," such that self-beliefs were predominantly social category membership and relational roles. The top words depicted in the word cloud include "daughter," "woman," "lover," "friend", "american." The highly associated self-concepts were predominantly loci—gender, race, and age—and family roles. Occupation and hobbies often appeared too. Cluster 3 is "Self-improvement and Growth."

The top words depicted in Figure B.3 word cloud include "future," "enjoys," "time," "life," "tired." This cluster shows emotional depth and aspirations (e.g., future-oriented). For example, the highest associated self-concept includes: "striving to be better," "fully capable of learning," goal-oriented," "struggling with creating a vision for my life." The other representative self-concepts for this cluster include "passionate" and gerunds like "finding," "struggling," "trying." Cluster 4 is "Positive, Mixture." The top words depicted in Figure B.3 include "friend," "strong," "smart," "happy," "lover." Overall, the top words and representative

self-concepts contain a mixture of positively valenced content: personality traits, loci (e.g., age, gender, and race), and family roles.

Figure B.6. Mean STM topic prevalence across GMM clusters



Across STM and LLM, I converged on the self-concept typologies: personality traits/attributes, emotional states/personal growth, social roles, and a well-rounded mixture of self-concept categories. Agentic and social traits (STM topics 1 and 6) were associated with GMM cluster personality traits and states (cluster 1), indicating the GMM consolidated the two topics into a cluster. Emotional states as a topic was positively associated with GMM cluster—self-improvement and personal growth (cluster 3). STM topic around social roles was highly associated with GMM cluster—social roles and relationships (cluster 2). STM topic 4 and GMM cluster 4 reflected a positive, mixture of self-concept categories (e.g., intelligent, sporty, tall, father, husband). Topic 5, however, uniquely focused on participants who described navigating multiple identities.

Table B.6. Contradictory pairs: Significant mean contrasts from Holm's post-hoc analysis

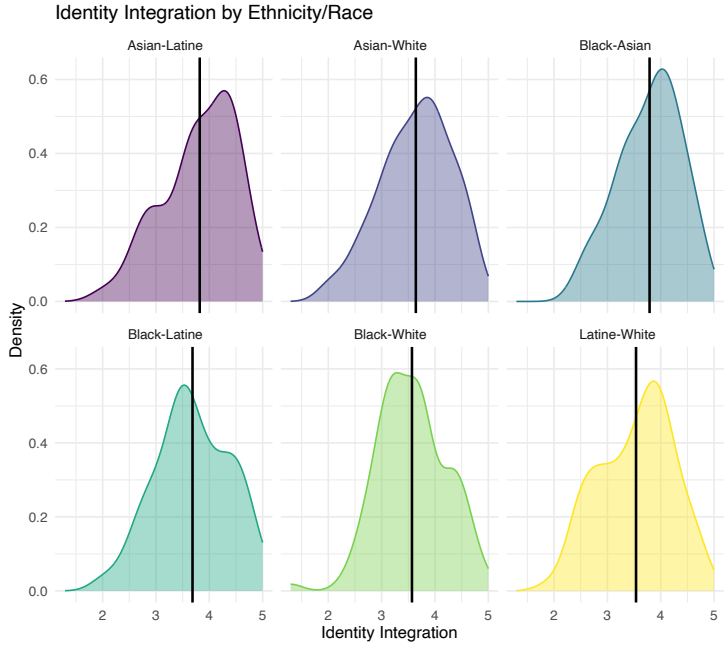
Contrast	Estimate	SE	df	t ratio	p-value	Cohens D
Asian-White – Black-Latine	-1.25	0.24	594	-5.18	<.001	-0.21
Asian-White -Black-White	-0.88	0.23	594	-3.79	0.007	-0.16
Asian-White - Latine	-1.15	0.35	594	-3.26	0.048	-0.13
Black-Latine – Latine-White	1.51	0.24	594	6.31	<.001	0.26
Black-White – Latine-White	1.14	0.23	594	4.97	<.001	0.20
Latine – Latine-White	1.41	0.35	594	4.01	0.00	0.16

Table B.6 summarizes significant pairwise differences in the number of contradictory pairs across racial groups. Overall, Asia-–White participants reported significantly fewer contradictory pairs than all other groups, including Black-Latine, Black-White, and Latine participants (ps < .05, ds ≈ .13–.21). In contrast, Black-Latine, Black-White, and Latine participants each reported significantly more contradictory pairs than Latine–White participants (ps < .001, ds ≈ .16–.26). These effects suggest that multiracial groups combining two minoritized backgrounds (e.g., Black-Latine) exhibit greater self-concept contradiction compared with those involving White heritage (e.g., Asian-White, Latine-White).

Table B.7. Self-concept and semantic coherence within participants

Self-Concept	Semantic Coherence
Shortsighted, naive, polite, meek, scatterbrained, literal, honest, unambitious, jejune, uncomplex, inconsistent, envious, observative, punctual, off-putting, organised, curious, solitary, lethargic, impatient.	0.03
A unique individual, defying categorization., A product of multiple heritages, A living testament to human interconnectedness, A complex identity shaped by various influences, A chameleon adapting to different environments., A puzzle with pieces from different backgrounds., A symphony of cultures harmonizing within me., A question mark exploring the meaning of identity., A blend of traditions and modern influences, A walking contradiction, embracing both similarities and differences, A mirror reflecting the complexities of society, A mosaic of experiences shaping my perspective, A global citizen with a local heart.	0.99

Figure B.7. Multiracial identity integration by ethnicity/race



APPENDIX C

Table C.1. Codebook for applied open-ended navigation strategies

Themes	Strategy	Definition	Sample Quote
Engagement with cultures	Practice customs	Practicing racial heritage through customs, behaviors, community gatherings, and passing traditions down to own children (e.g., cooking foods, attending cultural festivals, celebrating holidays, practicing activities).	"Participate in activities that involve both of my cultures (speak spanish around spanish people and my mom/ play basketball with my black friends)."
	Active learning	Explicit mention of learning about the customs, values, ancestry, and practices of ethnic/racial backgrounds (e.g., ancestral trees, learning about history). This includes learning a language *not* speaking it (see **Practice Customs** above).	"I try to keep learning about both of them. I have to remember that they are not static, and by changing, they allow me to find more harmony between them and myself."
Normalize Mixedness	Solidarity of similar or mixed background	Actively seeking and engaging with individuals who share similar multiracial or marginalized experiences to validate personal identity and normalize common struggles (e.g., explicit mention of feeling normal, common, not othered).	"I find people that share similar backgrounds and find common ground with them. It allows me to feel as if my experiences are normal, and that to struggle with our identity is something that isn't strange, but common. It makes my identity feel more normalized and accepted."

	Normalcy of mixedness in history	The individual alludes to how all are mixed in some way or another. Mixedness is embedded into our history.	"Almost all Hawaiians who are alive today are mixed, most probably mixed with some type of Asian, and Hawaii has its own culture based on all the ethnicities."
Social Behaviors	Frame-switching	Adaptively alternating behavior, language, and identity expression to align with varying social contexts while maintaining a core sense of self. For example, behaving in one context a certain way (e.g., very loud and open), and another way in other circumstances (e.g., polite and quiet). Language usage can allude to certain contexts: a behavioral shift is explicitly mentioned.	"I know that I do act differently when I am with certain people, but I am still true to me. I sometimes speak more Spanish with certain groups, whereas there are some friends that do not even know I am bi-lingual."
	Leave/avoid unwelcoming social situations	The individual recognizes and removes themselves from environments, interactions or social circles that undermine their well-being.	"I have just learned to accept who I am and <u>if someone doesn't want me around, then I leave</u>.
Bridge Identities	Cultural bridge/corridor	The individual implies distance between racial identities. They describe being on the fence, such that they bridge or serve as corridor between the two. For example, a participant describes racial identities as very culturally distinct, however, the individual creates the bridge themselves	"I find balance through being able to walk both paths of cultures. I feel that I can see both sides and walk between two separate lives as one person."

		("different cultures are pieces of a puzzle, and I solve how they fit together").	
	Shared values, histories, and traditions	Individual recognizes overlapping cultural values, historical narratives, and traditions that bind racial groups together (e.g., "my identities are not so different, same values and communities").	"I find they have a lot in common, their religiousness and also their love for flavorful food. They are both coming from underprivileged nations."
Ethnic/racial Identification	De-emphasis of racial identity	The individual de-emphasizes the role of race in defining one's self-concept, instead, highlighting individual traits and social roles over racial categorizations.	"I don't think about my racial and ethnic identities when I make decision. I view myself as an individual before these groups."
	Define what each racial identity means to oneself	The reflective process of recognizing, distinguishing, and internalizing each racial heritage into one's self-concept, acknowledging their unique significance. Compare and contrast what each identity contributes.	"Both identities are important to me. Asian is stronger culturally but Black is stronger in terms of how I feel I am seen in society."
	Lean into one identity over others	Selectively emphasizing one racial identity over others to align with personal comfort or social acceptance. One racial identity is prioritized.	"I tend to ignore my Black side as much as possible, and address it as needed. It hurts when people bring it up. I hope that I can find peace and harmony at some point in my life."

Finding Balance	Acceptance of dialectical self (Yin & Yang)	The individual goes with the flow, describing what feels "natural", "organic", "right." They say "I am who I am." No resistance, but self-acceptance, a hybridized "whole."	"Both races make up who I am and I don't struggle with my identity and think the balance I have is just a natural thing for me."
	Open-mindedness	A willingness to value and respect diverse cultural expressions and individual opinions. Embrace differences as aspects of human experience. It's outward focused, emphasizing open-mindedness not necessarily wholeness (e.g., "I am open to all perspectives", "all viewpoints value").	"I think less of balancing my own to cultures as much as valuing and respecting all cultures. In so viewing our differences as humans as something exciting."
	Imbalanced	Persistent difficulty reconciling multiple racial or cultural identities, leading to feeling imbalanced.	"I am always constantly struggling with it."

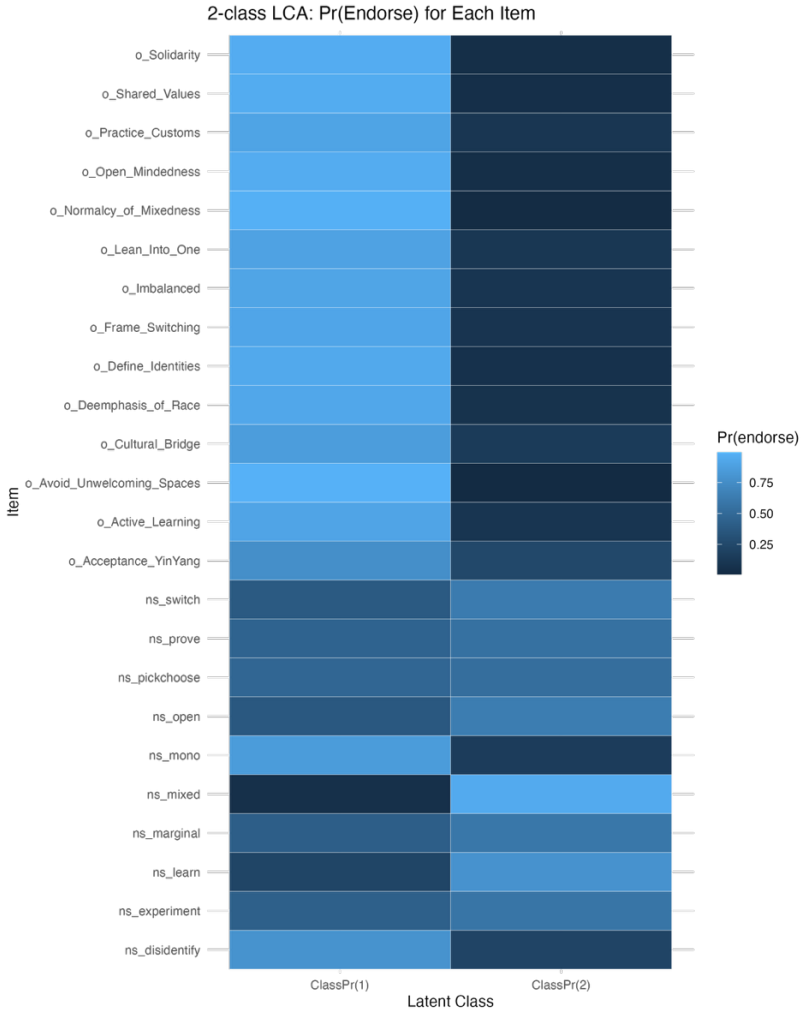
Table C.2. Simulated narrative generation and refinement

Phase	Description / Example	Reflection / Next Steps
Purpose	Simulate narratives to train research assistants (RAs) on identifying codebook themes and anticipating response variation.	Generated 25 narratives per theme to represent variability and realism.
Initial Query	"Here is the definition for codebook theme [insert theme]: [insert definition]. These narratives are [describe what makes this theme distinct]. Please create 50 narratives between 2–4 sentences."	Fifty were requested to allow for unusable outputs. In data, the mean sentence length was 2.17, between 1 to 10 sentences. This prompt helped anticipate likely narrative forms in the dataset.
Example Output	Theme: [insert Theme] Definition: [insert definition] Example narratives: [insert narratives of diverse sentence length and diverse ways of starting].	Retain a representative subset of generated examples for coder training materials.
Alternative Query (Noise-Added Simulation)	"I have a codebook theme called [insert theme] with the definition [insert definition]. Below are narratives of this theme. Can you add noise to the narratives and, in similar tone, simulate 15 new narratives with this theme?"	Another approach I tried was directly from de-identified narratives themselves. I gave GPT the narrative excerpt and asked to add noise to them instead of generate new ones entirely. Useful for preserving realistic linguistic tone.
Prompt Follow-Ups (Refinements)	1. Clarify theme boundaries: "If narratives overlap too heavily with other themes, here's how to distinguish them …" 2. Tone improvement: "Make it more conversational, less verbose, and more naturalistic." 3. Structural variety: "Avoid starting every sentence with 'I'; use diverse phrasing."	These iterative refinements improved realism and reduced thematic overlap. Follow-ups can be built into future semi-automated simulation pipelines.
Evaluation of Outputs	Compared multiple framing versions and retained prompts yielding the most humanlike and thematically distinct responses.	Identify objective metrics (e.g., linguistic diversity, sentiment range) to quantify narrative realism.
Next Steps	Fine-tune generation parameters (temperature, top-p). Use automated pre-	Incorporate this simulation framework as a reproducible

	screening before manual review. Expand simulation to underrepresented themes.	part of coder training and codebook calibration.

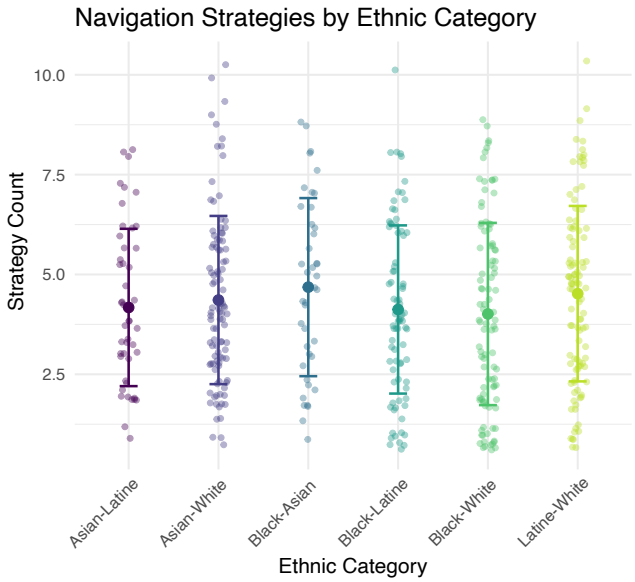
Table C.2 outlines the iterative, instructive process for simulating narratives as part of a training dataset. While promising, the approach remains barebones and needs some refinement. The simulated narratives for RAs to train captured key thematic patterns, but interrater agreement was lower than expected for group 2 in the real human coding, perhaps because some narratives were too straightforward, lacking the ambiguity of real data. I also recognize that working with fourteen themes posed challenges, yet this granularity was important: these categories represent the strategies described, and distinctions between emerging ideas were most informative. Collapsing them would risk erasing nuance in how participants express these kinds of strategies. After getting feedback, the RAs found it useful. With a broader and more nuanced narrative set, and more structured practice, perhaps future iterations could strengthen reliability and coder intuition.

Figure C.1. LCA using open- and close-ended strategies



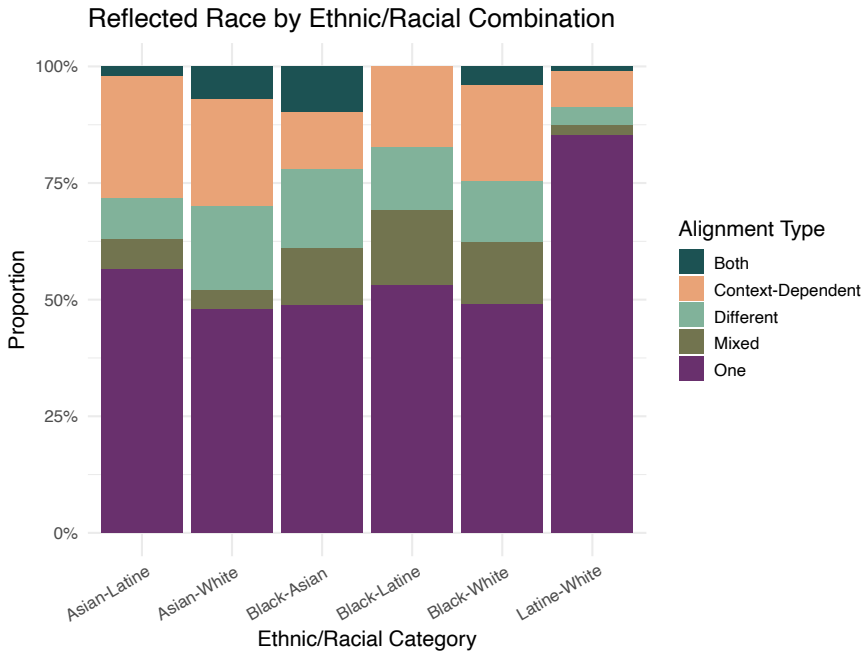
Initially, I ran LCA using the open- and close-ended navigation strategies. The fit indices showed the 2-class model as optimal; however, the classes are non-interpretable. Most participants' open-ended responses were coded as one strategy. It doesn't make sense that one entire class would have low endorsement of *all* open-ended strategies.

Figure C.2. Number of close- and open-ended navigation strategies endorsed by ethnic/racial combinations.



As shown in Figure C.2, ethnic/racial categories do not appear to notably vary in the number of navigation strategies endorsed, on average.

Figure C.3. Reflected race category by ethnic/racial combination



The reflected race is the race they think they are perceived by others. I coded across the following four categories: context-dependent (how they're perceived varies by context); different

(individuals are classified by others as a different racial identity than either of their racial identities); mixed (they're perceived by others as mixed); and one (they are perceived as only one of their racial identities). In Figure C.3, most participants, regardless of racial backgrounds, describe their reflected race as one racial background. Asian-Latine, Black-Asian, Black-Latine, and Black-White individuals describe themselves as perceived more as mixed in general than of their specific racial combinations. The inverse is found for Asian-White individuals. Most Latine-White individuals say they are perceived as one of their racial identities (and does not vary by context). Black-Asian and Black-White individuals describe their perceived race as different than their racial identities most often.

Figure C.4. Multiracial individuals who endorsed self-identifying as "monoracial" by ethnic/racial combination

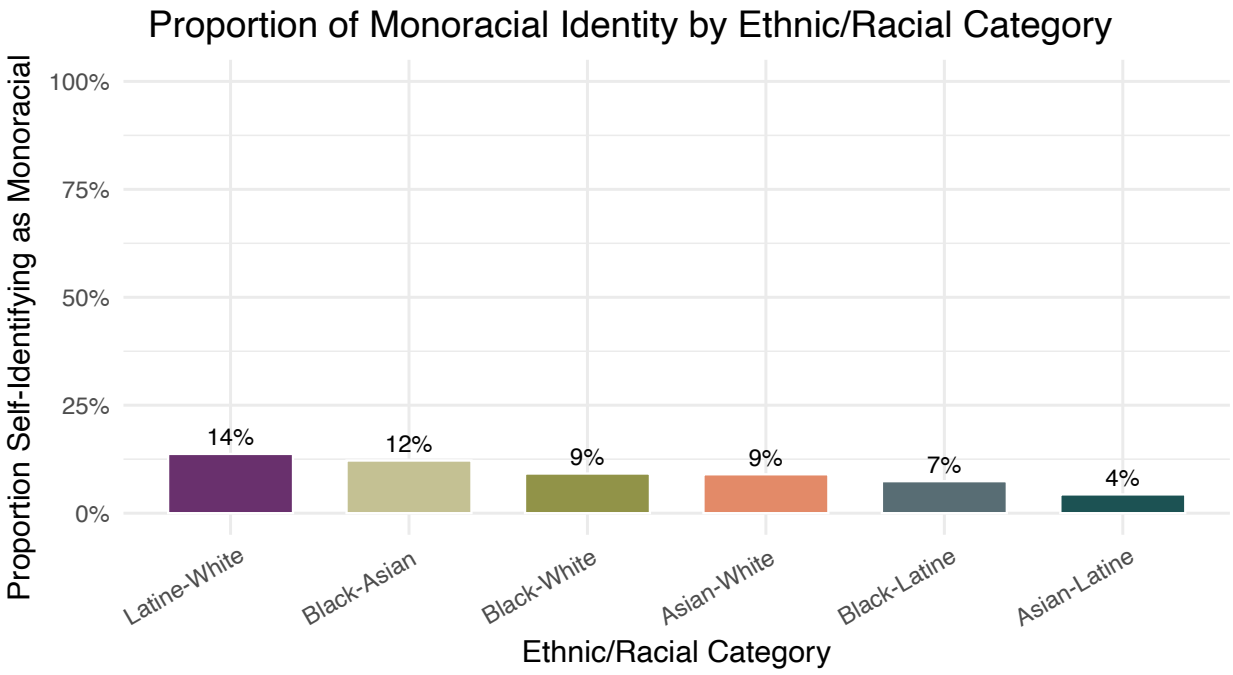


Figure C.5. Mean number of strategies endorsed per LCA class in 3-class model

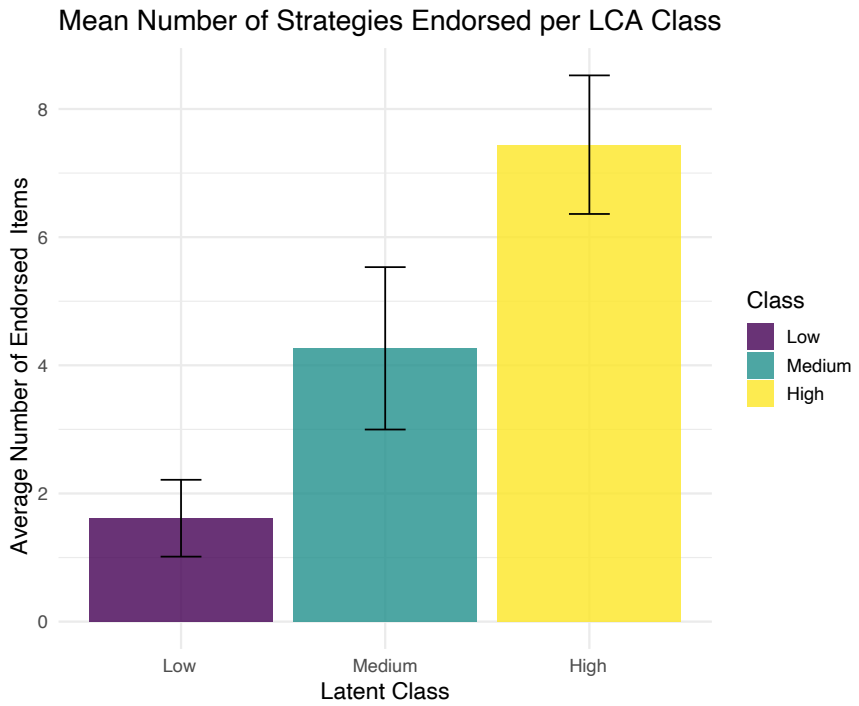
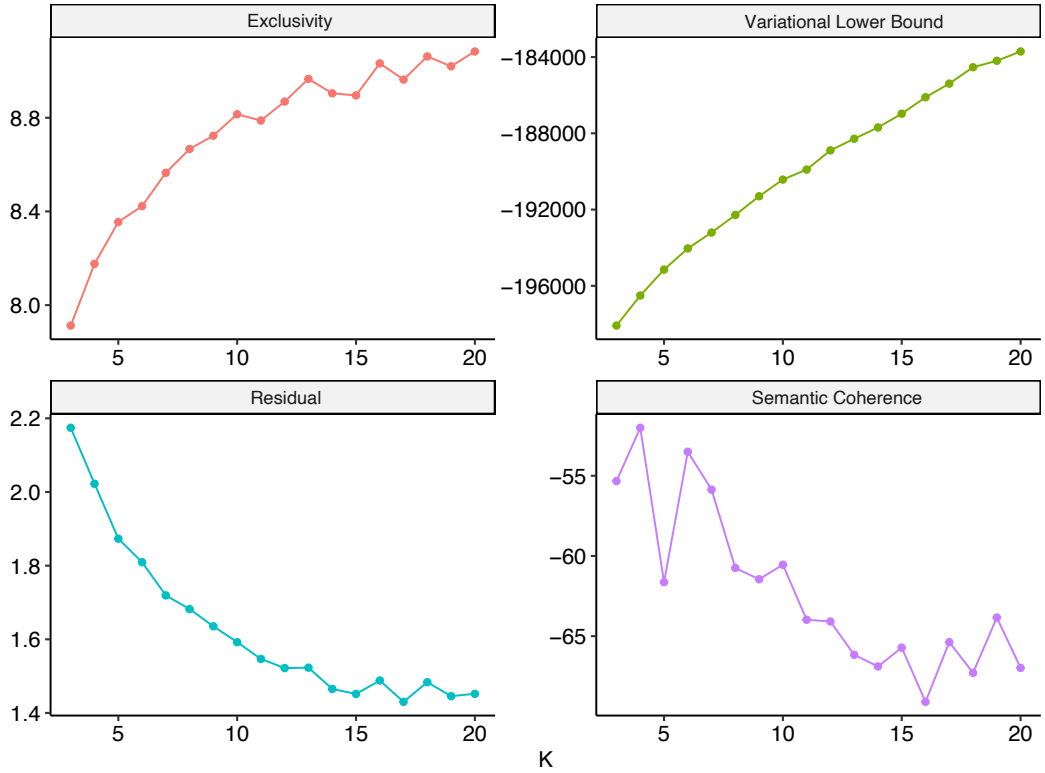


Figure C.6. Four metrics to select optimal number of topics

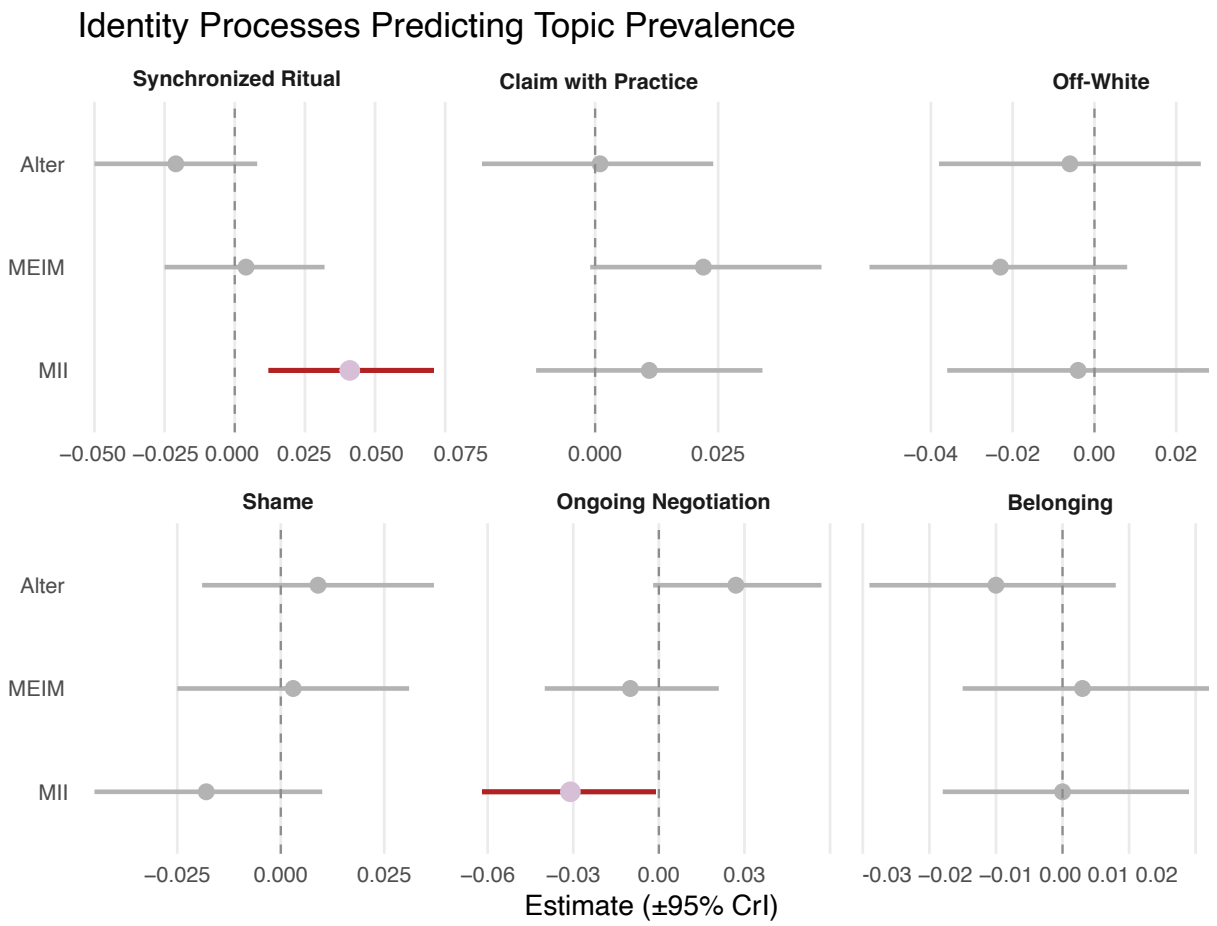


In Figure C.6, I evaluate the four diagnostic metrics introduced in Studies 1 and 2 (exclusivity, semantic coherence, residuals, and variational lower bound) to determine the optimal topic solution. Ideally, the selected model should exhibit high exclusivity and semantic coherence alongside low residuals and a low variational lower bound. In the six-topic model, semantic coherence increases sharply before declining, exclusivity and the variational lower bound continue to rise, and residuals drop steeply, indicating a well-balanced trade-off among these metrics. Based on this pattern, I selected the six-topic model as the optimal solution.

Figure C.7. Word clouds for types of rites of passage described by ethnicity/race

In Appendix C Figure 7 are FREX words as word clouds for ethnic/racial combination across two topics, *Synchronized ritual* and *Shame towards culture/history*. Participants universally describe family gatherings, words like "together," "party," and "wedding" signal rituals that foster relational cohesion and shared belonging. At the same time, terms such as "sides" and "traditions" highlight cultural negotiation, as individuals navigate and compare the norms of their heritages. The Latine-White combination introduces a more melancholic tone, with "sad" and "passed" suggesting funerals as cultural ritual. While *Synchronized Ritual* emphasizes togetherness and shared ritual, another theme reveals the more complicated emotional terrain of shame and cultural self-consciousness. Across groups, participants turn inward, reflecting on how appearance and knowledge shape their relationship to heritage. Words like *"feel," "look,"* and *"enough"* convey embodied awareness and perceived inadequacy, while *"history," "book", "taught,"* and *"raise"* suggest that learning and transmission of culture take place in this rite of passage.

Figure C.8. Rite of passage topic prevalence by identity processes



To examine how individual differences in identity processes relate to the rites of passage, six Bayesian linear regression models were estimated. Each model predicted the prevalence of a rite of passage derived from topic modeling, using the three continuous predictors that captured key aspects of identity development: MII, MEIM, and MISS (alternating identity styles). Models assumed a Gaussian likelihood with weakly informative priors centered at zero, and all parameters were estimated from the posterior distribution using MCMC sampling. Posterior means and 95% credible intervals (CrIs) were extracted to represent the most probable range of parameter values given the data. Model convergence and chain stability were confirmed by inspection of diagnostics and trace plots.

Figure C.8 displays the posterior distributions of regression estimates from the Bayesian model predicting topic prevalence from identity processes across six types of liminal life events. Each posterior represents the range of credible values for the effect of MII on a given thematic topic, with the mean (dot) and 95% credible intervals (horizontal bars) indicating the direction and uncertainty of the association. Across the six types of liminal life events, the posterior estimates suggest that multiracial identity integration is the most consistent predictor of thematic variation in rites of passage. Specifically, higher MII scores were positively associated with the prevalence of the "synchronized ritual" topic, indicating that participants who experience greater harmony between their racial identities tend to describe life events where their customs, practices, or traditions coexist. In contrast, MII was negatively associated with the prevalence of the "ongoing negotiation", suggesting that individuals with lower integration report more dynamic or unresolved identity experiences. The credible intervals for these effects indicate strong evidence for these associations.

REFERENCES

Abdulrahim, S., James, S. A., Yamout, R., & Baker, W. (2012). Discrimination and psychological distress: Does Whiteness matter for Arab Americans? *Social Science & Medicine*, *75*(12), 2116–2123. https://doi.org/10.1016/j.socscimed.2012.07.030

Ackerman-Barger, K., Boatright, D., Gonzalez-Colaso, R., Orozco, R., & Latimore, D. (2020). Seeking Inclusion Excellence: Understanding Racial Microaggressions as Experienced by Underrepresented Medical and Nursing Students. *Academic Medicine*, *95*(5), 758–763. https://doi.org/10.1097/ACM.0000000000003077

Adams, G., & Markus, H. R. (2004). Toward a Conception of Culture Suitable for a Social Psychology of Culture. In *The psychological foundations of culture.* (pp. 335–360). Lawrence Erlbaum Associates Publishers.

Adler, J. M., Dunlop, W., Fivush, R., Lilgendahl, J. P., Lodi-Smith, J., McAdams, D. P., McLean, K. C., Pasupathi, M., & Syed, M. (2017). *Research Methods for Studying Narrative Identity: A Primer* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/cz3q3

Adler, J. M., & Poulin, M. J. (2009). The Political Is Personal: Narrating 9/11 and Psychological Well-Being. *Journal of Personality*, *77*(4), 903–932. https://doi.org/10.1111/j.1467-6494.2009.00569.x

Akaike, H. (2003). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Albuja, A. F., Sanchez, D. T., & Gaither, S. E. (2018). Fluid racial presentation: Perceptions of contextual "passing" among biracial people. *Journal of Experimental Social Psychology*, *77*, 132–142. https://doi.org/10.1016/j.jesp.2018.04.010

Albuja, A. F., Sanchez, D. T., & Gaither, S. E. (2020). Intra-race intersectionality: Identity denial among dual-minority biracial people. *Translational Issues in Psychological Science*, *6*(4), 392–403. https://doi.org/10.1037/tps0000264

Albuja, A. F., West, A., & Gaither, S. E. (2022). *Two or More: A Comparative Analysis of Multiracial and Multicultural Research*. Cambridge University Press; Cambridge Core. https://doi.org/10.1017/9781009202695

Altweck, L., & Marshall, T. C. (2015). When You Have Lived in a Different Culture, Does Returning 'Home' Not Feel Like Home? Predictors of Psychological Readjustment to the Heritage Culture. *PLOS ONE*, *10*(5), e0124393. https://doi.org/10.1371/journal.pone.0124393

Amiot, C., de la Sablonnière, R., Terry, D., & Smith, J. (2007). Integration of Social Identities in the Self: Toward a Cognitive-Developmental Model. *Personality and Social Psychology Review : An Official Journal of the Society for Personality and Social Psychology, Inc*, *11*, 364–388. https://doi.org/10.1177/1088868307304091

Axt, J. R., Ebersole, C. R., & Nosek, B. A. (2014). The Rules of Implicit Evaluation by Race, Religion, and Age. *Psychological Science*, *25*(9), 1804–1815. https://doi.org/10.1177/0956797614543801

Bakan, D. (1966). *The duality of human existence: An essay on psychology and religion*. Rand Mcnally.

Bamberg, M. (2011). Who am I? Narration and its contribution to self and identity. *Theory & Psychology*, *21*(1), 3–24. https://doi.org/10.1177/0959354309355852

Bauer, J. J., & McAdams, D. P. (2004). Personal Growth in Adults' Stories of Life Transitions. *Journal of Personality*, *72*(3), 573–602. https://doi.org/10.1111/j.0022-3506.2004.00273.x

Baumeister, R. F. (1999). *The self in social psychology.* (pp. ix, 492). Psychology Press.

Bean, F. D. (2018). Growing U.S. Ethnoracial Diversity: A Positive or Negative Societal Dynamic? *The ANNALS of the American Academy of Political and Social Science*, *677*(1), 229–239. https://doi.org/10.1177/0002716218765415

Bean, F. D., & Lee, J. (2009). Plus ça Change…? Multiraciality and the Dynamics of Race Relations in the United States. *Journal of Social Issues*, *65*(1), 205–219. https://doi.org/10.1111/j.1540-4560.2008.01595.x

Benet-Martínez, V., & Haritatos, J. (2005). Bicultural Identity Integration (BII): Components and Psychosocial Antecedents. *Journal of Personality*, *73*(4), 1015–1050. https://doi.org/10.1111/j.1467-6494.2005.00337.x

Benet-Martínez, V., Hong, Y., Cheng, C.-Y., Lee, F., Benet-Martínez, V., & Huynh, Q.-L. (2014). Variations in Multicultural Experience. In V. Benet-Martínez & Y. Hong (Eds.), *The Oxford Handbook of Multicultural Identity*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199796694.013.025

Benkler, N., Friedman, S., Schmer-Galunder, S., Mosaphir, D., Sarathy, V., Kantharaju, P., McLure, M. D., & Goldman, R. P. (2022). Cultural Value Resonance in Folktales: A Transformer-Based Analysis with the World Value Corpus. *Social, Cultural, and Behavioral Modeling: 15th International Conference, SBP-BRiMS 2022, Pittsburgh, PA, USA, September 20–23, 2022, Proceedings*, 209–218. https://doi.org/10.1007/978-3-031-17114-7_20

Benoit, L., Harf, A., Sarmiento, L., Skandrani, S., & Moro, M. R. (2018). Shifting views and building bonds: Narratives of internationally adopted children about their dual culture. *Transcultural Psychiatry*, *55*(3), 405–427.

lei, David & Ng, Andrew & Jordan, Michael. (2001). Latent Dirichlet Allocation. *The Journal of Machine Learning Research. 3*. 601-608.

Bonilla-Silva, E. (2025). Rethinking racism again: Theorizing the racial structure "for real." *Ethnic and Racial Studies*, 1–17. https://doi.org/10.1080/01419870.2025.2570470

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, *18*(3), 328–352. https://doi.org/10.1080/14780887.2020.1769238

Brittian, A. S., Umaña-Taylor, A. J., & Derlan, C. L. (2013). An examination of biracial college youths' family ethnic socialization, ethnic identity, and adjustment: Do self-identification labels and university context matter? *Cultural Diversity & Ethnic Minority Psychology*, *19*(2), 177.

Brock, G., Pihur, V., Datta, S., & Datta, S. (2008). clValid: An R package for cluster validation. *Journal of Statistical Software*, *25*, 1–22.

Brondolo, E., Gallo, L. C., & Myers, H. F. (2009). Race, racism and health: Disparities, mechanisms, and interventions. *Journal of Behavioral Medicine*, *32*(1), 1–8. https://doi.org/10.1007/s10865-008-9190-3

Bronfenbrenner, U. (1992). Ecological systems theory. *Six Theories of Child Development: Revised Formulations and Current Issues.*, 187–249.

Brown, A., Zhu, J., Abdelwahab, M., Dong, A., Wang, C., & Rose, J. (2024). Generation, distillation and evaluation of motivational interviewing-style reflections with a foundational language model. *arXiv Preprint arXiv:2402.01051*.

Brown, T. N., Emily E. Tanner-Smith, Chase L. Lesane-Brown, & Ezell, M. E. (2007). Child, Parent, and Situational Correlates of Familial Ethnic/Race Socialization. *Journal of Marriage and Family*, *69*(1), 14–25. JSTOR.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*, 1–28.

Burnham, M. (2024). What is sentiment meant to mean to language models? *Research & Politics*, *11*(4), 20531680241307941. https://doi.org/10.1177/20531680241307941

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186. https://doi.org/10.1126/science.aal4230

Camarena, P. M., Sarigiani, P. A., & Petersen, A. C. (1990). Gender-specific pathways to intimacy in early adolescence. *Journal of Youth and Adolescence*, *19*(1), 19–32. https://doi.org/10.1007/BF01539442

Campbell, K. M. (2023). Mitigating the Isolation of Minoritized Faculty in Academic Medicine. *Journal of General Internal Medicine*, *38*(7), 1751–1755. https://doi.org/10.1007/s11606-022-07982-8

Cer, D., Yang, Y., Kong, S., Hua, N., Limtiaco, N., St. John, R., Constant, N., Guajardo-Cespedes, M., Yuan, S., Tar, C., Strope, B., & Kurzweil, R. (2018). Universal Sentence Encoder for English. In E. Blanco & W. Lu (Eds.), *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* (pp. 169–174). Association for Computational Linguistics. https://doi.org/10.18653/v1/D18-2029

Chae, D. H., Yip, T., Martz, C. D., Chung, K., Richeson, J. A., Hajat, A., Curtis, D. S., Rogers, L. O., & LaVeist, T. A. (2021). Vicarious racism and vigilance during the COVID-19 pandemic: Mental health implications among Asian and Black Americans. *Public Health Reports*, *136*(4), 508–517.

Charlesworth, T., Yang, V., Mann, T., Kurdi, B., & Banaji, M. (2021). Gender Stereotypes in Natural Language: Word Embeddings Show Robust Consistency Across Child and Adult Language Corpora of More Than 65 Million Words. *Psychological Science*, *32*, 095679762096361. https://doi.org/10.1177/0956797620963619

Charmaraman, L., & Grossman, J. M. (2010). Importance of race and ethnicity: An exploration of Asian, Black, Latino, and multiracial adolescent identity. *Cultural Diversity and Ethnic Minority Psychology*, *16*(2), 144–151. https://doi.org/10.1037/a0018668

Charmaz, K. (1995). The body, identity, and self: Adapting to impairment. *The Sociological Quarterly*, *36*(4), 657–680. https://doi.org/10.1111/j.1533-8525.1995.tb00459.x

Charmaz, K. (2006). Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis. In *Introducing Qualitative Methods* (Vol. 1).

Chavez, G. F., & Sanchez, D. T. (2010). A Clearer Picture of Multiracial Substance Use: Rates and Correlates of Alcohol and Tobacco Use in Multiracial Adolescents and

Adults. *Race and Social Problems*, *2*(1), 1–18. https://doi.org/10.1007/s12552-010-9023-1
Chávez-Moreno, L. C. (2021). The Problem With Latinx as a Racial Construct vis-à-vis Language and Bilingualism. In E. G. Murillo, D. D. Bernal, S. Morales, L. Urrieta, E. R. Bybee, J. S. Muñoz, V. Sáenz, D. Villanueva, M. Machado-Casas, & K. Espinoza, *Handbook of Latinos and Education* (2nd ed., pp. 164–180). Routledge. https://doi.org/10.4324/9780429292026-19
Cheng, C., & Lee, F. (2009). Multiracial Identity Integration: Perceptions of Conflict and Distance among Multiracial Individuals. *Journal of Social Issues*, *65*(1), 51–68. https://doi.org/10.1111/j.1540-4560.2008.01587.x
Cheryan, S., & Monin, B. (2005). Where are you really from?: Asian Americans and identity denial. *Journal of Personality and Social Psychology*, *89*(5), 717–730. https://doi.org/10.1037/0022-3514.89.5.717
Childs, E. C., Lyons, A., & Jones, S. L. (2021). Migrating mixedness: Exploring mixed identity development in New York City. *Journal of Ethnic and Migration Studies*, *47*(4), 782–801. https://doi.org/10.1080/1369183X.2019.1654153
Clasen, D. R., & Brown, B. B. (1985). The multidimensionality of peer pressure in adolescence. *Journal of Youth and Adolescence*, *14*(6), 451–468. https://doi.org/10.1007/BF02139520
Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104
Collins, L. M., & Lanza, S. T. (2009). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. John Wiley & Sons.
Collins, P. H., Da Silva, E. C. G., Ergun, E., Furseth, I., Bond, K. D., & Martínez-Palacios, J. (2021). Intersectionality as Critical Social Theory: Intersectionality as Critical Social Theory, Patricia Hill Collins, Duke University Press, 2019. *Contemporary Political Theory*, *20*(3), 690–725. https://doi.org/10.1057/s41296-021-00490-0
Condon, D. M. (2018). *The SAPA Personality Inventory: An empirically-derived, hierarchically-organized self-report personality assessment model*.
Connolly, J. A., & McIsaac, C. (2009). Romantic relationships in adolescence. In *Handbook of adolescent psychology: Contextual influences on adolescent development, Vol. 2, 3rd ed.* (pp. 104–151). John Wiley & Sons, Inc. https://doi.org/10.1002/9780470479193.adlpsy002005
Crenshaw, K. (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. *University of Chicago Legal Forum*, *1989*(1).
Cross, S. E., & Madson, L. (1997). Models of the Self: Self-Construals and Gender. *Psychological Bulletin*, *122*(1), 5–37. https://doi.org/10.1037/0033-2909.122.1.5
Cross, W. Jr., & Cross, T. (2008). Theory, research, and models. In *Handbook of race, racism, and the developing child* (Quintana SM&McKown C, pp. 154–181). Wiley.
Csizmadia, A., & Atkin, A. L. (2022). Supporting children and youth in multiracial families in the United States: Racial-ethnic socialization and familial support of multiracial experiences. *Journal of Child and Family Studies*, *31*(3), 664–674. https://doi.org/10.1007/s10826-022-02265-6

Cutler, A., & Condon, D. M. (2023). Deep lexical hypothesis: Identifying personality structure in natural language. *Journal of Personality and Social Psychology*, *125*(1), 173–197. https://doi.org/10.1037/pspp0000443

Defranza, D., Mishra, H., & Mishra, A. (2020). How Language Shapes Prejudice Against Women: An Examination Across 45 World Languages. *Journal of Personality and Social Psychology*, *119*, 7–22. https://doi.org/10.1037/pspa0000188

Denny, Matthew (Penn State University); Spirling, Arthur (New York University). (2017). *Replication Data for: Text Preprocessing For Unsupervised Learning: Why It Matters, When It Misleads, And What To Do About It* [Dataset]. Harvard Dataverse. https://doi.org/10.7910/DVN/XRR0HM

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *Bert: Pre-training of deep bidirectional transformers for language understanding*. 4171–4186.

Diedenhofen, B., & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS ONE*, *10*(4). https://doi.org/10.1371/journal.pone.0121945

Dunlop, W. L., & Tracy, J. L. (2013). Sobering stories: Narratives of self-redemption predict behavioral change and improved health among recovering alcoholics. *Journal of Personality and Social Psychology*, *104*(3), 576–590. https://doi.org/10.1037/a0031185

Dunlop, W. L., & Wilkinson Westberg, D. (2022). On stories, conceptual space, and physical place: Considering the function and features of stories throughout the narrative ecology. *Personality Science*, *3*, e7337. https://doi.org/10.5964/ps.7337

Eagly, A. H., & Revelle, W. (2022). Understanding the Magnitude of Psychological Differences Between Women and Men Requires Seeing the Forest and the Trees. *Perspectives on Psychological Science*, *17*(5), 1339–1358. https://doi.org/10.1177/17456916211046006

Erikson, E. H. (1968). *Identity youth and crisis* (Issue 7). WW Norton & company.

Ethayarajh, K. (2019). *How Contextual are Contextualized Word Representations? Comparing the Geometry of BERT, ELMo, and GPT-2 Embeddings* (p. 65). https://doi.org/10.18653/v1/D19-1006

Feinerer, I. (2011). *Introduction to the tm Package Text Mining in R*.

Fish, J., & Syed, M. (2021). Digital storytelling methodologies: Recommendations for a participatory approach to engaging underrepresented communities in counseling psychology research. *Journal of Counseling Psychology*, *68*(3), 271–285. https://doi.org/10.1037/cou0000532

Fraley, C., & Raftery, A. E. (1998). How Many Clusters? Which Clustering Method? Answers Via Model-Based Cluster Analysis. *The Computer Journal*, *41*(8), 578–588. https://doi.org/10.1093/comjnl/41.8.578

Franco, M. G., & O'Brien, K. M. (2018). Racial identity invalidation with multiracial individuals: An instrument development study. *Cultural Diversity and Ethnic Minority Psychology*, *24*(1), 112–125. https://doi.org/10.1037/cdp0000170

Fu, Q., Zhuang, Y., Gu, J., Zhu, Y., & Guo, X. (2021). Agreeing to Disagree: Choosing Among Eight Topic-Modeling Methods. *Big Data Research*, *23*, 100173. https://doi.org/10.1016/j.bdr.2020.100173

Gaither, S. E. (2015). “Mixed” Results: Multiracial Research and Identity Explorations. *Current Directions in Psychological Science*, *24*(2), 114–119. https://doi.org/10.1177/0963721414558115

Galliher, R. V., McLean, K. C., & Syed, M. (2017). An integrated developmental model for studying identity content in context. *Developmental Psychology*, *53*(11), 2011–2022. https://doi.org/10.1037/dev0000299

Garay, M. M., & Remedios, J. D. (2021). A review of White-centering practices in multiracial research in social psychology. *Social and Personality Psychology Compass*, *15*(10). https://doi.org/10.1111/spc3.12642

Gee, J. P. (2014). *An introduction to discourse analysis: Theory and method*. routledge.

Geertz, C. (1973). *The interpretation of cultures: Selected essays*. New York : Basic Books. https://search.library.wisc.edu/catalog/9910006648402121

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). *A weakly informative default prior distribution for logistic and other regression models*.

Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, *7*, 457–472.

Giamo, L. S., Schmitt, M. T., & Outten, H. R. (2012). Perceived discrimination, group identification, and life satisfaction among multiracial people: A test of the rejection-identification model. *Cultural Diversity and Ethnic Minority Psychology*, *18*(4), 319–328. https://doi.org/10.1037/a0029729

Glaser, B. (1992). *Basics of Grounded Theory Analysis*. Sociology Press.

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PloS One*, *12*(6), e0174035.

Gonzales-Backen, M. A. (2013). An Application of Ecological Theory to Ethnic Identity Formation Among Biethnic Adolescents. *Family Relations*, *62*(1), 92–108. https://doi.org/10.1111/j.1741-3729.2012.00749.x

Grotevant, H. D., & Cooper, C. R. (1986). Individuation in family relationships: A perspective on individual differences in the development of identity and role-taking skill in adolescence. *Human Development*, *29*(2), 82–100. https://doi.org/10.1159/000273025

Guimond, S., & Chatard, A. (2014). *Basic principles of social comparison: Does gender matter?* (pp. 205–229). https://doi.org/10.1017/CBO9781139035583.013

Gullickson, A., & Morning, A. (2011). Choosing race: Multiracial ancestry and identification. *Social Science Research*, *40*(2), 498–512. https://doi.org/10.1016/j.ssresearch.2010.12.010

Hammack, P. L. (2008). Narrative and the Cultural Psychology of Identity. *Personality and Social Psychology Review*, *12*(3), 222–247. https://doi.org/10.1177/1088868308316892

Harris, D. R., & Sim, J. J. (2002). Who is multiracial? Assessing the complexity of lived race. *American Sociological Review*, *67*(4), 614–627.

Harter, S. (2012). *The construction of the self: Developmental and sociocultural foundations, 2nd ed.* (pp. xxiii, 440). The Guilford Press.

Heine, S. J. (2001). Self as Cultural Product: An Examination of East Asian and North American Selves. *Journal of Personality*, *69*(6), 881–905. https://doi.org/10.1111/1467-6494.696168

Heine, S. J. (2012). *Cultural Psychology*. W.W. Norton. https://books.google.com/books?id=TkDmZwEACAAJ

Henning, G., Segel-Karpas, D., Bjälkebring, P., & Berg, A. I. (2022). Autonomy and loneliness – longitudinal within- and between-person associations among Swedish older adults. *Aging & Mental Health*, *26*(12), 2416–2423. https://doi.org/10.1080/13607863.2021.2000937

Herman, M. (2004). Forced to Choose: Some Determinants of Racial Identification in Multiracial Adolescents. *Child Development*, *75*(3), 730–748. https://doi.org/10.1111/j.1467-8624.2004.00703.x

Hitlin, S., Scott Brown, J., & Elder Jr., G. H. (2006). Racial Self-Categorization in Adolescence: Multiracial Development and Social Pathways. *Child Development*, *77*(5), 1298–1308. https://doi.org/10.1111/j.1467-8624.2006.00935.x

Hobson, D. G., Zhou, H., Ruths, D., & Piper, A. (2024). Story Morals: Surfacing value-driven narrative schemas using large language models. In Y. Al-Onaizan, M. Bansal, & Y.-N. Chen (Eds.), *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing* (pp. 12998–13032). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.emnlp-main.723

Hong, Y., Morris, M. W., Chiu, C., & Benet-Martínez, V. (2000). Multicultural minds: A dynamic constructivist approach to culture and cognition. *American Psychologist*, *55*(7), 709–720. https://doi.org/10.1037/0003-066X.55.7.709

Hordge-Freeman, E., & Veras, E. (2020). Out of the Shadows, into the Dark: Ethnoracial Dissonance and Identity Formation among Afro-Latinxs. *Sociology of Race and Ethnicity*, *6*(2), 146–160. https://doi.org/10.1177/2332649219829784

Huynh, Q.-L., Devos, T., & Smalarz, L. (2011). Perpetual foreigner in one's own land: potential implications for identity and psychological adjustment. *Journal of social and clinical psychology*, *30*(2), 133–162. https://doi.org/10.1521/jscp.2011.30.2.133

Impett, E., & Peplau, L. (2006). *"His" and "Her" Relationships? A Review of the Empirical Evidence* (pp. 273–292). https://doi.org/10.1017/CBO9780511606632.016

Iverson, S., Morning, A., Saperstein, A., & Xu, J. (2022). Regimes beyond the One-Drop Rule: New Models of Multiracial Identity. *Genealogy*, *6*(2). https://doi.org/10.3390/genealogy6020057

Jackson, J. J., & Harms, P. (2019). *An Idiographic Approach to Personality Development: Examining Change and Consistency in Self-Concept*. https://doi.org/10.31234/osf.io/qdp98

Jackson, K. F. (2012). Living the Multiracial Experience: Shifting Racial Expressions, Resisting Race, and Seeking Community. *Qualitative Social Work*, *11*(1), 42–60. https://doi.org/10.1177/1473325010375646

Jackson, K. F., Yoo, H. C. (Brandon), Guevarra Jr., R., & Harrington, B. A. (2012). Role of identity integration on the relationship between perceived racial discrimination and psychological adjustment of multiracial people. *Journal of Counseling Psychology*, *59*(2), 240–250. https://doi.org/10.1037/a0027639

Jacobson, R., Westberg, D. W., Chou, E., Syed, M., & Weston, S. J. (2024). Using structural topic modeling to understand ethnicity-related narratives. *Journal of Personality*, *92*(6), 1683–1703. https://doi.org/10.1111/jopy.12966

Johnson, K. (2024). Rejecting Multiracial Stereotypes: Parental Socialization Practices at the Intersection of Race and Gender. *Sociology of Race and Ethnicity*, *10*(4), 521–536. https://doi.org/10.1177/23326492241264219

Johnson, S. J., Murty, M. R., & Navakanth, I. (2024). A detailed review on word embedding techniques with emphasis on word2vec. *Multimedia Tools and Applications*, *83*(13), 37979–38007. https://doi.org/10.1007/s11042-023-17007-z

Jones, C. M., & Rogers, L. O. (2022). "There Are Stereotypes for Everything": Multiracial Adolescents Navigating Racial Identity under White Supremacy. *Social Sciences*, *11*(1), 19. https://doi.org/10.3390/socsci11010019

Jones, J. A., & Frank, R. (2023). Making Minorities or Honorary Whites? Examining Multiracial Self-Concept. *Social Psychology Quarterly*, *86*(3), 219–240. https://doi.org/10.1177/01902725231177646

Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2014). Pronoun Use Reflects Standings in Social Hierarchies. *Journal of Language and Social Psychology*, *33*(2), 125–143. https://doi.org/10.1177/0261927X13502654

Kalsner, L., & Pistole, M. C. (2003). College Adjustment in a Multiethnic Sample: Attachment, Separation-Individuation, and Ethnic Identity. *Journal of College Student Development*, *44*(1), 92–109. https://doi.org/10.1353/csd.2003.0006

Kashima, Y., Yamaguchi, S., Kim, U., Choi, S.-C., & Gelfand, M. J. (1995). Culture, Gender, and Self: A Perspective From Individualism-Collectivism Research. *Journal of Personality and Social Psychology*, *69*(5), 925–937. https://doi.org/10.1037//0022-3514.69.5.925

Keesing, R. M. (1974). Theories of Culture. *Annual Review of Anthropology*, *3*, 73–97. JSTOR.

Khanna, N. (2011). Ethnicity and race as 'symbolic': The use of ethnic and racial symbols in asserting a biracial identity. *Ethnic and Racial Studies*, *34*(6), 1049–1067. https://doi.org/10.1080/01419870.2010.538421

Koltcov, S., Ignatenko, V., & Koltsova, O. (2019). Estimating Topic Modeling Performance with Sharma–Mittal Entropy. *Entropy*, *21*(7), 660. https://doi.org/10.3390/e21070660

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kuhn, K. D. (2018). Using structural topic modeling to identify latent topics and trends in aviation incident reports. *Transportation Research Part C: Emerging Technologies*, *87*, 105–122.

Kuhn, M. H., & Mcpartland, T. S. (1954). An Empirical Investigation of Self-Attitudes. *American Sociological Review*, *19*, 68–76.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Light, R., & Odden, C. (2017). Managing the Boundaries of Taste: Culture, Valuation, and Computational Social Science. *Social Forces*, *96*, 877–908.

Lilgendahl, J. P., Benet-Martinez, V., Bishop, M., Gilson, K., Festa, L., Levenson, C., & Rosenblum, R. (2018). "So Now, I Wonder, What Am I?": A Narrative Approach to Bicultural Identity Integration. *Journal of Cross-Cultural Psychology*, *49*(10), 1596–1624. https://doi.org/10.1177/0022022118801555

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. *arXiv Preprint arXiv:1907.11692*.

Loveys, K., Torrez, J., Fine, A., Moriarty, G., & Coppersmith, G. (2018). Cross-cultural differences in language markers of depression online. *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*, 78–87. https://doi.org/10.18653/v1/W18-0608

Loyd, A. B., Westberg, D. W., Williams, L., Humphries, M., Meca, A., & Rodil, J. C. (2023). "I just want to be me, authentically": Identity shifting among racially and ethnically diverse young adults. *Journal of Youth and Adolescence*, *52*(4), 701–718. https://doi.org/10.1007/s10964-023-01744-3

Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin*, *18*(3), 302–318.

Luna, D., Ringberg, T., & Peracchio, L. A. (2008). One Individual, Two Identities: Frame Switching among Biculturals. *Journal of Consumer Research*, *35*(2), 279–293. https://doi.org/10.1086/586914

Luxburg, U. von. (2007). A Tutorial on Spectral Clustering. *Statistics and Computing*, *17*(4), 395–416. https://doi.org/10.1007/s11222-007-9033-z

Makowski, D., Ben-Shachar, M. S., Chen, S. A., & Lüdecke, D. (2019). Indices of effect existence and significance in the Bayesian framework. *Frontiers in Psychology*, *10*, 2767.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press; Cambridge Core. https://doi.org/10.1017/CBO9780511809071

Markus, H. R., & Hamedani, M. G. (2019). People are culturally shaped shapers: The psychological science of culture and culture change. *The Guilford Press*, *Handbook of cultural psychology*.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*(2), 224–253. https://doi.org/10.1037/0033-295X.98.2.224

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, *38*, 299–337. https://doi.org/10.1146/annurev.ps.38.020187.001503

McAdams, D. P. (1995). What do we know when we know a person? *Journal of Personality*, *63*(3), 365–396. https://doi.org/10.1111/j.1467-6494.1995.tb00500.x

McAdams, D. P. (2001). The Psychology of Life Stories. *Review of General Psychology*, *5*(2), 100–122. https://doi.org/10.1037/1089-2680.5.2.100

McAdams, D.P. & McLean, Kate. (2013). Narrative Identity. *Current Directions in Psychological Science. 22.* 233-238. 10.1177/0963721413475622.

McCabe, A., & Dinh, K. T. (2016). Agency and Communion, Ineffectiveness and Alienation: Themes in the Life Stories of Latino and Southeast Asian Adolescents. *Imagination, Cognition and Personality*, *36*(2), 150–171. https://doi.org/10.1177/0276236616648648

McCrae, R. R., & Costa, P. T. (1988). Age, Personality, and the Spontaneous Self-Concept. *Journal of Gerontology*, *43*(6), S177–S185. https://doi.org/10.1093/geronj/43.6.S177

McCutcheon, A. L. (1987). *Latent class analysis* (Vol. 64). Sage.
McLachlan, G., & Peel, D. (2000). *Finite Mixture Models*. John Wiley & Sons, Inc.
McLean, K. C., Lilgendahl, J. P., Fordham, C., Alpert, E., Marsden, E., Szymanowski, K., & McAdams, D. P. (2018). Identity development in cultural context: The role of deviating from master narratives. *Journal of Personality*, *86*(4), 631–651. https://doi.org/10.1111/jopy.12341
McLean, K. C., & Pratt, M. W. (2006). Life's little (and big) lessons: Identity statuses and meaning-making in the turning point narratives of emerging adults. *Developmental Psychology*, *42*(4), 714.
McLean, K. C., & Syed, M. (2015). Personal, Master, and Alternative Narratives: An Integrative Framework for Understanding Identity Development in Context. *Human Development*, *58*(6), 318–349. https://doi.org/10.1159/000445817
McLean, K. C., Syed, M., Pasupathi, M., Adler, J. M., Dunlop, W. L., Drustrup, D., Fivush, R., Graci, M. E., Lilgendahl, J. P., Lodi-Smith, J., McAdams, D. P., & McCoy, T. P. (2020). The empirical structure of narrative identity: The initial Big Three. *Journal of Personality and Social Psychology*, *119*(4), 920–944. https://doi.org/10.1037/pspp0000247
Mehra, A. (2023). At the Margins: A Distinctiveness Approach to the Social Identity and Social Networks of Underrepresented Groups. *Academy of Management Journal*.
Meyers, C., Williams, A., Pauker, K., & Apfelbaum, E. P. (2022). The impact of social norms on navigating race in a racially diverse context. *Group Processes & Intergroup Relations*, *25*(4), 853–870. https://doi.org/10.1177/1368430220984228
Mihoko Doyle, J., & Kao, G. (2007). Are Racial Identities of Multiracials Stable? Changing Self-Identification Among Single and Multiple Race Individuals. *Social Psychology Quarterly*, *70*(4), 405–423. https://doi.org/10.1177/019027250707000409
Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of Word Representations in Vector Space* (arXiv:1301.3781). arXiv. https://doi.org/10.48550/arXiv.1301.3781
Miller, B., Rocks, S., Catalina, S., Zemaitis, N., Daniels, K., & Londono, J. (2019). The missing link in contemporary health disparities research: A profile of the mental and self-rated health of multiracial young adults. *Health Sociology Review*, *28*(2), 209–227. https://doi.org/10.1080/14461242.2019.1607524
Miramontez, D. R., Benet-Martínez, V., & Nguyen, A.-M. D. (2008). Bicultural Identity and Self/Group Personality Perceptions. *Self and Identity*, *7*(4), 430–445. https://doi.org/10.1080/15298860701833119
Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, *102*(2), 246–268. https://doi.org/10.1037/0033-295X.102.2.246
Mitchell, L. L., Kathawalla, U.-K., Ajayi, A. A., Fish, J., Nelson, S. C., Peissig, L. H. M., & Syed, M. (2018). Ethnic-racial typicality and its relation to ethnic identity and psychological functioning. *Cultural Diversity and Ethnic Minority Psychology*, *24*(3), 400–413. https://doi.org/10.1037/cdp0000193
Moffitt, U., Rogers, L. O., Mzizi, Y., & Charlson, E. (2022). Race Talk During the 2020 U.S. Presidential Election: Emerging Adults' Critical Consciousness and Racial

Identity in Context. *Journal of Adolescent Research*, 074355842211450. https://doi.org/10.1177/07435584221145009
Morning, A. (2000). Who is multiracial? Definitions and decisions. *Sociological Imagination*, *37*(4), 209–229.
Morning, A., & Saperstein, A. (2018). The Generational Locus of Multiraciality and Its Implications for Racial Self-Identification. *The ANNALS of the American Academy of Political and Social Science*, *677*(1), 57–68. https://doi.org/10.1177/0002716218754774
Morrison, J. W. (1995). Developing identity formation and self-concept in preschool-aged biracial children. *Early Child Development and Care*, *111*(1), 141–152. https://doi.org/10.1080/0300443951110110
Nadal, K. L., Sriken, J., Davidoff, K. C., Wong, Y., & McLean, K. (2013). Microaggressions Within Families: Experiences of Multiracial People. *Family Relations*, *62*(1), 190–201. JSTOR.
Nagin, D. S. (2005). *Group-based modeling of development*. Harvard university press.
Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender Differences in Language Use: An Analysis of 14,000 Text Samples. *Discourse Processes*, *45*(3), 211–236. https://doi.org/10.1080/01638530802073712
Nicolas, G., Skinner, A. L., & Dickter, C. L. (2019). Other Than the Sum: Hispanic and Middle Eastern Categorizations of Black–White Mixed-Race Faces. *Social Psychological and Personality Science*, *10*(4), 532–541. https://doi.org/10.1177/1948550618769591
Noriega, E. (2016). *Exploring the "Chameleon Experience" Through the Lens of Multiracial People*. University of California, Santa Barbara. https://books.google.com/books?id=gGdT0AEACAAJ
Norman, J. B., & Chen, J. M. (2020). I am Multiracial: Predictors of Multiracial identification strength among mixed ancestry individuals. *Self and Identity*, *19*(5), 501–520. https://doi.org/10.1080/15298868.2019.1635522
Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(4), 535–569.
Omi, M., & Winant, H. (1986). *Racial formation in the United States: From the 1960s to the 1980s*. Routledge & Kegan Paul.
Ortiz, C. (2017). An Exploration of Parental Racial Socialization in Dual-Minority Multiracial Families [Ph.D., The University of Chicago]. In *ProQuest Dissertations and Theses* (1928531002). ProQuest Central; ProQuest Dissertations & Theses Global; Sociological Abstracts. https://ezproxy.cul.columbia.edu/login?url=https://www.proquest.com/dissertations-theses/exploration-parental-racial-socialization-dual/docview/1928531002/se-2?accountid=10226
Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, *128*(1), 3–72. https://doi.org/10.1037/0033-2909.128.1.3

Oyserman, D., & Destin, M. (2010). Identity-Based Motivation: Implications for Intervention. *The Counseling Psychologist*, *38*(7), 1001–1043. https://doi.org/10.1177/0011000010374775

Oyserman, D., Elmore, K., & Smith, G. (2012). Self, self-concept, and identity. *Handbook of Self and Identity, 2nd Ed.*, 69–104.

Oyserman, D., & Markus, H. R. (1990). Possible selves and delinquency. *Journal of Personality and Social Psychology*, *59*(1), 112–125. https://doi.org/10.1037/0022-3514.59.1.112

Pauker, K., Carpinella, C. M., Lick, David. J., Sanchez, D. T., & Johnson, K. L. (2018). Malleability in Biracial Categorizations: The Impact of Geographic Context and Targets' Racial Heritage. *Social Cognition*, *36*(5), 461–480. https://doi.org/10.1521/soco.2018.36.5.461

Peng, K., & Nisbett, R. E. (1999). Culture, Dialectics, and Reasoning About Contradiction. *American Psychologist*.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological Aspects of Natural Language Use: Our Words, Our Selves. *Annual Review of Psychology*, *54*(1), 547–577. https://doi.org/10.1146/annurev.psych.54.101601.145041

Pennington, J., Socher, R., & Manning, C. D. (2014). *Glove: Global vectors for word representation*. 1532–1543.

Perone, C., Silveira, R., & Paula, T. (2018). *Evaluation of sentence embeddings in downstream and linguistic probing tasks*. https://doi.org/10.48550/arXiv.1806.06259

Peters, M., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). *Deep contextualized word representations*. https://doi.org/10.48550/arXiv.1802.05365

Phinney, J. S. (1990). Ethnic identity in adolescents and adults: Review of research. *Psychological Bulletin*, *108*(3), 499–514. https://doi.org/10.1037/0033-2909.108.3.499

Phinney, J. S. (1992). The Multigroup Ethnic Identity Measure. *Journal of Adolescent Research*, *7*(2), 156–176.

Phinney, J. S. (1999). An intercultural approach in psychology: Cultural contact and identity. *Cross-Cultural Psychology Bulletin*, *33*, 24–31.

Poston, W. S. C. (1990). The Biracial Identity Development Model: A Needed Addition. *Journal of Counseling & Development*, *69*(2), 152–155. https://doi.org/10.1002/j.1556-6676.1990.tb01477.x

Prado, A. M. del, Church, A. T., Katigbak, M. S., Miramontes, L. G., Whitty, M. T., Curtis, G. J., Vargas-Flores, J. de J., Ibáñez-Reyes, J., Ortiz, F. A., & Reyes, J. A. S. (2007). Culture, method, and the content of self-concepts: Testing trait, individual–self-primacy, and cultural psychology perspectives. *Journal of Research in Personality*, *41*(6), 1119–1160. https://doi.org/10.1016/j.jrp.2007.02.002

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI Blog*, *1*(8), 9.

Ramirez, E. H., Brena, R., Magatti, D., & Stella, F. (2012). Topic model validation. *Neurocomputing*, *76*(1), 125–133. https://doi.org/10.1016/j.neucom.2011.04.032

Rathje, S., Mirea, D.-M., Sucholutsky, I., Marjieh, R., Robertson, C. E., & Van Bavel, J. J. (2024). GPT is an effective tool for multilingual psychological text analysis.

Proceedings of the National Academy of Sciences, *121*(34), e2308950121. https://doi.org/10.1073/pnas.2308950121

Remedios, J. D., & Chasteen, A. L. (2013). Finally, someone who "gets" me! Multiracial people value others' accuracy about their race. *Cultural Diversity & Ethnic Minority Psychology*, *19*(4), 453–460. https://doi.org/10.1037/a0032249

Renn, K. A. (2004). *Mixed race students in college: The ecology of race, identity, and community on campus*. Suny Press.

Renn, K. A. (2008). Research on biracial and multiracial identity development: Overview and synthesis. *New Directions for Student Services*, *2008*(123), 13–21. https://doi.org/10.1002/ss.282

Revelle, W., & Condon, D. M. (2019). Reliability from α to ω: A tutorial. *Psychological assessment*, *31*(12), 1395–1411. https://doi.org/10.1037/pas0000754

Rice, D. R., & Zorn, C. (2021). Corpus-based dictionaries for sentiment analysis of specialized vocabularies. *Political Science Research and Methods*, *9*(1), 20–35. Cambridge Core. https://doi.org/10.1017/psrm.2019.10

Richards, C., Bouman, W. P., Seal, L., Barker, M. J., Nieder, T. O., & T'Sjoen, G. (2016). Non-binary or genderqueer genders. *International Review of Psychiatry*, *28*(1), 95–102. https://doi.org/10.3109/09540261.2015.1106446

Richeson, J. A., & Sommers, S. R. (2016). Toward a Social Psychology of Race and Race Relations for the Twenty-First Century. *Annual Review of Psychology*, *67*(1), 439–463. https://doi.org/10.1146/annurev-psych-010213-115115

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). stm: R Package for Structural Topic Models. *Journal of Statistical Software*.

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural Topic Models for Open-Ended Survey Responses. *American Journal of Political Science*, *58*(4), 1064–1082. https://doi.org/10.1111/ajps.12103

Roberts, M. E., Tingley, D., Stewart, B. M., & Airoldi, E. M. (2013). *The Structural Topic Model and Applied Social Science*.

Robins, R. W., Tracy, J. L., & Trzesniewski, K. H. (2008). Naturalizing the self. In *Handbook of personality: Theory and research, 3rd ed.* (pp. 421–447). The Guilford Press.

Roccas, S., & Brewer, M. B. (2002). Social Identity Complexity. *Personality and Social Psychology Review*, *6*(2), 88–106. https://doi.org/10.1207/S15327957PSPR0602_01

Rockquemore, K. A., & Brunsma, D. L. (2002). Socially embedded identities: Theories, typologies, and processes of raical identity among Black/White biracials. *The Sociological Quarterly*, *43*(3), 335–356. https://doi.org/10.1111/j.1533-8525.2002.tb00052.x

Rockquemore, K. A., Brunsma, D. L., & Delgado, D. J. (2009). Racing to Theory or Retheorizing Race? Understanding the Struggle to Build a Multiracial Identity Theory. *Journal of Social Issues*, *65*(1), 13–34. https://doi.org/10.1111/j.1540-4560.2008.01585.x

Rogers, L. O., Niwa, E. Y., Chung, K., Yip, T., & Chae, D. (2021). M(ai)cro: Centering the Macrosystem in Human Development. *Human Development*, *65*(5–6), 270–292. https://doi.org/10.1159/000519630

Root, M. (1996). *The Multiracial Experience: Racial Borders as the New Frontier*. SAGE Publications, Inc. https://doi.org/10.4135/9781483327433

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Sanchez, D. T., Shih, M., & Garcia, J. A. (2009). Juggling multiple racial identities: Malleable racial identification and psychological well-being. *Cultural Diversity & Ethnic Minority Psychology*, *15*(3), 243–254. https://doi.org/10.1037/a0014373

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity*, *52*(4), 1893–1907. https://doi.org/10.1007/s11135-017-0574-8

Schaaff, K., Reinig, C., & Schlippe, T. (2023). *Exploring ChatGPT's empathic abilities*. 1–8.

Schwartz, S. J., Luyckx, K., & Vignoles, V. L. (2011). *Handbook of identity theory and research*. Springer.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 461–464.

Sclove, S. L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*, *52*(3), 333–343.

Sellers, R. M., Smith, M. A., Shelton, J. N., Rowley, S. A. J., & Chavous, T. M. (1998). Multidimensional Model of Racial Identity: A Reconceptualization of African American Racial Identity. *Personality and Social Psychology Review*, *2*(1), 18–39. https://doi.org/10.1207/s15327957pspr0201_2

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shih, M., & Sanchez, D. T. (2005). Perspectives and Research on the Positive and Negative Implications of Having Multiple Racial Identities. *Psychological Bulletin*, *131*(4), 569–591. https://doi.org/10.1037/0033-2909.131.4.569

Shih, M., & Sanchez, D. T. (2009). When Race Becomes Even More Complex: Toward Understanding the Landscape of Multiracial Identity and Experiences. *Journal of Social Issues*, *65*(1), 1–11. https://doi.org/10.1111/j.1540-4560.2008.01584.x

Singer, J. A. (2004). Narrative Identity and Meaning Making Across the Adult Lifespan: An Introduction. *Journal of Personality*, *72*(3), 437–459. https://doi.org/10.1111/j.0022-3506.2004.00268.x

Smith, J. A., Jarman, M., & Osborn, M. (1999). Doing interpretative phenomenological analysis. *Qualitative Health Psychology: Theories and Methods*, *1*(1), 218–240.

Smith, J. A., & Osborn, M. (2003). Interpretative phenomenological analysis. In *Qualitative psychology: A practical guide to research methods.* (pp. 51–80). Sage Publications, Inc.

Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A., & Potts, C. (2013). Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank. In D. Yarowsky, T. Baldwin, A. Korhonen, K. Livescu, & S. Bethard (Eds.), *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing* (pp. 1631–1642). Association for Computational Linguistics. https://aclanthology.org/D13-1170/

Song, M. (2016). Multiracial people and their partners in Britain: Extending the link between intermarriage and integration? *Ethnicities*, *16*(4), 631–648. https://doi.org/10.1177/1468796816638399

Song, M. (2025). The Double-Edged Nature of Whiteness for Multiracial People with White Ancestry in the US and UK. *Genealogy*, *9*(2). https://doi.org/10.3390/genealogy9020046

Spencer-Rodgers, J., Boucher, H. C., Mori, S. C., Lei Wang, & Kaiping Peng. (2009). The Dialectical Self-Concept: Contradiction, Change, and Holism in East Asian Cultures. *Personality and Social Psychology Bulletin*, *35*(1), 29–44. https://doi.org/10.1177/0146167208325772

Spencer-Rodgers, J., & Peng, K. (2005). The Dialectical Self: Contradiction, Change, and Holism in the East Asian Self-Concept. In *Cultural and social behavior: The Ontario Symposium, Vol 10.* (pp. 227–249). Lawrence Erlbaum Associates Publishers.

Starr, P., & Pao, C. (2024). The Multiracial Complication: The 2020 Census and the Fictitious Multiracial Boom. *Sociological Science*, *11*, 1107–1123. https://doi.org/10.15195/v11.a40

Stolcke, A., Shriberg, E., Bates, R. A., Ostendorf, M., Hakkani, D. Z., Plauche, M., Tür, G., & Lu, Y. (1998). *Automatic detection of sentence boundaries and disfluencies based on recognized words. 2*, 2247–2250.

Stonequist, E. V. (1935). The Problem of the Marginal Man. *American Journal of Sociology*, *41*(1), 1–12. JSTOR.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory, 2nd ed.* (pp. xiii, 312). Sage Publications, Inc.

Svensson, Y., & Syed, M. (2019). Linking self and society: Identity and the immigrant experience in two macro-contexts. *Journal of Applied Developmental Psychology*, *64*, 101056. https://doi.org/10.1016/j.appdev.2019.101056

Swann, W. B. (1987). Identity negotiation: Where two roads meet. *Journal of Personality and Social Psychology*, *53*(6), 1038–1051. https://doi.org/10.1037/0022-3514.53.6.1038

Swidler, A. (1986). Culture in Action: Symbols and Strategies. *American Sociological Review*, *51*(2), 273. https://doi.org/10.2307/2095521

Syed, M. (2015). Theoretical and methodological contributions of narrative psychology to ethnic identity research. In *Studying ethnic identity: Methodological and conceptual approaches across disciplines.* (pp. 27–54). American Psychological Association. https://doi.org/10.1037/14618-003

Syed, M., & Azmitia, M. (2008). A narrative approach to ethnic identity in emerging adulthood: Bringing life to the identity status model. *Developmental Psychology*, *44*(4), 1012–1027. https://doi.org/10.1037/0012-1649.44.4.1012

Syed, M., & Azmitia, M. (2010). Narrative and ethnic identity exploration: A longitudinal account of emerging adults' ethnicity-related experiences. *Developmental Psychology*, *46*(1), 208–219. https://doi.org/10.1037/a0017825

Syed, M., & Juang, L. P. (2014). Ethnic identity, identity coherence, and psychological functioning: Testing basic assumptions of the developmental model. *Cultural*

Diversity and Ethnic Minority Psychology, *20*(2), 176–190. https://doi.org/10.1037/a0035330

Syed, M., & McLean, K. C. (2016). Understanding identity integration: Theoretical, methodological, and applied issues. *Journal of Adolescence*, *47*, 109–118.

Syed, M., & McLean, K. C. (2021). *Who Gets to Live the Good Life? Master Narratives, Identity, and Well-Being within a Marginalizing Society* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/4pjuf

Tajfel, H., & Turner, J. C. (1979). An integrative theory of inter-group conflict. In *The social psychology of inter-group relations* (W.G. Austin&S. Worchel, pp. 33–47). Brooks/Cole.

Tamm, A., Tõugu, P., & Tulviste, T. (2024). Self-concept at different stages of life: How do early and late adolescents and young, middle-aged, and older adults describe themselves? *International Journal of Behavioral Development*, *48*(5), 434–441. https://doi.org/10.1177/01650254241254103

Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, *29*(1), 24–54. https://doi.org/10.1177/0261927X09351676

Thompson, B., Roberts, S. G., & Lupyan, G. (2020). Cultural influences on word meanings revealed through large-scale semantic alignment. *Nature Human Behaviour*, *4*(10), 1029–1038. https://doi.org/10.1038/s41562-020-0924-8

Törngren, S. O. (2018). Ethnic Options, Covering and Passing. *Asian Journal of Social Science*, *46*(6), 748–773. https://doi.org/10.1163/15685314-04606007

Törngren, S. O., Irastorza, N., & Rodríguez-García, D. (2021). Understanding multiethnic and multiracial experiences globally: Towards a conceptual framework of mixedness. *Journal of Ethnic and Migration Studies*, *47*(4), 763–781. https://doi.org/10.1080/1369183X.2019.1654150

Törngren, S. O., & Sato, Y. (2021). Beyond being either-or: Identification of multiracial and multiethnic Japanese. *Journal of Ethnic and Migration Studies*, *47*(4), 802–820. https://doi.org/10.1080/1369183X.2019.1654155

Townsend, S. S. M., Markus, H. R., & Bergsieker, H. B. (2009). My Choice, Your Categories: The Denial of Multiracial Identities. *Journal of Social Issues*, *65*(1), 185–204. https://doi.org/10.1111/j.1540-4560.2008.01594.x

Tran, A., Miyake, E., Martinez-Morales, V., & Csizmadia, A. (2015). "What Are You?" Multiracial Individuals' Responses to Racial Identification Inquiries. *Cultural Diversity & Ethnic Minority Psychology*, *22*, 26–37. https://doi.org/10.1037/cdp0000031

Tsai, A., Straka, B., & Kimura, D. (2022). Mixed-Heritage Individuals and Systemic Risk During Negotiation of Identity. *Journal of Language, Identity & Education*, *23*(6), 895–909. https://doi.org/10.1080/15348458.2022.2065277

Tsai, A., Straka, B., & Kimura, D. (2024). Mixed-Heritage Individuals and Systemic Risk During Negotiation of Identity. *Journal of Language, Identity & Education*, *23*(6), 895–909. https://doi.org/10.1080/15348458.2022.2065277

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory.* (pp. x, 239). Basil Blackwell.

Turner, V. (1969). *The Ritual Process. Structure and Anti-Structure*. Brunswick & London: Aldine Transaction.

Umaña-Taylor, A. J., Quintana, S. M., Lee, R. M., Cross, W. E., Rivas-Drake, D., Schwartz, S. J., Syed, M., Yip, T., Seaton, E., & Ethnic and Racial Identity in the 21st Century Study Group. (2014). Ethnic and Racial Identity During Adolescence and Into Young Adulthood: An Integrated Conceptualization. *Child Development*, *85*(1), 21–39. https://doi.org/10.1111/cdev.12196

van Gennep, A. (1960). *The Rites of Passage*. University of Chicago Press. https://books.google.com/books?id=LuEOWytNUhQC

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P.-C. (2021). Rank-normalization, folding, and localization: An improved Rˆ for assessing convergence of MCMC (with discussion). *Bayesian Analysis*, *16*(2), 667–718.

Villegas-Gold, R., & Tran, A. G. T. T. (2018). Socialization and well-being in multiracial individuals: A moderated mediation model of racial ambiguity and identity. *Journal of Counseling Psychology*, *65*(4), 413–422. https://doi.org/10.1037/cou0000277

Wallach, H. M. (2006). Topic modeling: Beyond bag-of-words. *Proceedings of the 23rd International Conference on Machine Learning - ICML '06*, 977–984. https://doi.org/10.1145/1143844.1143967

Ward, C., Ng Tseung-Wong, C., Szabo, A., Qumseya, T., & Bhowon, U. (2018). Hybrid and Alternating Identity Styles as Strategies for Managing Multicultural Identities. *Journal of Cross-Cultural Psychology*, *49*(9), 1402–1439. https://doi.org/10.1177/0022022118782641

Webber, M., & Waru-Benson, S. (2022). The role of cultural connectedness and ethnic group belonging to the social-emotional wellbeing of diverse students. In T. A. Olsen & H. Sollid (Eds.), *Indigenising Education and Citizenship* (pp. 294–312). Scandinavian University Press. https://doi.org/10.18261/9788215053417-2022-16

Westberg, D. W. (2022). Stories off the beaten path: Exploring the content and correlates of cultural deviation narratives. *Journal of Research in Personality*, *100*, 104266. https://doi.org/10.1016/j.jrp.2022.104266

Weston, S. J., Shryock, I., Light, R., & Fisher, P. A. (2023). Selecting the Number and Labels of Topics in Topic Modeling: A Tutorial. *Advances in Methods and Practices in Psychological Science*, *6*(2), 251524592311601. https://doi.org/10.1177/25152459231160105

Wetherell, M., & Potter, J. (1992). *Mapping the language of racism: Discourse and the legitimation of exploitation.* (pp. ix, 246). Columbia University Press.

Wiberg, M., Molenaar, D., González, J., Böckenholt, U., & Kim, J.-S. (Eds.). (2021). *Quantitative Psychology: The 85th Annual Meeting of the Psychometric Society, Virtual* (Vol. 353). Springer International Publishing. https://doi.org/10.1007/978-3-030-74772-5

Williams, C. D., Byrd, C. M., Quintana, S. M., Anicama, C., Kiang, L., Umaña-Taylor, A. J., Calzada, E. J., Pabón Gautier, M., Ejesi, K., Tuitt, N. R., Martinez-Fuentes, S., White, L., Marks, A., Rogers, L. O., & Whitesell, N. (2020). A Lifespan Model of Ethnic-Racial Identity. *Research in Human Development*, *17*(2–3), 99–129. https://doi.org/10.1080/15427609.2020.1831882

Willig, C. (2003). Discourse analysis. *Qualitative Psychology: A Practical Guide to Research Methods*, 159–183.

Wilson Van Voorhis, C. R., & Morgan, B. L. (2007). Understanding Power and Rules of Thumb for Determining Sample Sizes. *Tutorials in Quantitative Methods for Psychology*, *3*(2), 43–50. https://doi.org/10.20982/tqmp.03.2.p043

Xu, J., Saperstein, A., Morning, A., & Iverson, S. (2021). Gender, Generation, and Multiracial Identification in the United States. *Demography*, *58*(5), 1603–1630. https://doi.org/10.1215/00703370-9334366

Yip, T., Cheah, C. S. L., Kiang, L., & Hall, G. C. N. (2021). Rendered invisible: Are Asian Americans a model or a marginalized minority? *American Psychologist*, *76*(4), 575–581. https://doi.org/10.1037/amp0000857