NORTHWESTERN UNIVERSITY

Causal Heterogeneity in Social Essentialism: Shared Experiences and Shared Genes

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Psychology

By

Natalie M. Gallagher

EVANSTON, ILLINOIS

September 2021

Abstract

We structure our lives around social groups – belonging to them and thinking about them. In this dissertation, I develop a new stereotype content measure to assess the attributes associated with groups in America today, propose and support a theory of sociocultural essentialism, and explore the strategic activation of sociocultural essentialism among members of marginalized groups. Together, these studies contribute to psychological research on the relationship between ontological belief and stereotyping, and their functions in society today.

In the first chapter, I comprehensively update the adjective checklist, an existing measure to assess stereotype content about different groups in American society today. I first discuss the affordances and critiques of the measure and conduct a narrative review of its history, before generating a new measure integrating content from in-lab and online text sources. I then gather and use norming data to pick the best set of terms for inclusion in the measure and validate the resulting list.

In the second chapter, I use the new adjective checklist measure to test the theory of sociocultural essentialism. I argue that existing research on essentialist beliefs about social categories have over-emphasized biogenetic beliefs and that the perception of shared social experience or common fate may similarly ground essentialist perceptions. I update an existing experimental task to test this proposition across six social domains, asking participants to attribute traits to an individual who had switched between two social groups in a given domain. Findings support the existence of sociocultural essentialist reasoning while also showing that participants' ideas about the process of switching groups impacted attributions beyond the stereotype of each group.

In the third chapter, I explore the role of sociocultural essentialism among members of marginalized groups. With LGB+ participants, I test whether reminders of group devaluation and group denial change levels of sociocultural and biogenetic essentialist thinking and whether changes in essentialist beliefs predict ingroup attitudes or support for policy change. Though the experimental manipulation failed to produce the expected differences in essentialist thinking, I discuss correlational and exploratory results suggesting that biogenetic and sociocultural essentialism play distinct roles in this population, and that reminders of group denial produce sociocultural essentialism among younger participants, while reminders of group devaluation do so among older participants. Finally, I review the three chapters and suggest future avenues for development of the research program.

Acknowledgements

To my advisors, Doug Medin and Galen Bodenhausen, thank you for all that you've taught me, including all I didn't realize I needed to learn. Thank you especially for always reminding me to ask not just *if* but *why* and *for whom* – it has made me a better researcher and a better person. Thank you also to Leoandra Onnie Rogers, a member of my committee whose perspective and feedback within and beyond this project have been invaluable.

The MOSAIC and Social Cognition labs have been wonderful environments in which to grow during graduate school. I thank Jennie Woodring, Alissa Baker, Jordan Daley, Janene Cielto, Amanda Herrera, bethany ojalehto, and Kalonji Nzinga for this. Special thanks to Alexandra Garr-Schultz, Chris Petsko, and Francisco Maravilla for guiding me through the program, and Kyle Nolla and Nathan Couch for countless conversations on statistics, methods, and measurement.

To my dad, Steve, and sister, Halley – our trio grounds everything for me; I couldn't have done it without you. To my partner, Aistė, thank you for believing in me always, and in this project when I doubted.

This work would not have been possible without the generous support of the National Science Foundation, Northwestern's Graduate Research Grant program, and the Dispute Resolution Research Center at Kellogg. Thanks also to the Summer Institute in Computational Social Science and the Kellogg Linux Cluster at Northwestern University for training and support. Finally, my sincere gratitude to every coffee shop in Chicago that harbored me during the writing of this dissertation.

Table of Contents

List of Figures

List of Tables

Chapter I: Developing an Updated Adjective Checklist Measure

Our use of social categories brings order to a wildly variable interpersonal environment. Shaped by our culture, our context, and our social learning processes, it provides structure for how we understand ourselves as group members, and for our interactions with others. In this chapter, I address the importance of measuring stereotype content, discuss the affordances and critiques of the adjective checklist measure of stereotyping, and conduct a narrative review of its use from 1933 until today. Finally, I use a multi-method approach to develop an updated adjective checklist suitable for use across an array of social domains in the US cultural context.

Stereotype Content

Stereotypes contain both affective content (i.e., general positive or negative attitudes towards a group) and descriptive content (i.e., traits, attributes, or behaviors associated with a group). Just as the boundary lines defining particular social groups vary across societies, the link between a group and the stereotype applied to it is contingent and contextual. Stereotypes emerge from systems of sociohistorical power (e.g., white supremacy, patriarchy, colonialism), and should be understood with reference to those systems (e.g., Auguostinos & Walker, 1998; Buchanan, 1951; Embrick & Henricks, 2013; Jost & Banaji, 1994). People in higher-power social positions become defined as the norm, with other groups stereotyped in reference to – as verbally marked deviations from – this group (Bailey et al., 2020; Brekhus, 1998). This can also be seen in the content of verbally marked groups: beyond the marked membership, the stereotyped group is assumed to be normative (e.g., stereotypes about nationalities in general overlap more with stereotypes about men in a country than women in a country; Eagly & Kite, 1987). It is unlikely all members of a society *agree* with hegemonic stereotypes, and people may hold their own personal beliefs about a particular group even while being aware of cultural

stereotypes (Devine & Elliot, 1995). Stereotype awareness can provoke resistance to negative stereotypes (e.g., Perry et al., 2019; Rogers & Way, 2018), but unchallenged hegemonic stereotypes help reproduce existing systems of power (e.g., Steele, 1997, 2010).

The descriptive content of stereotypes can predict the qualitative experiences and treatment of that group in society (Bodenhausen et al., 1998). For example, Asian, Latinx, and Native American people in the US experience discrimination on the basis of being perceived as "culturally foreign", while Black people experience discrimination on the basis of being perceived as "inferior" (Nosek et al., 2007; Zou & Cheryan, 2017). Stereotype content is flexible to changes in society (e.g., the increasing presence of women in the workforce is associated with more agentic stereotypes of women; Eagly et al., 2020; Jones et al., 2020), but also to the ways inequity persists over time (e.g., a history of slavery predicts pro-White bias among White Americans, partly through indicators of contemporary structural racism; Payne et al., 2019).

Studying the stereotype content applied to groups within American society has been an extremely fruitful avenue for understanding disparate treatment based on group membership (e.g., Bodenhausen & Richeson, 2010; Cuddy et al., 2007). It has provided a framework through which to understand why Americans provide racially-biased medical treatment (Staton et al., 2007; van Ryn et al., 2011), feel pressure to behave in gender-normative ways (Way et al., 2014), socially exclude elderly people (Cuddy et al., 2005; Vale et al., 2020), and support military campaigns in majority-Muslim nations (Sides & Gross, 2013).

Measuring Stereotype Content with the Adjective Checklist

One frequently-used measure of stereotype content is the adjective checklist (Katz & Braly, 1933, 1935). In this measure, participants see a list of human characteristics (e.g., *proud*, *happy*, *yielding*) and choose which of them they associate with a certain group or individual

group member (e.g., gay Black men; Petsko & Bodenhausen, 2019). If the participant has chosen more than five traits to describe a particular group, they may be asked to choose the five traits that are most typical of the group. In some cases, participants are also able to add adjectives that they deem missing from the overall list. This data can be used to assess stereotype consensus – the extent to which multiple participants from a given cultural setting choose the same traits to describe a certain group – to qualitatively describe the stereotypes applied to each group, and to test whether different groups are more or less associated with particular attributes (e.g., Gilbert, 1951; Karlins et al., 1969; Katz & Braly, 1935).

The approach can offer further insight by combining responses with data from norming participants. Typically, a set of norming participants rates all of the characteristics on the extent to which they are connected to a unifying construct (e.g., stereotypic Blackness; Petsko & Bodenhausen, 2019). The traits chosen by the study participants can then be assessed in terms of that unifying construct, for example, the extent to which the traits participants chose to describe gay Black men are (or are not) stereotypically Black. Though gathering the norming ratings can be resource-intensive, these ratings can then be used in combination with the trait list for multiple projects and between research teams (e.g., Petsko & Bodenhausen, 2019; Wages et al., 2020). In the next two sections, I discuss the affordances and critiques of the adjective checklist approach to measuring stereotype content, before proceeding to a narrative review of the measure's use since 1933.

Affordances of the Adjective Checklist

The adjective checklist measure of assessing stereotype content balances the holistic assessment of stereotype content with the efficiency and manageability often desired by

researchers. It has three particular benefits: a relatively high level of participant agency, high analytic flexibility for the researcher, and high efficiency for the participant and the researcher.

Participant agency. In some stereotype content measures, participants have low agency. For example, the gender-science implicit association test quantifies the participant's relative association between women, men, art, and science – there is no opportunity for the participant to offer a more holistic impression of their view of the group (Nosek et al., 2009). In others, participants have high agency. For example, some researchers have participants freely generate a description of a group (Niemann et al., 1994), draw a group member (Miller et al., 2018), or use the reverse correlation technique to generate a mental image of a face (Dotsch & Todorov, 2012). In general, high-agency measures require substantially more work by the researchers after gathering the initial data – the freely generated stereotypes must be categorized, coded, or rated prior to comparative analysis. Low-agency measures require substantially less work for the researcher, because they are oriented towards a specific operationalization of the stereotype. The adjective checklist is at an intermediate level of participant agency: it allows a wide array of responses – a list of 84 terms offers over a septillion (10^{24}) possible unique combinations; if participants can add terms, this expands even further – while also putting guardrails around the process to help researchers in the analysis process.

Analytic flexibility. When analyzing the output of an adjective checklist, the researcher can operationalize individual-level stereotypes at multiple levels of complexity. At the simplest level, a researcher can consider the association of a group with an individual trait (e.g., *materialistic* or *practical*, S. A. Haslam & Turner, 1995), without using any norming data. Traits can also be combined together to form dimensions, using norming data about how a given trait is associated with a particular dimension (e.g., stereotypical Blackness, Petsko & Bodenhausen,

2019). Finally, a researcher could focus on the total set of traits an individual participant associated with a particular group, emphasizing that those are tied together by an underlying idea about the group (Brown & Turner, 2002; Kunda et al., 1997; Saenger & Flowerman, 1954; Weitz & Gordon, 1993). For example, *outspoken and creative* suggests a very different kind of person than *outspoken and rude* or even *creative and rude* (Asch & Zukier, 1984; Zanna & Hamilton, 1977).

In contrast, indirect measures like the IAT (Greenwald et al., 2003) or the semantic misattribution procedure (Ye & Gawronski, 2018) define the association between a given target group (or groups) and one dimension or trait. Similarly, shorter scales designed to assess the association between a group and a specific dimension, as in the Stereotype Content Model (Fiske et al., 2007) do not incorporate other relevant traits or the individual participant's holistic stereotype. By allowing analysis at multiple levels of interest, the adjective checklist measure enables researchers to use convergent approaches to answer their questions.

Participant ease of use. The adjective checklist is extremely straightforward for participants to complete. By relying on a binary choice format, it is simpler than: (a) asking participants to think in frequencies as in the diagnostic ratio (McCauley & Stitt, 1978), (b) having participants rate each included trait on unipolar (e.g., Abele et al., 2016) or bipolar scales (e.g., semantic differential; Gardner et al., 1972), or (c) using indirect measures which require many trials (Dotsch & Todorov, 2012; Greenwald et al., 2003; Ye & Gawronski, 2018). This ease of use is a particular benefit because the results of adjective checklist methods are very similar to other measures of explicit stereotype assessment (McCauley & Stitt, 1978; Stangor & Lange, 1994). The simplicity of the measure reflects the ethical need for a researcher to reduce burden on participants while also making it likely that the data will be of better quality and

enabling the repeated use of the measure (e.g., to assess multiple target groups, to assess the same stereotype over time, or to assess the same stereotype in different contexts; Eisele et al., 2020; Sharp & Frankel, 1983).

Efficient use of researcher resources. As with any psychological measure that can be used across projects, repeated use of the same adjective checklist across different studies has benefits. In contrast with fully open-ended measures of stereotype content (e.g., Niemann et al., 1994), which require norming data be gathered for each study, the adjective checklist facilitates re-use of the same norming data between projects (e.g., Petsko & Bodenhausen, 2019; Wages et al., 2020). Measure re-use also facilitates the comparison of stereotype content across time (e.g., the Princeton Trilogy studies; Devine & Elliot, 1995; Gilbert, 1951; Karlins et al., 1969; Katz & Braly, 1933), subgroup (Bayton & Byoune, 1947; Katz & Braly, 1933), or context.

Critiques of the Adjective Checklist

Despite these affordances, the adjective checklist faces four substantive areas of methodological critique. For each, I introduce the critique from the literature, discuss its merits, and propose methodological solutions to the concern as appropriate.

Interpreting Chosen Traits. When somebody chooses an adjective to describe a group, what does that mean? Some scholars have criticized this measure as ambiguous because it does not define a comparison group (Blake & Dennis, 1943; Duijker & Frijda, 1960), or because it does not require that the trait apply to a majority of the target group (Brigham, 1969, 1971). Other approaches, including the diagnostic ratio or percentage estimates (Ganong & Coleman, 1995; Hopkins & Murdoch, 1999; McCauley et al., 1980), avoid this critique by specifying the comparison group and the asking participants to estimate the percentage of group members who have a trait. However, if the specified comparison group is not the group against which the

participant regularly (implicitly) evaluates people, its inclusion may distort participant responses. By not specifying a comparison group, the adjective checklist allows participants to use the default comparison group they would outside of the experimental context. In general, traits chosen in the checklist task contain those which the participant considers distinctive to the group (relative to other groups) and those common to the group (a high percentage of all group members are estimated to have them; McCauley & Stitt, 1978). If there is a particular comparison group of interest, gathering information about both comparison groups is useful; beyond that, this concern does not merit the costs of using more complex measures.

Individual-level stereotypes. Various scholars have claimed that the adjective checklist is unsuitable for assessing stereotypes held by particular individuals (Correll et al., 2010; Six & Eckes, 1991), with McCauley, Stitt, and Segal going as far as to say "there is no way, using the checklist method, to measure personal stereotypes, since the choice of five unusual traits may be only a random selection in the absence of any stereotype" (1980, p. 197). However, random responses are a concern for all explicit measurement of stereotype content, including the diagnostic ratio advocated by McCauley and colleagues, and can be statistically assessed when considering adjective checklist responses (i.e., as in Prothro & Melikian, 1955). A more worrying part of this critique regards the arbitrary requirement for participants to pick exactly five traits to characterize the target group. Some work shows that there is greater between-participant consensus about top-five traits than all traits chosen (Schoenfeld, 1942), while other authors mention that there is little difference in comparing results between adjectives chosen in phase 1 and phase 2 (Meenes, 1943; Petsko & Bodenhausen, 2019; Schneider & Bos, 2011; Stopar, 2015). The two-step procedure was begun by Katz & Braly (1933), but there is no particular defense of the practice in that article, and it seems to have been continued more as a

matter of precedent than theory. Other adjective checklists do not impose this constraint, notably the 300-adjective checklist measure of Gough & Heilbrun (1983) (initially developed as a measure of personality, this and modifications of it have also been used to measure stereotype content, e.g., Hassell & Smith, 1975; O'Connor & D'Angelo, 2013). Removing this constraint on participants is a good option, allowing for maximal self-expression within the boundaries of the task and allowing the number of adjectives chosen to indicate the richness or paucity of an individual participant's stereotype about the target group.

List Construction & Comprehensiveness. The utility of the adjective checklist measure lies squarely in the appropriateness of the adjectives in the list. A list of terms totally unfamiliar to participants, for example, is unlikely to be informative about their stereotypes of any groups at all. The general point of concern here is coverage – whether the terms in the list are comprehensive enough that participants can express their holistic stereotype of the group (Ganong & Coleman, 1995). This can be broken down into the familiarity of the terms to current participants, and the extent to which participants can fully characterize the target group(s) with terms from the list. Scholars looking at stereotype change across time have noted that using outdated terms may reduce the appearance of stereotype consensus among a population (e.g., Devine & Elliot, 1995; Madon et al., 2001). Some scholars have intentionally modernized prior lists (e.g., Oulmokhtar et al., 2011), while others have updated existing checklist measures to incorporate language appropriate to measuring stereotypes of their target social dimensions (e.g., Horch, 2011; Weitz & Gordon, 1993). One solution to concerns about coverage is to allow participants to add adjectives to the list that they consider lacking. By looking at the percentage of participants in a sample that opted to modify the list, the researcher can get a sense of whether

the list was suitable to measure stereotypes about the target group within the research population; if many participants opted to add in terms, the list probably has inadequate coverage.

Inflating Stereotyping Tendencies. The adjective checklist asks participants to make generalizations about the target group, which may encourage stereotypic thinking or responding, overestimating the strength of stereotypes that participants hold (Brigham, 1971; Buchanan, 1951; Ehrlich & Rinehart, 1965b). This is a measurement issue common to all stereotyping research – in order to assess stereotypes about a group, the researcher in some way reifies the group. This may be done by using photos of multiple group members (e.g., the semantic misattribution procedure or IAT), naming the group directly (e.g., scale measures, having participants describe groups), or even having participants interact with a confederate of a given group membership. While the adjective checklist may inflate apparent stereotyping, it does not seem any more likely to do so than other measures in this area. Moreover, by allowing participants to choose any number of traits as recommended above, participants may opt simply to choose no traits if they have no notions whatsoever about a given target group.

Part I: Narrative History of the Adjective Checklist

Having discussed the affordances and critiques of the adjective checklist measure as it is commonly used, I now turn to a comprehensive review of the 107 papers that have used it to assess stereotype content, ranging from 1933 up until a 2021 preprint (see Appendix A: Locating Papers Using the Adjective Checklist Measure). Some papers include multiple studies, and not all studies are relevant to all characterizations (e.g., Katz & Braly, 1935, gathered trait-level norming data rather than gathering data about any particular target group). In this section, I describe the characteristics of the measure as used across the reviewed studies, and how its use

has changed over time. As with much psychological research, these studies were primarily conducted in the US (64.57%) and commonly relied on undergraduate student samples (65.87%).

Measurement Variation

Source of Terms. The largest plurality of adjective checklist studies re-used the Katz & Braly, 1933 checklist, though 46.67% of these used a modified version (see Table 1 for sources of terms). When studies did generate new lists or use modified existing lists, content was often gathered using a free-response pilot (47.62%). Using a pre-existing measure enables comparability across time and target group. However, it introduces the possibility that language use has shifted in the interim and may not apply well to groups beyond those for which it was designed. Despite many studies using a modified Katz & Braly scale, the authors themselves described it as a list of "racial characteristics (1935, p. 183); it is unlikely to be as informative for research about other groups, or to fully capture racial stereotypes in the present.

Table 1. Where adjective checklist studies have sourced their adjectives.

Adjective Source	Frequency	Proportion of Studies Using this Which Use a Modified Version
New Measure	21.77%	--
Katz & Braly	48.39%	46.67%
Gough & Heilbrun	16.13%	10.00%
Other Re-Used Measure	13.71%	35.30%

Instructions. Though the adjective checklist always requires participants to choose traits to associate with a particular group, instructions vary in three ways: (1) the basis on which participants are told to make their judgments, (2) whether participants are restricted in the total number of traits chosen, and (3) whether participants can add traits. Of reviewed studies, 53.22% asked participants to report their personal beliefs about targets, while 12.90% asked participants to report the cultural stereotype of the target (7.26% did both). This distinction is often attributed

to Devine & Elliot (1995), who found that personal beliefs about Black people were markedly more positive than cultural stereotypes as reported by their participants. The authors therefore argued that the prior Princeton Trilogy studies (Gilbert, 1951; Karlins et al., 1969; Katz & Braly, 1933) showed a decline in consensus of *personal beliefs* about racial groups rather than a change in consensus of *cultural stereotypes*. Other scholars have opted for the cultural stereotype instructions to minimize desirability bias – in this view, the "cultural stereotype" instructions license participants to express their true beliefs (Bjerstedt, 1960; Petsko & Bodenhausen, 2019; Schneider & Bos, 2011, 2014). While the social desirability concern is serious, there is also substantial evidence that people can recognize stereotypes without personally endorsing them (e.g., Rogers & Way, 2016). It seems advantageous to find other means to reduce desirability concerns without relying on cultural stereotype instructions (for a similar argument, see Kotzur et al., 2020).

In 58.97% of studies, participants were restricted in the final number of traits they could choose; in 19.66%, they were able to add adjectives to the list. However, only two studies provided any analysis of these added-in adjectives (Dempsey, 1992; Ehrlich & Rinehart, 1965b). As discussed earlier, allowing participants to choose any number of traits is preferable to restricting participants to an arbitrary number, and allowing participants to add traits and reporting the percentage of participants who did so can indicate the checklist's coverage.

List Length. The number of traits in the list ranged from 12 to 300, with a median of 84. Though one author raised concerns about list length after noting order effects in a 96-trait list (Berreman, 1958), the largest constraint on length seems to be participant time and attention.

How the Measure's Use Has Changed Over Time

It is also worth considering how the use of the adjective checklist has changed over time – in what ways does it reflect changing interests of psychological science? I focus on the target groups about which stereotypes are being measured and the kinds of analysis being conducted.

Target Groups. The adjective checklist measure has been used to study a wide array of social dimensions (see Table 2). The classification of a given social group in a given dimension is not always straightforward – "Japanese" can be viewed both as a nationality and a race/ethnicity, while "Jewish" can be viewed both as a religion and a race/ethnicity – and this has differed across context and history. In this analyses, social groups were coded as representing as many dimensions as might apply. Looking only at domains that are included in at least 5% of the total body of literature, we can see that the initial use of the measure was very focused on race/ethnicity and nationality, and in the time since it has been employed to study a wider array of social dimensions (see Figure 1). In the first 30 years, almost all studies looked at race/ethnicity, nationality, and religion, in line with Katz & Braly (1933).

Table 2. The proportion of adjective checklist studies focused on different social dimensions. Each study could include multiple dimensions.

Social Dimension	Proportion of Studies Addressing Stereotypes in this Dimension
Race or Ethnicity	53.23%
Nationality	43.55%
Gender or Sex	23.39%
Occupation	19.35%
Religion	15.32%
Sexuality	8.87%
Manipulated Ingroup	4.84%
Age	3.23%
(Dis)ability status	3.23%
Parenthood	3.23%
Political Group	2.42%
Class	1.61%

Social Dimension	Proportion of Studies Addressing Stereotypes in this Dimension
Intelligence	1.61%
Language spoken	1.61%
Marital status	1.61%
Physicality/Health	1.61%
Other	6.45%

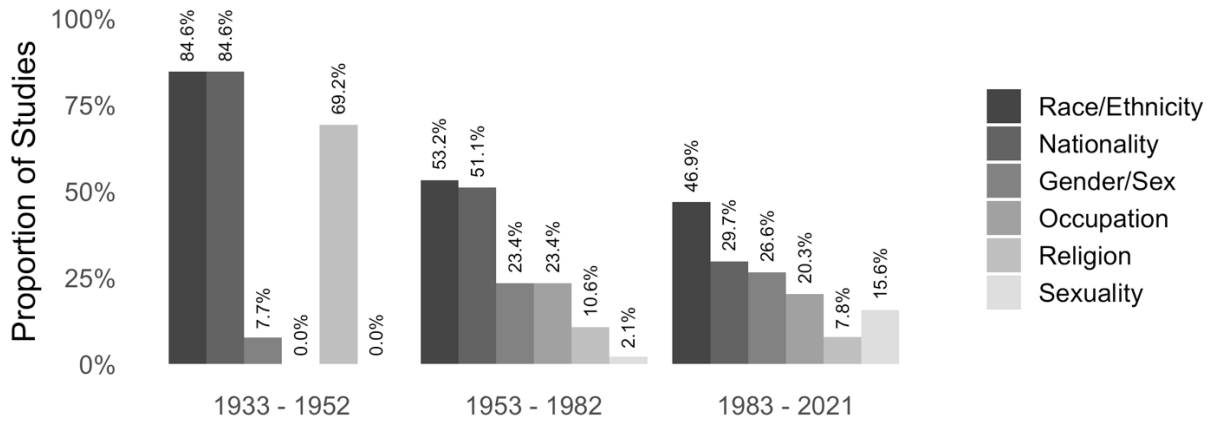

Figure 1. Proportion of Adjective Checklist studies in certain time periods assessing stereotypes of groups in different social domains.

Analytic Approach. The second major shift over time is in the analyses applied to adjective checklist data (see Figure 2). While the use of consensus analysis – addressing the percentage of participants that use the same term to describe the same group – is fairly consistent across time, we can see that four other kinds of analysis trade off over time. Historical analysis, in which the researcher emphasizes the sociohistorical context in which participants are living, has decreased over time. For instance, initial adjective checklist studies which looked at nationality emphasized the international relations between different nations at the time of data collection (e.g., Buchanan, 1951; Seago, 1947). Descriptive analysis, in which the researcher describes the total impression given by the traits chosen for a group – for example, Katz & Braly saying that "the picture of the Japanese seems more clear cut with some recognition of the

westernization of Japan" (1933, p. 287) – has also decreased over time. In their place, norming analysis and statistical target group comparison have increased. This may reflect the growing emphasis within stereotyping research on cognitive processes (cf. Baars, 1986) and statistical rigor (cf. Shrout & Rodgers, 2018), as well as increasing computational power for statistical analyses. However, while norming analysis and statistical target group comparison may serve a similar role to descriptive analyses, they cannot do the same work as sociohistorical analysis (cf. Nzinga et al., 2018; Pettigrew, 2018). Moving forward, researchers should incorporate sociohistorical analysis in their presentation of adjective checklist results.

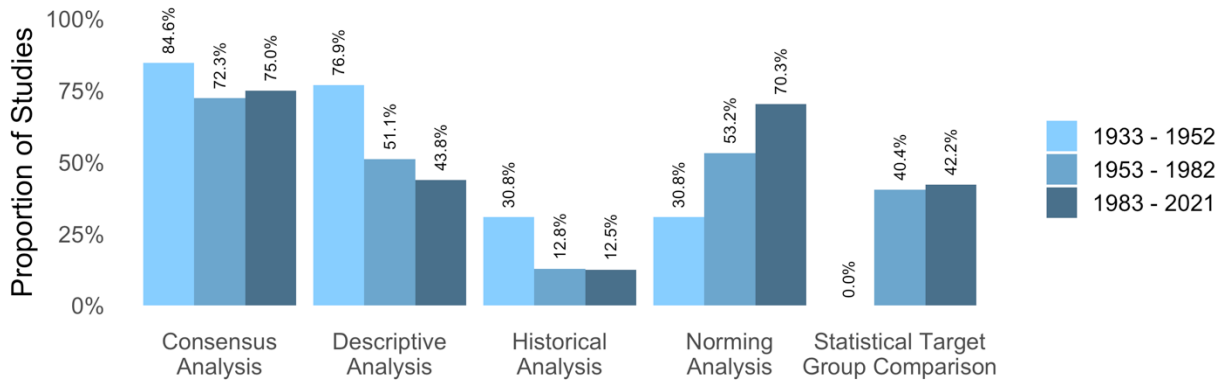

Figure 2. Proportion of adjective checklist studies using different kinds of analysis.

Moving Forward with the Adjective Checklist

Taking stock of the adjective checklist's advantages, disadvantages, and history, a series of recommendations emerge for the continued use of this measure to assess stereotype content. Two are methodologically straightforward: participants should be able to add adjectives to the list as they deem necessary and should not be restricted in the number of adjectives they may choose. By reporting the proportion of participants that opt to add adjectives to the list, the researcher can understand how appropriate the list was for a particular participant population and

target group. By allowing the participant to choose as many adjectives as desired, the researcher avoids shoehorning the participant's genuine response into a more-stereotypical form. A third methodological point concerns the instructions used by the researcher: telling participants to report their own beliefs or the cultural stereotype of a given group. The use of the cultural stereotype instructions should not be assumed to inevitably result in the expression of personal biases; researchers who use these instructions should explain why and, if possible, provide evidence supporting their assumption of how those instructions will be used.

The next three recommendations concern analysis. First, scholars should re-incorporate sociohistorical analysis as a component of understanding stereotype content results. Second, the sharing and re-use of norming data within a given population offers value for analysis relying on adjective checklist measures, and should be pursued. Finally, it is worth asking whether the meaning of certain traits on the list changes based on the other traits that were chosen with them (e.g., *outspoken* and *creative* vs. *outspoken* and *rude;* Asch & Zukier, 1984; Zanna & Hamilton, 1977). One approach to analyzing this would be gathering norming data on the traits that a given participant or sample used to describe a given group (e.g., the words that more than 80% of the sample chose to describe women). Another would be having participants generate synonyms to adjectives as they make their sections – this could illuminate, for instance, that participants are choosing *outspoken* in its meaning of *candid* to describe artists, but in its meaning of *vocal* to describe politicians. Both of these approaches would require data beyond the basic adjective checklist itself, and very directly connect to adjective polysemy. A final possibility, relying only on the adjective checklist data, is to see whether the correlation in choosing given adjectives differs by target group. In describing one group, if the choice of several terms is highly correlated, such that they are regularly chosen (or not chosen) together, this suggests they are

semantically associated in the context of evaluating this target group. If selection of those same terms is uncorrelated when evaluating another target group, it suggests they are not semantically associated with respect to this target group. This means that the level of association between the two terms depends on the target group – the target group is providing a frame of interpretation for the adjectives. For example, if the choice of *outspoken* was highly correlated with the choice of *reflective* for describing artists, but not for describing politicians, that would provide additional information about how to interpret the choice of *outspoken* in each case. Such analysis could be done by looking at the correlation in choice of particular adjectives (e.g., *outspoken* and *reflective*), or by comparing factor analysis or semantic network analysis on adjective co-selection across target groups.

I now turn to a final recommendation: using a list of terms appropriate for contemporary reference across a wide array of social domains. In 1965, Ehrlich & Rinehart asked: "Why is it that an instrument composed of adjectives culled from the open-ended responses of 25 students attending college in 1932 has been used so predominantly with little modification?" (1965a, p. 421). Though the Katz & Braly measure has been updated and amended (e.g., Devine & Elliot, 1995; Galinsky et al., 2013), their central critique holds. In the next section, I describe the development of an updated version of the adjective checklist measure that is applicable across an array of social dimensions.

Part II: Developing a New Checklist

In this section, I describe the generation of a new adjective checklist measure for assessing stereotype content. This is a six-step process: defining groups of interest, sourcing adjectives, deciding which groups should be retained, selecting a subset of adjectives, gathering norming data about the potential list, and selecting the final list. The goal is to obtain a list that

(a) is applicable to a wide array of social groups, (b) uses contemporary language, (c) is of a reasonable length, and (d) has convergent validity with other measures of attitudes towards a target group. All analysis was conducted in R, using the tidyverse, tidytext, psych, jsonlite, lme4, and lmerTest packages (Bates et al., 2015; Kuznetsova et al., 2017; Ooms, 2014; R Core Team, 2018; Revelle, 2019; Silge & Robinson, 2016; Wickham et al., 2019).

Defining Groups of Interest

I chose 17 social domains of interest for construction of the new measure (see Table 3). This includes five of the social domains represented by at least 5% of adjective checklist studies to date: race/ethnicity, gender/sex, occupation, religion, and sexuality. Nationality was omitted, and twelve other domains were added. Domains were intentionally chosen so as to vary in everyday social salience, perceived category origin, and personal experience. In some domains, social media data was gathered for more target groups than free-response data.

Table 3. Social Groups about which stereotype content was gathered from each source. Italicized groups have data from Twitter and Reddit but not the free-response.

Social Dimension	Social Groups
Body type	Fat people, Muscular people, Short people, Skinny people, Tall people
Diet	Dieters, Meat-eaters, Omnivores, *Pescatarians, Vegans,* Vegetarians
Employment class	Blue collar workers, Unemployed people, White collar workers
Eye color	Blue-eyed people, Brown-eyed people, Green-eyed people
Gender	*Boys, Girls, Gender nonconforming people,* Genderqueer people, Men, Non-binary people, Transgender people, *Transmen, Transwomen,* Women
Handedness	Ambidextrous people, Left-handed people, Right-handed people
Hobbies	Gardeners, Movie buffs, Sports fans, Woodworkers,
Kinship	Aunts, *Brothers, Children,* Cousins, Daughters, Fathers, *Grandchildren, Granddaughters,* Grandfathers, Grandmothers, Grandparents, *Grandsons*, Mothers, Parents, Siblings, *Sisters*, Sons, Uncles
Occupations	Administrative assistants, Bankers, Construction workers, *Secretaries*, Students, Tech entrepreneurs
Political party	*Conservatives*, Democrats, *Liberals,* Libertarians, Progressives, Republicans
Race/ethnicity	Asian people, Black people, Hispanic people, Native American people, White people

Social Dimension	Social Groups
Religion	*Agnostics*, *Atheists,* Buddhists, Catholics, *Christians, Hindus, Jewish people,* Muslims, Protestants, *Quakers*
School tropes	Jocks, Loners, Nerds, Stoners
Sexuality	Bisexual people, *Gay people,* Heterosexual people, Homosexual people, *Pansexual people,* Queer people
Sexuality-Gender	Bisexual men, Bisexual women, Gay men, Lesbian women
US region	East coasters, Midwesterners, Southerners, West coasters
Wealth class	Middle class people, Poor people, Rich people, *Working class people*

Sourcing Adjectives

To gather potential adjectives for a new checklist measure, I relied on free-response data from lab participants (as in the prior literature). However, I expanded this traditional approach by incorporating stereotyping content from two online sources: Reddit and Twitter. In recent years, modern technology and advancing computational techniques are providing researchers an opening to leverage massive routinely-occurring corpora (Iliev et al., 2015), and there is some evidence that analysis of such data capture stereotype content better than lab-based ones (Nicolas et al., 2020). Public text analyses have shown, for example, that the stereotyping terms most connected to gender and racial/ethnic groups have shifted over the past 100 years in ways that can be predicted by national demographic and employment changes (Bhatia & Bhatia, 2020; Garg et al., 2018).

For someone who wants to use text analysis to make decisions (e.g., with resume review, Dastin, 2018), stereotype content in language is a problem because algorithms carry forward stereotype-based inequities (Mitchell et al., 2021). This is particularly the case because the producers of online content are no more a random sample of the population than is a class of psychology undergraduate students, and online spaces can recreate existing social hierarchies (Bender et al., 2021). The use of freely-generated public online text data is not a panacea for the

sampling issues that plague psychology (Henrich, 2010; Hruschka et al., 2018; Medin et al., 2010; Rad et al., 2018), but it is worth incorporating because it captures culturally dominant narratives about different social groups. Online content is disproportionately produced by those with social power, whose stereotypes are also most likely to influence the lives of those with less power. For someone who wants to understand language biases as indices of stereotype content, online data and text analysis expand the realm of analytic possibility.

For this project, I source adjectives using both lab-based and web-based approaches, casting the net wide. I focus on Reddit and Twitter specifically because these are readily available text sources, ensuring the work can be replicated and checked by other scholars, or compared across different periods of history. The process for extracting adjectives from social media is described in brief in the next few sections of this section, and in detail in Appendix C: Data Processing Pipeline for Social Media Data.

Online Speech: Reddit Post Extraction. Reddit, which calls itself the "front page of the Internet", was founded in 2005. As of October 2019, it reported having 430 million monthly active users (Murphy, 2019). It is centrally organized around forums about particular topics (also called subreddits; for example, r/aww is all posts about adorable things). People contribute to the site either by starting subreddits, posting within subreddits (a *submission*), or responding in a thread resulting from a post (a *comment*). Posts can be up to 40,0000 characters. Reddit is anonymous in that people's submissions and comments are connected only with their usernames. Though this means we cannot know more about the individual person posting (e.g., where in the world they are), we do know that Reddit is the sixth most popular website in the United States, that about 50% of users are in the US (the next largest representation is 8% from the UK), and that Reddit is generally more popular among younger than older age groups (Duggan & Smith,

2013; Sattelberg, 2020; *Who Uses YouTube, WhatsApp and Reddit*, 2019). Reddit data has appeared in many social science studies and been used as an informative source of publicly-accessible naturally-occurring language and social interaction (e.g., Apostolou, 2019; Cinelli et al., 2020; Datta & Adar, 2019; Unkel & Kümpel, 2020).

I relied on the Pushshift Reddit dataset (Baumgartner et al., 2020), which is an independently collated source of all Reddit posts & comments. As an independent archive, this retains data from subreddits that have been deleted from the main website (for example, subreddits which were banned for hate speech in 2020, like r/The_Donald or r/LGBdroptheT), or posts which people have gone back and deleted. My sampling population is all of the comments and submissions produced during 2019, as indexed in the Pushshift dataset. All of 2019 comprises 2.78 terabytes of data. Instead of processing all of 2019, two chunks of 5500000 comments and one chunk of 2060000 submissions (about 6.5 GB per chunk) were randomly selected from each month of that year, such that roughly 234 GB of Reddit data were selected as the sample. Comments and submissions were filtered for references to the target groups of measure development, including synonyms and alternate spellings (see Appendix B: All Search Terms Used for Online Adjective Sourcing). This led a corpus of 14,552,285 Reddit posts, where each social group appeared in 6 to 6,480,762 posts (see Table 4).

Online Speech: Twitter Post Extraction. Twitter, a social media microblogging website, was founded in 2006. As of 2019, twitter had 126 million daily active users (Shaban, 2019), and 22% of American adults used the website (Wojcik & Hughes, 2019). People contribute to the site by posting short messages of 240 characters or less. These posts can be independent, or in response to posts by other Twitter users, and users can also retweet something another user posted. Twitter users are generally identifiable, unlike Redditors. In 2019, about

26.36% of Twitter users were in the US (Iqbal, 2018). Twitter data has appeared in many social science studies and been used as an informative source of publicly-accessible naturally-occurring language and social interaction (e.g., Hswen et al., 2021; Jaidka et al., 2020; Joseph et al., 2017).

I used Twitter's academic-track API, with which Twitter allows academic researchers to collect up to 10 million tweets per month (*Advancing Academic Research with Twitter Data*, 2021). Researchers can request tweets from the Twitter archive, which goes back to the start of the website. However, tweets are omitted if they are: from suspended or banned accounts (e.g., Donald Trump's account), deleted, or private (13% of adult users in the US have their accounts private). My sampling population is all tweets produced during 2019, as indexed by date and time in the archive. To accommodate the API's rate limit, I randomly selected one six-hour period from each month of 2019 to make up my sample. For each window of time, every new English-language tweet which referenced one of the target groups (or a synonym or alternate spelling; see Appendix B: All Search Terms Used for Online Adjective Sourcing) was extracted from the API. After gathering the tweets, I looked at the location of the person tweeting – this is mostly unknown (95.87%), but where the location was known and was outside the US (1.71%), I removed the tweet from the corpus. This led to a corpus of 7,555,933 Twitter posts, where each social group appeared in 3 to 2,350,786 posts (see Table 4).

Table 4. The percentage of posts in each corpus referencing each group.

Social Domain	Social Group	Proportion of Twitter Posts about Each Group	Proportion of Reddit Posts about Each Group
Body Type	Fat people	0.0215%	0.0726%
	Muscular people	0.0138%	0.0172%
	Short people	0.0119%	0.0140%
	Skinny people	0.0097%	0.0131%
	Tall people	0.0091%	0.0156%
Diet	Dieters	0.0060%	0.0085%
	Meat-eaters	< 0.0001%	0.0002%

Social Domain	Social Group	Proportion of Twitter Posts about Each Group	Proportion of Reddit Posts about Each Group
	Omnivores	0.0000%	< 0.0001%
	Pescatarians	0.0056%	0.0050%
	Vegans	0.4811%	0.6165%
	Vegetarians	0.1129%	0.1812%
Employment Class	Blue collar workers	0.0045%	0.0088%
	Unemployed people	0.1618%	0.1488%
	White collar workers	0.0007%	0.0024%
Eye Color	Blue-eyed people	0.0003%	0.0008%
	Brown-eyed people	0.0002%	0.0003%
	Green-eyed people	0.0001%	0.0003%
Gender	Boys	6.3926%	4.2077%
	Gender non-binary people	0.0522%	0.1073%
	Gender nonconforming people	0.0027%	0.0066%
	Genderqueer people	0.0039%	0.0087%
	Girls	8.4968%	7.1010%
	Men	31.1118%	44.5343%
	Transgender people	0.6030%	1.2111%
	Transmen	0.0047%	0.0128%
	Transwomen	0.0152%	0.0324%
	Women	11.9489%	11.4935%
Handedness	Ambidextrous people	0.0035%	0.0118%
	Left-handed people	0.0238%	0.0566%
	Right-handed people	0.0106%	0.0382%
Hobbies	Gardeners	0.0382%	0.0344%
	Movie buffs	0.0016%	0.0028%
	Sports fans	0.0163%	0.0235%
	Woodworkers	0.0584%	0.0668%
Kinship	Aunts	0.3984%	0.3921%
	Brothers	2.6849%	2.9102%
	Children	11.4152%	14.7034%
	Cousins	0.5444%	0.7599%
	Daughters	1.3327%	1.5801%
	Fathers	4.5060%	6.0332%
	Grandchildren	0.1163%	0.1381%
	Granddaughters	0.0427%	0.0304%
	Grandfathers	0.2717%	0.4959%
	Grandmothers	0.7176%	0.9754%
	Grandparents	0.1265%	0.2820%
	Grandsons	0.0670%	0.0504%

Social Domain	Social Group	Proportion of Twitter Posts about Each Group	Proportion of Reddit Posts about Each Group
	Mothers	5.0948%	6.6281%
	Parents	2.3348%	5.0116%
	Siblings	0.3995%	0.4505%
	Sisters	1.6106%	1.9504%
	Sons	2.2219%	2.5906%
	Uncles	0.4551%	0.6110%
Occupation	Administrative assistants	0.0139%	0.0050%
	Bankers	0.0911%	0.0626%
	Construction workers	0.3270%	0.4439%
	Secretaries	0.3107%	0.1404%
	Students	3.6613%	2.5714%
	Tech entrepreneurs	0.0028%	0.0009%
Political Ideology	Conservatives	0.5247%	0.4645%
	Democrats	2.4265%	1.1754%
	Liberals	0.4463%	0.5310%
	Libertarians	0.0206%	0.0915%
	Progressives	0.0764%	0.0988%
	Republicans	1.3391%	0.8259%
Race-Ethnicity	Asian people	0.0653%	0.1911%
	Black people	4.0444%	0.9600%
	Hispanic people	0.1140%	0.1140%
	Native American people	0.0958%	0.1044%
	White people	0.6104%	0.7527%
Religion	Agnostic people	0.0174%	0.0754%
	Atheists	0.1084%	0.3843%
	Buddhists	0.0350%	0.0781%
	Catholics	0.2861%	0.4424%
	Christians	0.9244%	1.1436%
	Hindus	0.3878%	0.1161%
	Jewish people	0.4331%	0.5733%
	Muslims	0.9433%	0.7232%
	Protestants	0.0226%	0.0747%
	Quakers	0.0118%	0.0117%
School Tropes	Jocks	0.0871%	0.0538%
	Loners	0.0192%	0.0360%
	Nerds	0.4184%	0.5310%
	Stoners	0.0451%	0.0916%
Sexuality	Bisexual people	0.0971%	0.2008%
	Gay people	0.2161%	0.3481%

Social Domain	Social Group	Proportion of Twitter Posts about Each Group	Proportion of Reddit Posts about Each Group
	Heterosexual people	0.0846%	0.1783%
	Homosexual people	0.1252%	0.3222%
	Pansexual people	0.0161%	0.0316%
	Queer people	0.0251%	0.0286%
Sexuality-Gender	Bisexual men	0.0042%	0.0133%
	Bisexual women	0.0034%	0.0148%
	Gay men	0.0455%	0.1093%
	Lesbian women	0.2929%	0.3773%
US Region	East Coasters	0.0015%	0.0033%
	Midwesterners	0.0052%	0.0074%
	Southerners	0.0189%	0.0275%
	West Coasters	0.0014%	0.0024%
Wealth Class	Middle class people	0.1516%	0.2213%
	Poor people	0.1122%	0.1904%
	Rich people	0.1144%	0.2169%
	Working class people	0.1243%	0.1512%

Online Speech: Adjective Extraction. In each corpus, I extracted the sentence (or sentences) in each post that referenced the target group, then split that sentence into its individual strings (this included some non-words, such as *abcnews*). Strings which appeared in the search terms were excluded (e.g., *caucasian*).

Lab Data: Free-Response. In line with prior adjective checklist development (e.g., Garcia-Marques et al., 2006; Prothro & Melikian, 1954; Sinha & Upadhyaya, 1960; Udolf, 1973), I recruited participants to generate multi-adjective descriptions of each of the groups of interest. Participants were undergraduate students at Northwestern University, who participated for course credit in March 2020. Two hundred and fifty-nine participants completed the study; none were excluded (M_{age} = 18.65 years, SD_{age} = 0.86; 59.07% female, 40.15% male; 46.72% White or Caucasian, 25.48% Asian or Asian American, 11.97% Multiracial, 8.11% Black, African American, or African, 7.33% Latino or Latina). Following the method of Niemann et al. (1994), participants were asked to list the first five adjectives that came to mind when they

thought of members of a particular social group. Each participant completed this description for fifteen groups, randomly selected and in a random order out of the total set of groups. Each group was described by 37 to 67 participants.

Choosing Groups to Include

Up to this point, the measure development process has cast a wide net, looking for stereotype content about an array of groups. However, it is worth asking whether all of these groups are equally suitable for inclusion in a stereotype content measure; are they all sufficiently socially salient? To assess general social salience, I look at the number of posts about each group in the Twitter and Reddit data. The number of Twitter posts and Reddit posts about a group are strongly correlated ($r(98) = 0.98$, $p < 0.0001$), suggesting that group salience is very similar in the two settings. Though this correlation is quite high, some differences appear when considering the relative frequency of different social groups in each medium. For instance, kinship groups tend to be more frequently-referenced on Reddit (Men, Children, Parents, Mothers, Fathers), while political groups (Democrats, Republicans), Black people, Girls, and Boys are more frequently referenced on Twitter (see Table 4). This suggests that, while both data sources have been used to understand "everyday language," the discussions occurring in each case do differ, as the sites differ in membership in function – for the purpose of measure development, this has the advantage of incorporating content from both mediums.

To allow comparison between the two sources, I define target group frequency as the total number of posts referencing a group within the corpus, divided by the total number of unique posts in the corpus. There are many more low-frequency than high-frequency groups in both media sources (see Table 4), and groups defined primarily by gender and kinship have the highest overall frequency (see Figure 3). Groups defined by eye color, US region, handedness,

body type, and hobby appear quite rarely. Being conservative in removing groups from consideration, I retained all groups which appeared in at least 0.1% of the gathered posts in either corpus, retaining 54 of the original 101 groups.

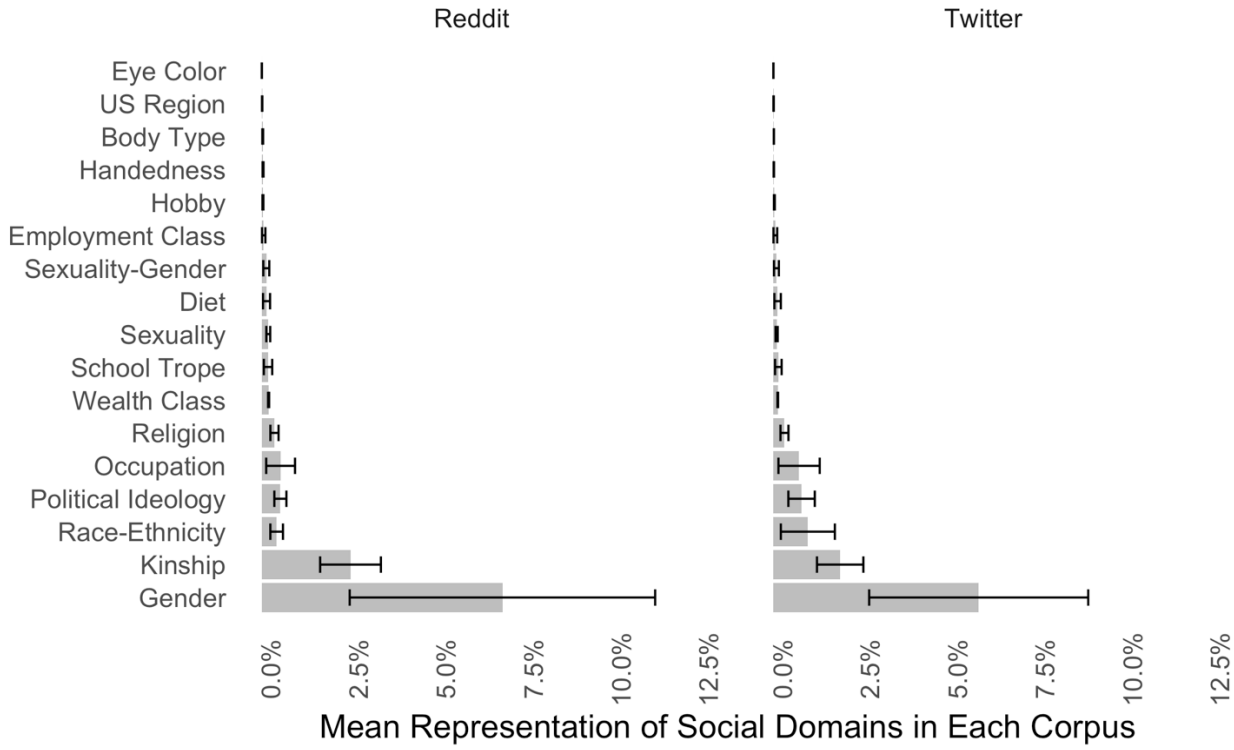

Figure 3. Frequency of references to a social domain in each corpus. Mean is calculated as the mean of all target groups within the social domain, and error bars show standard error based on that calculation.

From there, I excluded target groups if they were the only remaining group in their social domain. *Unemployed People* was the only Employment Class group that appeared frequently enough (*Blue Collar Workers* and *White Collar Workers* both appeared too rarely; see Table 4), so it was reassigned to the Wealth Class domain. *Nerds* was the only school trope retained (*Jocks*, *Loners*, and *Stoners* all appeared too infrequently; see Table 4); as it did not fit well into any of the other social domains, it was removed from further measure development. This produced a list of 53 groups (see Table 7).

Defining Adjectives of Interest

After reducing the total number of groups, I selected the top 10 adjectives associated with each group in each corpus (whether a string was an adjective or not was defined according to Google's English Dictionary, which is based on the Oxford Languages dictionary; (Oxford Languages, 2021). Adjectives were omitted if (a) they occurred less than once per 1,000,000 words in a public corpus of online speech (e.g., *zionist*; Gimenes & New, 2016), (b) they referenced a specific location or group (e.g., *caribbean, buddhist*), (c) they referenced a consistent physical feature (e.g., *muscular*), (d) they referenced one of the retained target groups (e.g., *islamist, girly*; this includes all search terms for the retained groups as listed in Appendix B: All Search Terms Used for Online Adjective Sourcing), (e) they were comparative (e.g., *younger*), or (f) they seemed unlikely to apply to people as standalone adjectives (e.g., *fiscal*; see Table 5 for all words in this final category). Though the adjectives in this last group did commonly appear in the same sentences as some target groups, this is most likely an effect of other nouns that frequently co-occurred with the target group in the corpora (e.g., *fiscal policy* co-occurring with Democrats and Republicans), or of the term appearing as a noun (e.g., *economy* co-occurring as a noun with Political Ideology groups), or to emphasize other terms (e.g., *fucking, very*). I therefore opted to selectively remove these (Table 5). Several adjectives were quite frequent across groups in the corpora – these are not retained for a particular group because they are somewhat generic, but are listed separately in Table 6 and included in the list of terms for the measure. Table 7 presents the adjectives that were selected from each source about each retained group. In cases where the tenth most-frequent adjective shared frequency with other adjectives, all were selected (italicized in the table).

Table 5. Terms that were omitted from consideration

Actual	Cardinal	Economy	Fucking	Looking	Our	Same	Themed
Animal	Chicken	Electrical	Future	Master	Outside	Serial	Through
Attacking	Coming	Elementary	Going	Missing	Own	Specialist	Top
Away	Commercial	Family	Handmade	Name	Past	State	Turned
Baby	Congressional	Federal	High	New	Plus	Still	Under
Becoming	Daily	Few	Human	Next	Possible	Supporting	Vegetable
Born	Dairy	Fiscal	Inside	Nothing	Prime	Taking	Veggie
Bought	Demographic	Former	Like	Now	Public	Talking	Very
Budget	Done	Found	Likely	Nuclear	Racial	Tech	Weekly
Capital	Due	Founding	Living	Only	Raised	Telling	

Table 6. Terms that appeared frequently across groups within the corpus and were retained for the measure

Bad	False	Great	Kind	Married	Pretty	Together	Well
Dead	Free	Happy	Left	Mean	Real	True	Whole
Different	Game	Hard	Live	Nice	Right	Trying	Young
Down	Good	Just	Lost	Old	Sure	Used	

Table 7. Top adjectives for all groups from each source. Adjectives are listed in order of frequency.

Social Domain	Target Group	Reddit	Twitter	Free-Response
Diet	Vegans	Healthy, easy, ethical, choice, wrong, friendly, able, raw, moral, low	Healthy, organic, delicious, easy, natural, friendly, raw, amazing, total, available	
	Vegetarians	Healthy, certain, easy, adequate, strict, choice, dietary, ethical, low, appropriate	Healthy, delicious, easy, organic, available, suitable, tasty, choice, gluten-free, *low, yummy*	Healthy, caring, smart, annoying, active, conscious, fit, picky, religious, weak
Gender	Boys	Magic, mad, wrong, proud, fine, cute, sweet, sorry, healthy, blue	Proud, cute, beautiful, sweet, handsome, sad, ready, wrong, sorry, mad	
	Gender Non-Binary People	Feminine, masculine, wrong, valid, fluid, comfortable, fit, social, feeling, neutral	Valid, inclusive, feminine, fluid, neutral, sexual, masculine, okay, amazing, disabled	Brave, unique, confusing, confident, cool, misunderstood, fluid, smart, strong, *annoying, expressive, honest, normal, progressive, sensitive, weird*
	Girls	Magic, cute, beautiful, interested, attractive, wrong, thinking, weird, sexy, dated	Beautiful, cute, sweet, sexy, amazing, fine, proud, crazy, sorry, okay	
	Men	Sorry, wrong, cool, able, thinking, awesome, weird, funny, level, amazing	Sorry, wrong, amazing, funny, cool, united, crazy, mad, beautiful, awesome	Strong, smart, masculine, powerful, hardworking, privileged, assertive, aggressive, arrogant, loud
	Transgender People	Mental, wrong, medical, sexual, social, biological, able, general, feminine, thinking	Disabled, wrong, medical, mental, sexual, biological, legal, activist, important, *able, safe*	Brave, strong, cool, courageous, confident, unique, *independent, interesting, misunderstood, normal, proud*

Social Domain	Target Group	Reddit	Twitter	Free-Response
	Women	Sexual, attractive, pregnant, beautiful, wrong, strong, able, average, interested, social	Beautiful, strong, amazing, racist, pregnant, sexual, sexy, wrong, lovely, crazy	Strong, caring, intelligent, smart, powerful, beautiful, hardworking, independent, loving, *compassionate, diverse, emotional, understanding*
Kinship	Aunts	Cool, crazy, decided, able, entitled, wrong, weird, sorry, visiting, front	Grand, beautiful, cool, sorry, proud, favorite, ready, amazing, dear, crazy	Caring, loving, cool, funny, friendly, distant, supportive, *chill, compassionate, generous, hardworking, helpful, sweet*
	Brothers	Sorry, able, wrong, decided, strong, alone, magic, thinking, cool, front	Dear, proud, sorry, cool, welcome, strong, wrong, amazing, sad, thinking	
	Children	Adult, able, wrong, serious, stupid, thinking, cool, alone, giving, normal	Stray, adult, safe, cool, sick, wrong, able, amazing, social, sexual	
	Cousins	Distant, removed, weird, able, decided, wrong, sorry, thinking, cool, front	Favorite, proud, distant, grand, beautiful, cute, okay, crazy, funny, sorry	Friendly, funny, distant, loving, smart, caring, cool, crazy, loud, *annoying, supportive*
	Daughters	Able, magic, pregnant, wrong, adult, front, alone, sorry, beautiful, giving	Beautiful, proud, amazing, grand, lovely, cute, sorry, able, killing, favorite	Caring, sweet, loving, feminine, smart, cute, hardworking, innocent, delicate, *playful, responsible*
	Fathers	Able, sorry, wrong, abusive, biological, decided, thinking, growing, early, alone	Proud, sorry, dear, holy, heavenly, late, sad, okay, grand, amazing	Caring, loving, hardworking, strong, protective, funny, *absent, busy, strict, supportive, wise*

Social Domain	Target Group	Reddit	Twitter	Free-Response
	Grandchildren	Able, giving, wanting, adult, chance, alone, wrong, pregnant, sad, important	Racist, able, legacy, proud, united, beautiful, sorry, wonderful, thinking, giving	
	Grandfathers	Late, sorry, able, alive, available, cool, early, decided, dying, racist	Late, proud, magic, racist, sorry, crazy, sad, thinking, favorite, grave	Wise, caring, sweet, loving, funny, tired, elderly, generous, smart, stubborn
	Grandmothers	Able, sorry, late, decided, sweet, thinking, front, sick, dying, alone	Sorry, proud, cute, sweet, beautiful, late, favorite, funny, okay, ready	Caring, loving, sweet, wise, gentle, helpful, warm, frail, supportive, *compassionate, cute, elderly, funny, smart, thoughtful, welcoming*
	Grandparents	Able, visiting, alive, growing, spent, early, late, decided, alone, wrong	Important, visiting, proud, able, special, forever, alive, united, growing, late	Loving, caring, wise, sweet, funny, friendly, tired, warm, *cute, elderly, experienced, generous, relaxed, smart, supportive*
	Mothers	Able, sorry, wrong, alone, decided, pregnant, abusive, thinking, front, crying	Proud, sorry, beautiful, dear, mad, cute, amazing, okay, crying, wrong	Caring, loving, strong, hardworking, warm, protective, supportive, helpful, smart, compassionate
	Parents	Serious, removed, able, independent, adult, wrong, divorced, entitled, abusive, shitty	Proud, able, wrong, social, important, private, amazing, adult, paid, sad	Loving, caring, hardworking, strict, protective, supportive, responsible, stressed, helpful, strong

Social Domain	Target Group	Reddit	Twitter	Free-Response
	Siblings	Able, growing, normal, adult, alone, weird, wrong, decided, fine, abusive	Vital, grand, favorite, cute, growing, fighting, weird, annoying, sad, wrong	Loving, annoying, caring, friendly, supportive, funny, helpful, playful, *dependable, dependent, smart, trustworthy, understanding*
	Sisters	Able, decided, wrong, pregnant, alone, magic, sorry, weird, thinking, front	Dear, beautiful, proud, amazing, sorry, sweet, cute, excited, funny, *late, okay*	
	Sons	Holy, able, wrong, fair, sorry, magic, proud, adult, alone, giving	Proud, holy, sorry, beautiful, amazing, crying, ready, favorite, wrong, cute	Strong, masculine, smart, hardworking, messy, obedient, *active, adventurous, athletic, caring, loud, loving, loyal, responsible*
	Uncles	Cool, creepy, drunk, weird, crazy, racist, able, sorry, wrong, favorite	Creepy, grand, drunk, racist, crazy, cool, sorry, proud, favorite, late	Funny, friendly, caring, distant, hardworking, supportive, cool, crazy, weird, *helpful, loud, loving*
Occupa-tion	Construction Workers	Independent, paid, private, base, general, able, local, hired, level, specific	Independent, perfect, private, general, local, paid, concrete, available, custom, national	Hardworking, strong, tough, brave, tired, motivated, diligent, dirty, masculine, *dedicated, determined, persistent, skilled, smart*
	Secretaries	General, foreign, united, national, slutty, acting, anal, chief, executive, international	General, foreign, chief, acting, national, united, permanent, executive, legal, interior	
	Students	Graduate, able, medical, level, paid, current, average, private, important	Amazing, international, proud, medical, welcome, paid, social, academic, excited	Stressed, tired, depressed, anxious, smart, busy, curious, *ambitious, intelligent, sad*

Social Domain	Target Group	Reddit	Twitter	Free-Response
Political Ideology	Conservatives	Social, political, racist, religious, general, wrong, progressive, stupid, moderate, funny	Racist, political, social, general, wrong, stupid, safe, fake, national, decent	
	Democrats	Social, political, democratic, moderate, progressive, corporate, running, racist, base, chance	Racist, socialist, political, illegal, running, corrupt, democratic, stupid, wrong, fake	Diverse, progressive, smart, educated, blue, caring, loud, aware, passionate, political
	Liberals	Political, social, stupid, racist, classical, wrong, progressive, mad, leftist, dumb	Racist, stupid, fake, political, crazy, socialist, mad, corrupt, social, wrong	
	Republicans	Democratic, political, racist, moderate, base, wrong, able, stupid, partisan, fair	Racist, criminal, stupid, corrupt, defending, democratic, political, questioning, wrong, afraid	Stubborn, religious, selfish, ignorant, traditional, political, racist, sexist, *arrogant, closeminded, hardworking, rude, wrong*
Race-Ethnicity	Asian People	Racist, crazy, average, general, social, central, smart, common, affirmative, considered	Crazy, racist, funny, proud, beautiful, general, low, wrong, average, fellow	Smart, hardworking, quiet, talented, diverse, driven, intelligent, reserved, strict, *dedicated, diligent, disciplined, funny, motivated, successful*
	Black People	Racist, wrong, violent, civil, average, stupid, social, funny, dark, general	Racist, mad, crazy, wrong, funny, dumb, stupid, weird, fake, lame	Hardworking, strong, athletic, loud, smart, friendly, intelligent, *beautiful, normal, passionate, resilient*
	Hispanic People	Racist, illegal, mixed, general, united, speaking, considered, average, central, common	Racist, illegal, speaking, political, united, low, important, winning, democratic, *legal, national*	Hardworking, friendly, loud, diverse, funny, passionate, smart, spicy, *caring, religious, traditional, warm*

Social Domain	Target Group	Reddit	Twitter	Free-Response
	Native American People	Racist, united, killing, cultural, forced, modern, ethnic, wrong, general, national	Racist, sorry, deep, forced, tribal, national, united, general, acknowledged, cultural	Traditional, spiritual, oppressed, hardworking, hurt, smart, strong, tribal, *caring, cultural, determined, disadvantaged, loyal, respectful*
	White People	Racist, wrong, average, okay, oppressed, social, general, superior, stupid, able	Racist, dear, wrong, crazy, stupid, funny, mad, thinking, dumb, weird	Privileged, hardworking, ignorant, normal, annoying, bland, cool, racist, rude, smart
Religion	Atheists	Religious, wrong, moral, secular, thinking, edgy, spiritual, common, militant, stupid	Religious, secular, wrong, moral, anonymous, offended, simply, funny, honest, intelligent	
	Catholics	Orthodox, religious, holy, devout, wrong, private, traditional, sexual, local, practicing	Religious, holy, central, sexual, social, orthodox, faithful, wrong, national, traditional	Religious, strict, traditional, faithful, spiritual, *boring, friendly, holy, honest, loving, moral, old-fashioned, rigid*
	Christians	Religious, wrong, early, orthodox, modern, holy, devout, pagan, moral, political	Fake, religious, holy, wrong, racist, political, social, killing, early, spiritual	
	Hindus	Magic, religious, nationalist, ancient, secular, wrong, political, killing, sacred, holy	Secular, religious, terrorist, converted, fake, killing, nationalist, proud, forced, political	
	Jewish People	Religious, killing, orthodox, racist, ethnic, political, wrong, modern, considered, common	Racist, killing, wrong, democratic, orthodox, political, religious, alone, fine, fake	

Social Domain	Target Group	Reddit	Twitter	Free-Response
	Muslims	Religious, terrorist, racist, magic, wrong, killing, political, radical, famous, united	Terrorist, religious, killing, racist, illegal, radical, political, wrong, secular, innocent	Religious, strict, devout, misunderstood, dedicated, faithful, reserved, strong, *devoted, hardworking, quiet*
Sexuality	Bisexual People	Sexual, wrong, opposite, interested, thinking, attractive, okay, romantic, confused, fine	Cute, sexual, valid, horny, sissy, proud, okay, wrong, sexy, fellow	Normal, cool, confident, *adventurous, brave, curious, fluid, friendly, loving, quirky, strong, unique, weird*
	Gay People	Wrong, average, religious, killing, normal, common, sexual, okay, fine, political	Okay, killing, sorry, wrong, religious, social, funny, winning, racist, made	
	Heterosexual People	Sexual, normal, opposite, okay, wrong, romantic, able, limited, equal, social	Sexual, okay, weird, sorry, wrong, dumb, normal, ugly, *cute, funny, uncomfortable*	Normal, boring, accepted, regular, common, average, typical, loving, *cool, funny, ordinary, privileged, standard, traditional*
	Homosexual People	Sexual, average, wrong, common, low, opposed, modern, insulting, fine, religious	Sexual, wrong, cute, official, okay, sexy, fine, killing, dear, dumb	Normal, outgoing, friendly, loving, proud, cool, funny, brave, caring, confident
Sexuality -Gender	Gay Men	Sexual, feminine, masculine, anal, attractive, wrong, interested, general, curious, social	Disabled, sexual, wrong, proud, thinking, funny, acting, sorry, okay, *sick, turbulent*	Normal, funny, proud, confident, fashionable, feminine, honest, *energetic, friendly, outgoing, strong*
	Lesbian Women	Welcome, magic, butch, sexual, negative, wrong, interested, feminist, thinking, masculine	Butch, sexy, teen, anal, cute, horny, amateur, disabled, okay, sexual	Strong, independent, bold, cool, masculine, proud, tough, *athletic, funny, progressive, quirky*

Social Domain	Target Group	Reddit	Twitter	Free-Response
Wealth Class	Middle Class People	Average, low, social, economic, able, growing, educated, suburban, shrinking, paid	Low, social, private, paid, shrinking, educated, economic, average, giving, taxing	Hardworking, average, comfortable, normal, content, educated, common, traditional, *diligent, friendly, shrinking, smart, stressed, struggling, stuck, tired*
	Poor People	Able, giving, social, lazy stupid, homeless, expensive, low, cheap, economic	Giving, financial, able, paid, social, sick, stupid, expensive, political, *racist, sad, sorry*	Hardworking, sad, unfortunate, unlucky, dirty, stressed, homeless, struggling, depressed, tough
	Rich People	Able, expensive, giving, private, social, average, fair, stupid, paid, normal	Giving, private, paid, smart, stupid, fair, able, social, expensive, *average, running*	Hardworking, arrogant, privileged, luxurious, rude, smart, successful, *entitled, generous, greedy, lucky, selfish*
	Unemployed People	Homeless, employed, able, low, paid, social, depressed, spent, lazy, disabled	Graduate, homeless, employed, busy, low, disabled, paid, social, level, sick	Lazy, unlucky, unfortunate, sad, struggling, stressed, tired, difficult, *depressed, desperate, determined, scared, unhappy*
	Working Class People	Social, capitalist, political, economic, socialist, democratic, ruling, average, low, rural	Elite, political, social, socialist, racist, weak, published, helpless, reproductive, private	

It is clear that the adjectives differ by source. This is likely a result of (a) the difference between co-occurrence and description, and (b) the authors, audience, and context of text production. While there is a growing body of literature studying stereotype content in online text (see Sourcing Adjectives), the form and nature of this data is simply different from data generated by participants in response to a direct prompt. In online text data, many analytic approaches emphasize co-occurrence: if two terms consistently co-occur with each other, or co-occur similarly with other words, this is worth considering and interpreting (this is the basis of topic analysis and word embedding; see Alghamdi & Alfalqi, 2015; Bhatia & Bhatia, 2020). For instance, Bhatia and Bhatia looked at co-occurrence patterns between definitionally gendered terms (e.g., she, her, mother, girl, his, him, father, boy) and stereotypically gendered terms (e.g., affectionate, imaginative, aggressive, analytical) in large corpora from Google News, Google Books, the Corpus of Historical American English, and the New York Times. Across the time period from 1910 to 2010, they found that the *stereotypically* gendered terms (especially feminine ones) became less-differentially-associated with the *definitionally* gendered terms. They interpreted this pattern to mean that – at least in this data – linguistic stereotyping has diminished (though not vanished).

Looking at co-occurrence in text relies on the assumption that text reflects the (potentially implicit) mental associations of the author (such as, for instance, when journalists use more aggressive language to describe outgroup members committing crimes than ingroup members; Vaes et al., 2019). This not-particularly-radical assumption is the basis of the approach I took to extracting adjectives from the online sources. However, it is no more radical to note that some co-occurrence is unlikely to reflect stereotype content. This was partly noted by omitting adjectives that seemed likely to represent other entities in the candidate sentences (e.g., *fiscal*

policy in connection with Democrats and Republicans; see Table 5). However, in other cases it is less clear. For instance, *racist* co-occurs frequently with all the racial-ethnic groups in the online data, but only the White People target group in the free-response data. Based on simple co-occurrence, it is unclear whether all of the racial-ethnic groups are being characterized as racist, or if, instead, some of the racial-ethnic groups are being discussed in the context of racist behavior. The phrase "J.D. is a racist Black person" and the phrase "J.D. is racist against Black people" both index an association between *racist* and *Black people*. However, the nature of that association differs meaningfully across phrases, and that distinction is lost in this method of sourcing adjectives from online content. In contrast, the lab data is easier to interpret because participants were explicitly asked to characterize target groups with adjectives – we know that each term is being used to describe the target group, rather than frequently occurring near to the target group.

Beyond the conceptual differences between co-occurrence and description, the text sources differ in their authors, audiences, and context. Each lab participant was a Northwestern University undergraduate student who completed the study for class credit, and generated responses for a researcher. They did this on their own, in a deliberative process, without explicit interaction partners. People represented in the Reddit and Twitter data include a much wider array and number of people (though, again, these are neither random nor representative samples of the US population, see Sourcing Adjectives) spontaneously generating text for a wide variety of reasons. A critical feature of the online data is that both Twitter and Reddit data are online social spaces, designed for conversation and interaction with (anonymous or identified) other people. The people in these data sets are not speaking to the researcher, but to each other.

This difference in population and motives may explain the adjectives associated with *Transgender People* and *Gender Non-Binary People*. In the free-response data, participants mostly emphasized the bravery and uniqueness of these groups – the two are described quite similarly. In contrast, the online data shows that overwhelmingly negative and othering terms co-occurred with *Transgender People*, and overwhelmingly positive and inclusive terms co-occurred with *Gender Non-Binary People*. Social desirability is likely one component of this, with students wanting to describe the groups positively in order to reflect their view of themselves as tolerant. In the online data, a potential explanation of the valence difference between transgender-associated terms and nonbinary-associated terms is in who is speaking about them. Transgender people are currently the subject of poisonous social debate, while Gender Non-Binary people are less well-recognized. This may mean that the online adjectives associated with transgender people reflect negative outgroup stereotypes about this group, while the online adjectives about gender non-binary people represent ingroup members supporting each other.

Each adjective source has advantages and disadvantages. The lab-based data is very clear to interpret and consistent with prior literature, but only includes adjectives participants were willing to explicitly associate with the target group and represents a very small and homogeneous section of the US population. The online data sources include data from a more heterogeneous set of people and can catch less-explicit group-term associations, but index group-term association in a rather blunt manner. By combining these sources, I hope to leverage their relative advantages proceeding into the norming and validation process. This was especially important because people completing adjective checklists may be willing to endorse a trait-group association (particularly a negative one) that they would not have personally generated.

The terms for Secretaries make clear that the social media data gathered for that group focused on international relations (e.g., Secretary of State), and so this target group is excluded from further consideration. As some adjectives are repeated between target groups, this provided a list of 383 unique adjectives. To make the list manageable for norming participants, I removed synonyms (e.g., *awesome* and *amazing*), ending up with 213 terms.

Gathering Norming Data

After isolating the list of adjectives associated with 52 groups, the next step was to gather norming data. This serves three purposes: (1) to shorten the overall adjective list while guided by data, (2) to provide norming data to other researchers about the association between a given adjective and a given target group, and (3) to demonstrate the methodological suggestions made in Moving Forward with the Adjective Checklist. This norming process was conducted in Amazon Mechanical Turk to enable efficient data gathering in a population that is frequently recruited to psychology studies (Bohannon, 2016; Paolacci & Chandler, 2014).

Group-Rating Participants. 324 participants were recruited from Amazon Mechanical Turk. 16 were excluded (12 failed the attention check and 4 did not indicate they took the study seriously), for a final N of 308 (M_{age} = 38.32, SD_{age} = 9.93; 61.36% men, 38.64% women; 78.25% White, 10.06% Black, 7.79% Asian, 5.84% Latinx, 1.30% Native American, 0.32% Pacific Islander; $M_{duration}$ = 18 minutes, $SD_{duration}$ = 10 minutes).

Group-Rating Method. After passing the screener (see Appendix D: Online Screening), participants completed an online consent form. Each participant completed the adjective checklist for eight of the fifty-two possible social groups (each group was evaluated by 32 to 53

participants)[1], and was randomly assigned to one of ten random orders of the terms in the adjective checklist. For each target group, the participant chose all the adjectives they thought were associated with the group (see Appendix E: Adjective Checklist Measure Instructions). The participant then completed a feeling thermometer to indicate their attitude towards "people in general" and each target group (Liu & Wang, 2015). Next, participants indicated whether they were a member of any of the eight groups ("Are you a member of any of these groups?") and, if so, the identity centrality of each group in which they were a member. Centrality was measured on a 0-100 slider with a single item drawn from existing research ("Being a [group member] is an important part of how I see myself"; Leach et al., 2008). Finally, participants reported their race/ethnicity, political ideology, sexual orientation, and education (gender and age were reported in the screener). There was an attention check embedded in the demographics section, and after completing their demographics participants completed the seriousness check (see Appendix F: Attention Check and Seriousness Check).

List Coverage

Only 9.09% of participants opted to include extra adjectives about one or more target groups. This varied some by group (most common for *Jewish People*, 9.30% of participants), but is low overall, suggesting the 213-adjective list had good coverage for the 52 groups.

---

[1] Because of a coding issue within Qualtrics, the adjective checklist task about the group *Daughters* originally only allowed participants to select one adjective. This data was excluded from analysis. Subsequently, to make sure that *Daughters* was evaluated by enough participants, all participants did the measures for *Daughters* as well as eight other randomly-selected groups. As it is unlikely that order of participation would strongly affect stereotype content for *Daughters*, I proceeded with other analyses as planned.

Data-Driven List Reduction

A sequence of steps was conducted to reduce the total length of the list while retaining its coverage; based on prior use of the measure, I aimed to reduce the list to 100 or fewer terms. First, I defined the strength of association between each adjective and each group as the percentage of participants who chose that adjective while evaluating that group, noted *A*. For example, 39.62% of participants who completed the adjective checklist for Hindu people chose the adjective *cultural* ($A_{hindu\text{-}cultural} = 0.3962$), while only 2.13% of those evaluating Brothers did so ($A_{brothers\text{-}cultural} = 0.0213$). Traits which were not at least 15% associated with at least one group were omitted from further analysis, as they did not contribute substantially to coverage. This excluded 31 adjectives from further consideration: *absent, amateur, anonymous, bad, boring, cheap, confused, creepy, dark, distant, divorced, fake, glutenfree, hurt, illegal, late, mad, magic, medical, mental, opposed, pagan, published, quiet, right, sorry, spicy, stupid, terrorist, ugly,* and *weird*.

This leaves a list of 182 terms, from which I aimed to choose the best subset. How can the best subset be recognized and selected? I first approach the question of recognition, operationalizing what it means for some subset of adjectives to be better than another subset for this measure. I then describe the process of selection I used to arrive at a final list of adjectives in the measure.

Recognizing Better Subsets. In order for the adjective checklist to function well across target groups, it should be similarly relevant to each target group. This is critical to avoid a situation as with the Katz & Braly list, where the terms were well-suited to measure racial-ethnic stereotypes but may not be well-suited to measure stereotypes about other groups. For a given target group, I can define the distribution of associative strength to some set of adjectives (the

distribution of *A* for the group, based on a specific set of adjectives). Figure 4 shows the distribution of *A* for the groups Heterosexual People, Jewish People, and Women, based on the remaining 182 adjectives. We can see that the distributions have different shapes: Heterosexual People are weakly-associated with more adjectives, while Jewish people and Women have more adjectives with which they are strongly associated. From the figure, we can see that *A* is more similarly distributed for Women and Jewish people than Heterosexual people.

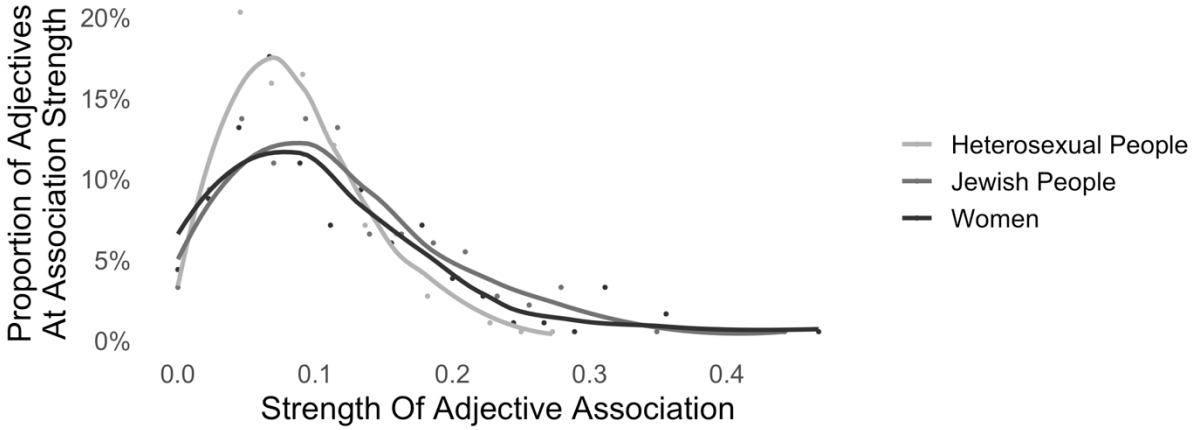

Figure 4. The distribution of associative strength between the 182 retained adjectives for the target groups Heterosexual people, Women, and Jewish people. Dots represent proportion of 182 adjectives at a particular value of A for target group, lines represent approximation of the distribution.

To measure the similarity between two distributions of associative strength, I relied on the Kolmgorov-Smirnov statistic, a non-parametric test that compares two empirical cumulative distribution functions (Goodman, 1954). It does this by calculating the distance metric *D*, the largest divergence between the two cumulative distributions. *D* ranges between 0 and 1, with smaller values indicating more-similar distributions. For Jewish People and Heterosexual People across the182 adjectives, $D = 0.2802$. For Jewish People and Women, $D = 0.1758$. For any two target groups, *D* quantifies the extent to which they show similar patterns of associative strength with a set of adjectives. There are 1326 pairwise combinations of the 52 target groups. For a given set of adjectives, I defined: (1) the average distribution similarity is the mean of *D* across

the 1326 possible group pairs, noted M_D, and (2) the spread of distribution similarity is the standard deviation of D, noted SD_D. A good set of adjectives for this measure will have low M_D and low SD_D, showing that the target groups are similarly associated with the set of adjectives. If both metrics are low, it means that the measure can be used with similar efficacy to measure stereotypes across all of the target groups. If both metrics are high, it means that the measure is much better at measuring stereotypes for some target groups than other target groups.

Along with being *equivalently* relevant to the different target groups (as indexed by M_D and SD_D), a good set of adjectives should be *relevant* to the different target groups in the first place. In general, the strength of association between the set of adjectives and a target group should be high (calculated as the mean of A, noted S_A). Looking again at Figure 4, we can see that the set of 182 adjectives is more associated with Jewish People and Women than Heterosexual People ($S_{A\text{-}jewish} = 0.1183$, $S_{A\text{-}women} = 0.1137$, $S_{A\text{-}heterosexual} = 0.0894$). For a given set of adjectives, I defined the average associative strength of the set with the target groups as the mean of S_A across target groups, noted M_S. A good set of adjectives for this measure will have higher M_S, showing that the adjectives are useful for describing the target groups.

Finally, a good set of adjectives should not be too long. Adjective checklist measures have ranged in length, with common ones having 80-100 terms (see Measurement Variation). Though longer measures can be more comprehensive, they place more burden on participants and are less efficient. Participant fatigue may mean, for instance, that not every adjective is carefully considered. In concert with other optimization variables, a good measure will have a lower number of adjectives (noted L).

These four optimization variables – M_D, SD_D, M_S, and L – answer the first part of the optimization question, allowing us to recognize better and worse subsets of terms for use in the

measure. They can be defined for every possible subset of the 182 adjectives. However, it is computationally intractable to calculate them all. Even if I only considered combinations of adjectives with L from 80 and 100, that is more than 10^{54} possible combinations[2].

Selecting a Subset. As it is intractable to know which of all of the possible subsets of adjectives are best, selecting a subset requires a process for going through some but not all of those subsets. As a first step, I selected the two adjectives that are most strongly associated with each of the target groups, marking them for retention in the final list. Including these 54 adjectives is intended to ensure that M_S does not drop too low. As a second step, I restricted L to values between 60 and 100, deciding based on the literature that these are appropriate bounds for the possible length of the list. Together, these meant that I needed to pick a subset of 6-46 adjectives out of 128. M_D, SD_D, M_S, and L for a given subset are defined including the subset and the 54 retained adjectives.

To understand how the four optimization variables relate to each other, I generated 100 random subsets of adjectives for L from 60 to 100 and calculated the optimizer variables for each. In Figure 5, we see that L is monotonically associated with M_S; as more adjectives are included in the measure, the overall strength of association between the measure and the target groups decreases. There is a strong linear trend ($\beta = -2.86$), and a smaller curvilinear one ($\beta = 1.91$), such that M_S decreases more quickly at lower values of L and more slowly at higher values

[2] Best subset selection in regression is a related statistical problem, where the goal is to choose the best subset of variables to serve as predictors for some dependent variable. A variety of algorithmic solutions have been proposed for this problem (e.g., Hocking & Leslie, 1967; Xu & Zhang, 2001), but are centrally concerned about the correlation between different variables (the predictor variables and the outcome variables), not with the distribution of attributes over the retained or omitted variables.

of *L*. In Figure 6, we see that M_D and SD_D are positively related across different values of *L,* such that smaller values of *L* are worse for both metrics.

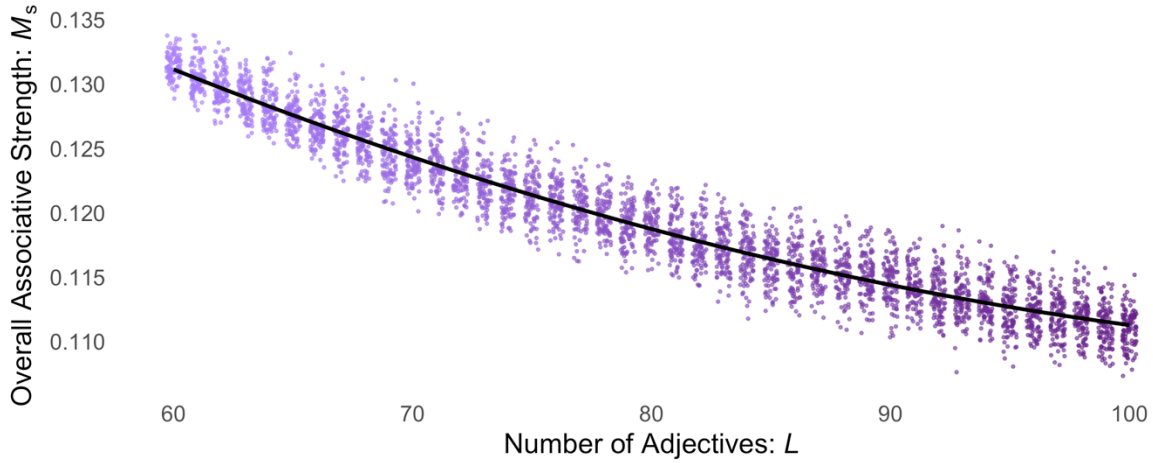

Figure 5. M_S is monotonically associated with L. Each data point represents a single random combination of adjectives at a particular value of L. Line shows a best-fit curvilinear model of the relationship between the two outcome metrics.

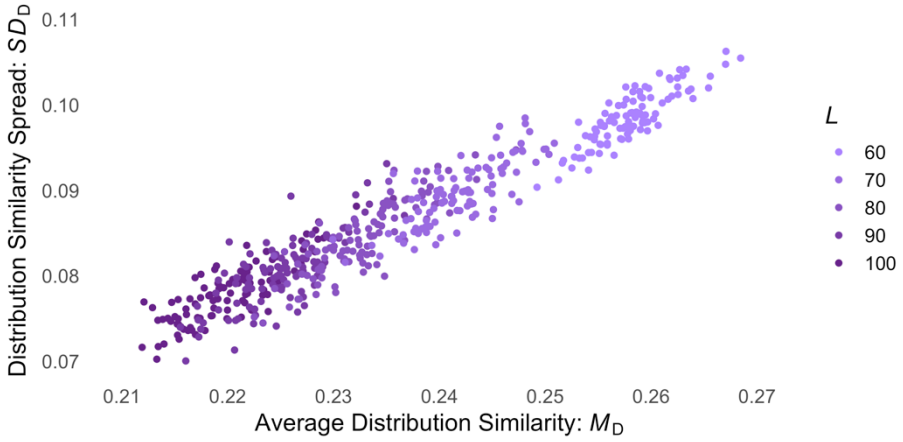

Figure 6. M_D and SD_D are positively related, and both are worse at lower levels of L.

These patterns mean that instead of simultaneously addressing all four optimizers (M_D, SD_D, M_S, and *L*) as holding unique information about the quality of a particular set of words, I can think of M_D and SD_D as very similar indicators and M_S and *L* as very similar indicators. I therefore decided to proceed with optimization focused on M_D across different values of *L*, while assessing how this optimization process influenced the value of SD_D and M_S.

To do this, I used a backward elimination algorithm. This starts with the largest number of adjectives (L = 128 + 54 retained adjectives) and defines every possible combination of size L – 1. For the first step, this selects all 128 possible combinations of 127 words out of the 128 that may be included. The combination with the lowest M_D is selected for continuation. At the next step, another word is removed in the same fashion. This step-wise process, eliminating one adjective at a time, continues until L = 60 (54 retained adjectives + 6 selected based on stepwise M_D). This approach is path-dependent; the removal of an adjective may influence which adjectives are removed at later steps. As a point of comparison, for each value of L, I also selected the adjectives that had the highest A across the 52 groups; this approach is path-independent.

Figure 7 displays the optimization variables for 80 adjective subsets: one per choice process at each value of L from 60 to 100. As L increases, the backwards elimination algorithm provides better subsets than highest association in terms of M_D and SD_D. In contrast, the highest association subsets perform better on M_S. Around 70 words, the three optimization metrics stabilize. Noting this, and keeping in mind that a shorter measure is easier for researchers and participants, I selected the terms that were picked out by the stepwise process for inclusion in a 70-term measure; all terms displayed in Table 8.

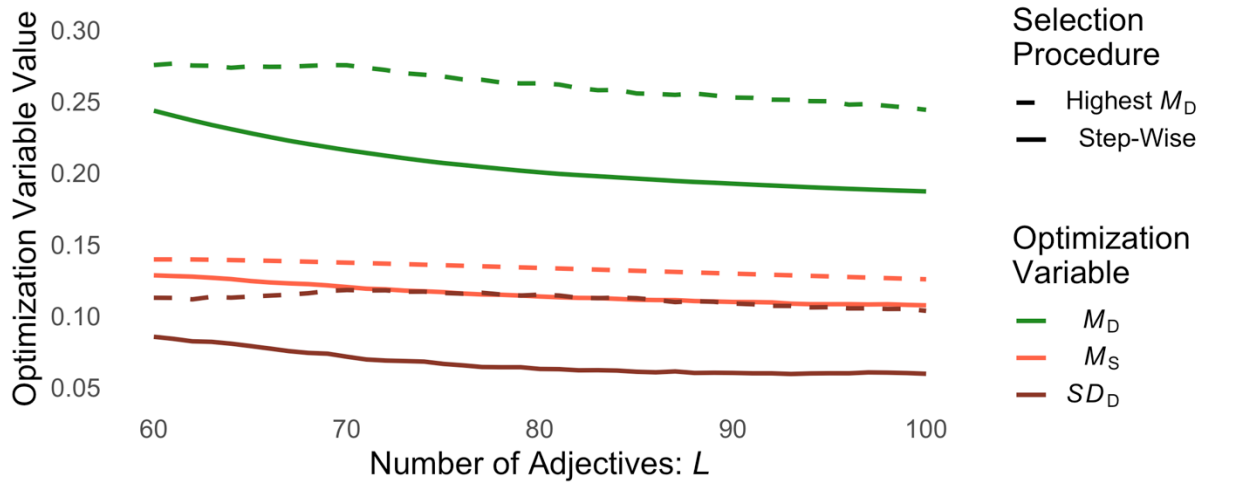

Figure 7. Results of optimizing with step-wise algorithm compared to picking highest-associated subset for every value from L = 60 to L = 100.

Table 8. Terms retained for use in the 70-adjective checklist measure

adult	curious	feminine	important	misunderstood	religious	stressed
aggressive	dedicated	feminist	innocent	motivated	respectful	struggling
alone	democratic	friendly	interesting	nationalist	responsible	successful
anxious	difficult	gentle	lazy	natural	ruling	thoughtful
athletic	dirty	handsome	loving	political	scared	traditional
available	disabled	happy	low	powerful	selfish	trustworthy
brave	emotional	hardworking	loyal	pretty	sexual	uncomfortable
comfortable	employed	healthy	lucky	privileged	sick	unfortunate
common	expressive	holy	married	protective	skilled	unique
cultural	fashionable	homeless	masculine	proud	smart	young

Validation

To validate this measure, I tested whether participants chose more positive adjectives to describe groups they felt more positively about (as indexed by the feeling thermometer; for a similar approach see Parish et al., 1976). I further tested whether participants chose more positive adjectives to describe groups in which they were a member, and whether this was especially pronounced for groups that were very important to their sense of self. In each case, I tested whether this pattern appeared using the full set of 213 adjectives, and whether it remained

using the subset of 70 adjectives. To pursue this validation strategy, I first gathered norming data about the valence of each adjective when applied to people.

Valence-Rating Participants. 103 participants were recruited from Amazon Mechanical Turk. Two were excluded for failing the attention check (all said they took the study seriously, and all had variability in their scale responses), for a final N of 101 (M_{age} = 38.09, SD_{age} = 11.36; 70.30% men, 29.70% women; 66.33% White, 16.83% Black, 15.84% Latinx, 4.95% Asian; $M_{duration}$ = 11 minutes, $SD_{duration}$ = 8 minutes).

Valence-Rating Method. After passing the screener (see Appendix D: Online Screening), participants completed an online consent form. Each participant then rated the valence of 107 of the adjectives, on a 1-7 scale from Very Negative to Very Positive ("How positive or negative is this term when it's used to describe a person?"). Adjectives were randomly selected from the full set of 213 (each adjective was rated by 37 to 63 participants). Finally, participants reported their race/ethnicity, political ideology, sexual orientation, and education (gender and age were reported in the screener). There was an attention check embedded in the demographics section, and after completing their demographics participants completed the seriousness check (see Appendix F: Attention Check and Seriousness Check). The mean of participants' ratings for a given adjective was defined as adjective's normative valence. On average, the adjectives were significantly more positive than the scale midpoint of 4 ($M_{valence}$ = 4.54, $t(212) = 6.914$, $p < 0.0001$; see mean and standard deviation of valence for each adjective in Appendix G: Valence of the 213 Adjectives).

Results. All analyses are conducted using a multi-level modeling approach, with a random intercept for participant and for target group (to control for repeated measures; Raudenbush & Bryk, 2002). The mean valence of all individual adjectives chosen by a

participant to describe a certain group was defined as the valence of their stereotype content about that group. I used the participant's feeling thermometer score about the group to predict the valence of their stereotype content (see Figure 8). When including information from all adjectives used in the norming process, feeling thermometer ratings did predict stereotype valence ($t(2349) = 20.29$, $p < 0.0001$). When including information only from adjectives retained for the shorter list, feeling thermometer ratings still predicted stereotype valence ($t(2106) = 15.94$, $p < 0.0001$). The similar effects suggest that the shortened list retains the predictive utility of the task with the full set of terms.

Figure 8. Feeling thermometer ratings predict stereotype valence on the full set of 213 adjectives and the selected subset of 70 adjectives. Lines show estimated marginal means, and shading shows standard error.

Next, I tested whether participants expressed more positive stereotypes about their ingroups than their outgroups, and whether this was especially the case for high-identity-centrality ingroups. Using a multi-level model with a random intercept for participant and target group, ingroups were stereotyped in more positive terms than outgroups for both the full 213 adjectives ($t(2358) = 5.281$, $p < 0.0001$) and the subset of 70 adjectives ($t(2326) = 4.028$, $p < 0.0001$; see Figure 9), though this effect is smaller than the relationship between stereotype valence and feeling thermometer ratings. To follow up on this, I focus on only the data in which

participants chose adjectives to describe an ingroup. Using a multi-level model with a random intercept for participant and target group, identity centrality predicted positively stereotyping the ingroup based on both the full set and final set of adjectives (full set: $t(633) = 6.614$, $p < 0.0001$; subset: $t(588) = 4.711$, $p < 0.0001$; see Figure 9). These results again suggest that the subset of adjectives retains the predictive utility of the full list.

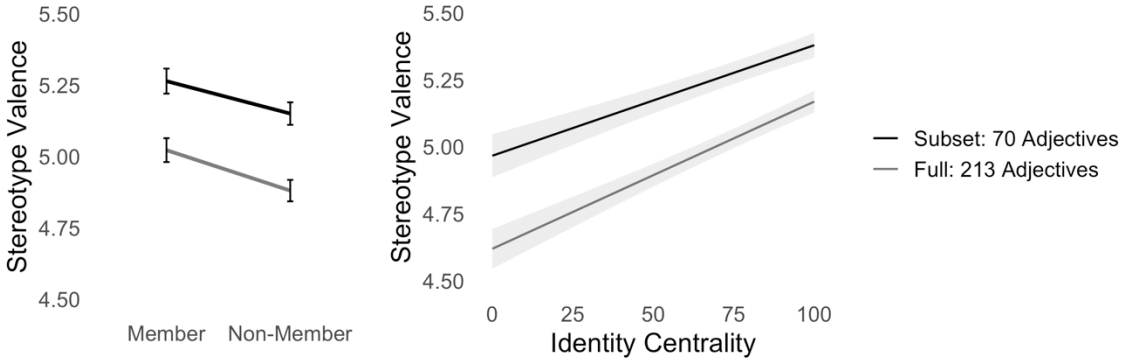

Figure 9. Participants chose more positive stereotype attributes for their ingroups than outgroups, and this was even more the case for memberships more central to their sense of self.

Task Analysis

In Moving Forward with the Adjective Checklist, I suggested that researchers using the adjective checklist take should take a series of analytic steps. To demonstrate these, I will present analyses asking: are lesbian women or gay men stereotyped more similarly to homosexual people overall? All analyses are conducted with the adjective checklist results about Homosexual People, Lesbian Women, and Gay Men, using the 70 terms chosen to include in the final adjective checklist ($N = 128$).

If gay men are stereotyped more similarly to homosexual people than are lesbian women, this will support the argument that many stereotypes are characterized by androcentrism (e.g., Bailey et al., 2019), and could provide a valuable framework for understanding how people in the US today approach conversations about gay rights. From a historical perspective, this

analysis contributes to a conversation about who is socially foregrounded in the LGBTQ+ rights movement. This is intended to be a preliminary investigation into this question based solely on adjective checklist data, not a conclusive one.

Stereotype Similarity Across Participants

I first ask whether, on average, gay men or lesbian women are stereotyped more similarly to homosexual people overall. As an exploratory step common to use of the adjective checklist measure, I consider the most-frequently chosen adjectives for each group (see Table 9). From this, we can see the groups share meaningful stereotype content: *loving* and *sexual* appear for all three. We also see that none of the adjectives are chosen by a majority of participants; this may represent a lack of cultural-level stereotype consensus, but it also may be an artifact of having presented participants with such a long list of terms.

Table 9. Five most frequent adjectives chosen for each group, and the percentage of participants who chose them.

Homosexual People		**Gay Men**		**Lesbian Women**	
expressive	22.00%	expressive	27.66%	feminist	31.25%
proud	22.00%	sexual	25.53%	sexual	25.00%
loving	20.00%	cultural	21.28%	interesting	22.92%
sexual	20.00%	fashionable	21.28%	friendly	20.83%
cultural	18.00%	feminine	21.28%	loving	20.83%
interesting	18.00%	loving	21.28%	feminine	18.75%
selfish	18.00%	misunderstood	21.28%	responsible	16.67%
healthy	16.00%	democratic	19.15%	successful	16.67%
handsome	14.00%	motivated	19.15%	trustworthy	16.67%
hardworking	14.00%	protective	19.15%	brave	14.58%
important	14.00%	responsible	19.15%	democratic	14.58%
respectful	14.00%	struggling	19.15%	proud	14.58%
scared	14.00%			thoughtful	14.58%
				young	14.58%

To statistically assess whether gay men or lesbian women are stereotyped more similarly to homosexual people overall, I calculate each adjective's strength of association with each of

the three groups (A, as in Data-Driven List Reduction). Then, across the 70 adjectives, I test whether $A_{homosexual}$ predicts A_{gaymen}, and whether $A_{homosexual}$ predicts $A_{lesbianwomen}$. While the adjective's association with homosexual people predicts its association with gay men ($r(68) = 0.52$, $p < 0.0001$), this pattern does not emerge for the adjective's association with lesbian women ($r(68) = 0.16$, $p = 0.18$; see Figure 10). This pattern occurs despite a correlation between $A_{lesbianwomen}$ and A_{gaymen} ($r(68) = 0.31$, $p = 0.01$), suggesting that the stereotype of homosexual people is related to the unique content of the gay man stereotype, rather than the content it shares with the lesbian women stereotype.

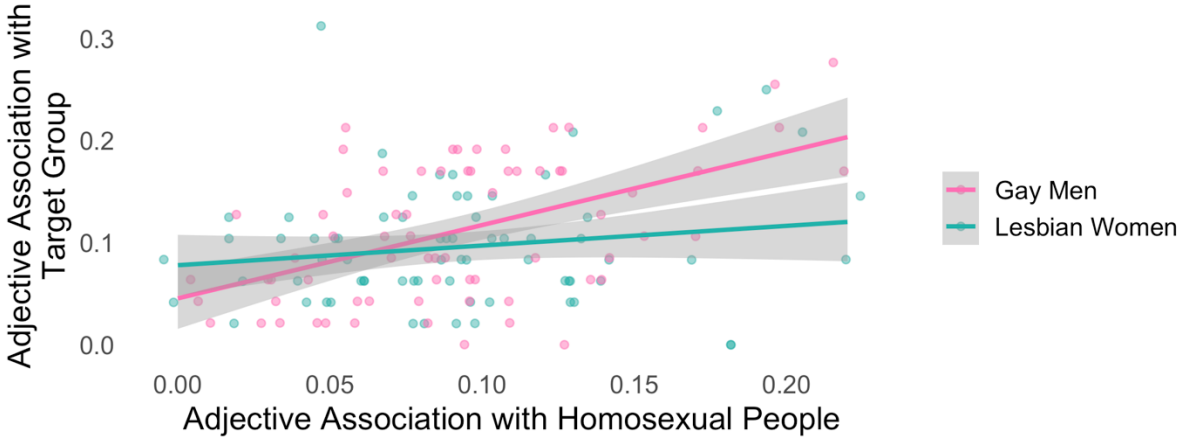

Figure 10. The extent to which an adjective is stereotypically associated with homosexual people predicts its stereotypical association with gay men, but not with lesbian women.

Number of Adjectives Chosen from List

I earlier suggested that researchers allow participants to choose any number of adjectives, rather than restricting it to five or ten. Figure 11 shows that participants chose more terms to describe gay men and lesbian women than homosexual people, though these differences are not significant (ps > 0.3). The majority of participants chose more than five adjectives to describe gay men or lesbian women (gay men: 58.70%; lesbian women: 52.27%), while a large plurality

did so to describe homosexual people (37.50%). This supports the idea that restricting to five adjectives curbs participants' reporting of stereotype content.

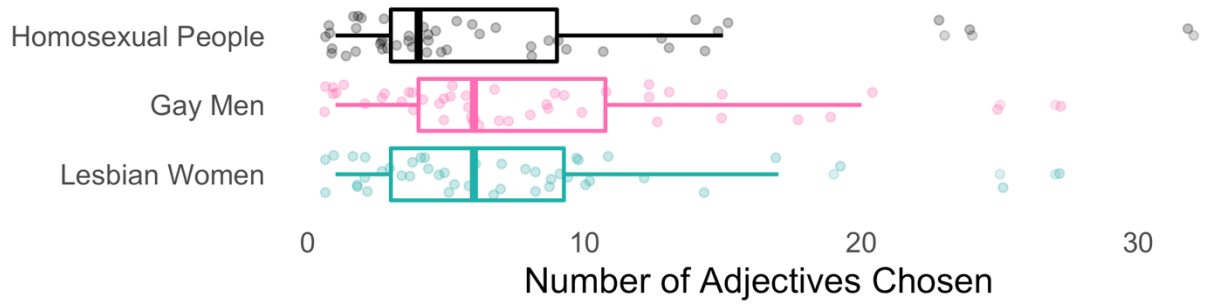

Figure 11. The number of adjectives chosen by participants to describe each group. Many participants chose more than five adjectives.

Term Co-Occurrence

I earlier suggested that it is valuable to consider how the co-occurrence of different terms varies by different target group. Though many analyses of this measure aggregate across the terms to define some aspect of stereotype content (e.g., valence), it is also worthwhile to consider whether participants use the terms differently when evaluating different groups. I look at whether *powerful* and *emotional* are similarly related across the three target groups, using a chi-squared test of choice of each term for each target group. For lesbian women, but not gay men or homosexual people in general, choosing *emotional* strongly predicts choosing *powerful* (lesbian women: $\chi^2(1) = 12.685$, $p = 0.0004$; gay men: $\chi^2(1) < 0.001$, $p > 0.8$; homosexual people: $\chi^2(1) = 0.030$, $p > 0.8$; see Figure 12). In this particular instance – the extent to which participants perceived a group's power as connected to its emotionality – stereotypes of lesbian women again differed from those of gay men and homosexual people.

Figure 12. The extent to which emotional and powerful are associated depends on target group. When evaluating lesbian women, choosing emotional as an adjective strongly predicted choosing powerful as an adjective. When evaluating gay men or homosexual people overall, this was not the case.

Discussion

In the first part of this chapter, I argued for the importance of measuring stereotype content, described the advantages and pitfalls of the adjective checklist measure, and conducted a narrative review of the measure's use since 1933. In the second part, I gathered new adjectives from online and experimental sources, gathered norming data from participants on Amazon Mechanical Turk, and used a data-driven process to select a set of 70 adjectives for use in an adjective checklist measure that is applicable across a wide array of social groups. I validated this measure by showing that the valence of chosen adjectives was predicted by feeling thermometer ratings, and that participants chose more-positive-adjectives to describe their ingroups and especially ingroups that felt very important to them. Finally, I demonstrated several analyses using data from those 70 terms, showing an androcentric bias in the stereotype of homosexual people overall, which more closely resembled that of gay men than lesbian women. In the process, I provided a series of resources for ongoing research: the measure itself, norming data about the association of each adjective with each target group, and norming data about the valence of each adjective.

How New Is This Adjective Checklist?

What is the added value of this adjective checklist? I first consider how much of an update it is with respect to the well-established Katz & Braly list, with semantic distance as calculated by SemDis (Beaty & Johnson, 2021). This metric ranges from 0 to 2; higher scores show that the two texts are more distantly related. The Devine and Elliot (1995) update on the Katz and Braly (1933) has a distance of 0.02. Galinsky et al.'s 2013 update of the 1995 list has a distance of 0.01. These updates, which involved adding relatively small sets of adjectives onto a relatively long existing list, only slightly changed the measure's semantic content. In contrast, moving from the 2013 list to the 70-term list proposed here represents a semantic jump of 0.24 – it is a much bigger jump than prior updates have been.

It is also worth considering the extent of the overlap between this adjective list and common dimensional measures of stereotype content – the Stereotype Content Model (SCM; Fiske et al., 2007) and the Agency-Beliefs-Communion model (ABC model; Koch et al., 2016). In both models, agency and communion represent important facets of stereotype content (sometimes with sub-facets as well; Abele et al., 2016), while the ABC model includes political beliefs (conservative – progressive) as a third facet (for more on the integration of these and other dimensional models of stereotype content, see Abele et al., 2021). Some adjectives in the list of 70 seem well-explained by these three dimensions. *Emotional*, *friendly*, and *loving* are closely connected to communion, while *aggressive*, *hardworking*, and *responsible* likely represent agency and *feminist*, *nationalist*, and *political* could be connected to political beliefs. Other terms fit less neatly into these three dimensions – for instance, *cultural*, *misunderstood*, and *uncomfortable*. While dimensional measures often emphasize dispositional adjectives (e.g., *selfish*), the checklist also contains situational ones (e.g., *homeless*), expanding the range of

potential attributes to include group-linked circumstances as well as dispositions. Thus, while the relationship between this adjective checklist and dimensional approaches to stereotype content is an empirical question for further research, the checklist seems to share ground with both the SCM and ABC while also accommodating a wider spectrum of content.

Optimization and Tradeoffs

Any process of optimization requires tradeoffs, and this is certainly the case with the measure development process in this paper – no such process is theoretically neutral. In this case, I have consistently emphasized central tendencies and majority-selection. When gathering adjectives from Reddit, Twitter, and the Free Response data, I chose the most-common adjectives to use in the norming process. When using the norming data, I defined *A* as the raw proportion of participants choosing an adjective to describe a group. In both cases, I focused on raw frequency rather than investigating how subgroups of individuals might stereotype the target groups differently. This means that, as with the adjective checklist in general, this measure development process collapsed across potentially rich linguistic diversity and variability (Duijker & Frijda, 1960). It also relied on participants from Amazon Mechanical Turk, with a majority of White participants and a majority of participants self-identifying as men. In many ways, this weakness is the weakness of any standardized measure – pursuing the goal of functioning as well as possible for the largest possible proportion of a population, I averaged across heterogeneity. Future work may explore how the total list should be adapted in different populations or sub-populations.

I also focused on developing a measure that was equally suitable to an array of groups across several social dimensions. If I used the same optimization process but aimed to develop the best measure for a smaller set of target groups (e.g., only gender groups, or only kinship and

wealth class groups), that could produce a different list. As all data and code for the optimization process are freely accessible, other scholars may reproduce the process to define the most-suitable subset of terms for their particular project.

Finally, the choice of optimization metrics is itself a theoretical decision. Having decided that it was theoretically important for the list of adjectives to be similarly associated with each target group, I optimized on M_D. In another case, it might be advantageous to pick the terms that best *distinguish* between target groups. For this, one would focus not on the distribution of A, but rather on the difference between A_{Group1} and A_{Group2}, selecting adjectives which maximize this difference. In yet another case, it might be important to balance the valence of terms stereotypically associated with different groups. By incorporating the valence norming data with the values of A, one could minimize expected differences in stereotype valence across the groups. In this case, I focused on the list's association with each target group in order to develop a measure that could be flexibly used across different areas of stereotyping research.

Corpora for Future Use

Beyond the measure itself, the corpus I constructed in this chapter could be a useful resource for further research. Online text, like all qualitative data, is rich and deep. Though there are a variety of public corpuses (including, for example, the Pushshift Reddit dataset; Baumgartner et al., 2020), the one in this chapter has been constructed to suit the needs of researchers who study stereotyping, prejudice, and bias. The affordances of online data – in particular, that it is generated during social interaction – mean that it could be especially useful to study sub-communities of consensus across which the stereotype of a given group varies (e.g., stereotypes of Democrats in conservative or liberal subreddits), as well as the communication and negotiation of stereotype content. Other analyses could take advantage of the time-course

represented in the data, testing whether the stereotype content associated with a particular group changed in response to particular social events (e.g., did the stereotype content associated with Rich People shift during the college admissions scandal? Did gender stereotypes shift during the US women's undefeated soccer season?). Beyond the association between target groups and adjectives, this data could offer insight into linguistic biases more broadly, such as the linguistic intergroup bias (Maass et al., 1989) or the tendency to use culture to explain the behavior of racial minority group members, but not racial majority group members (Causadias et al., 2018).

Recommendations for Future Use

In closing, I will make several recommendations for further use of this measure. Methodologically, researchers should (a) allow participants to add words to the list, using the proportion of participants who did so as an indication of the measure's efficacy for their research project, (b) allow participants to select any number of adjectives, using the number of adjectives chosen as a proxy for the depth of the person's stereotype, and (c) articulate their reason for using instructions asking about participants' personal beliefs or perceptions of cultural stereotypes. Analytically, researchers should (a) incorporate historical analysis, (b) share norming data, and (c) consider how adjective meaning may shift based on a particular target group. Beyond these specific suggestions, I also recommend that this measure is used alongside others, especially measures of stereotype content that allow for high participant agency such as dialect analysis and free-response.

Chapter II: An Experimental Test of Sociocultural Essentialism in Social Categories

In this chapter, I argue that existing research on social essentialism has overemphasized essentialist beliefs founded in the perception of shared genes and suggest that shared experience may serve as another basis for social essentialism. I then propose three specific hypotheses about essentialist reasoning about social categories perceived to be non-biological (e.g., social class). I used a norming process to identify appropriate target groups, and conducted two studies testing the possibility that shared social experience can serve as a basis for social essentialism.

Psychologically essentialized categories are understood as having an underlying causal essence (Medin & Ortony, 1989). The relevant question is not whether the essences *truly exist*, but rather whether people *act as though* and *believe* they do. Within an individual being, this essence is perceived to be immutable: if Rex is a tiger, he will be a tiger for his entire life. This essence is believed to cause the common attributes of the category (e.g., Rex has stripes and a tail because he is a tiger), constrain those attributes (e.g., Rex should not have a mane because he is a tiger), and provide information about the common life trajectory of category members (e.g., a tiger cub will grow into a hunter because it is a tiger, Medin & Ortony, 1989).

Inductive potential is a central affordance of essentialized categories – because essences causally produce category-linked attributes, learning about stable attributes of one member of a category provides useful information about other category members (e.g., if one pig has a heart that flattens as it sleeps, other pigs probably do as well). However, this is limited to central attributes (Medin & Ortony, 1989); accidental or incidental attributes are not generalized (e.g., if a tiger is living in a zoo, that does not mean all tigers live in zoos; Gelman, 2003). The inductive utility of essentialized categories serves valuable purposes, but can also lead to over-generalization and failure to account for change over time or context. For example, essentialist

representations of animals and chemicals can lead to misunderstandings of evolutionary and chemical change (Gelman & Rhodes, 2012; Leslie, 2013).

Though essentialism appears consistently in representations of natural kind categories (e.g., animals), it is not a universal feature of category representation. It is used by some people more than others, and elicited by some task situations more than others (Kalish, 2002). It does not typify all categories, either: artifact categories tend not to be essentialized (Boyer, 1996). They are instead grounded in lay theories of function (e.g., an object is a chair because you can sit on it) or creator's intention (e.g., an object is a chair because someone made it for sitting; Bloom, 1996; Gelman, 2013; Putnam, 1975). While these lay theories satisfy the cognitive impulse to *causal* lay theories, they do so without assigning an essence to a category (Matan & Carey, 2001), though for a different perspective see (Gelman, 2013). Unlike essentialist lay theories, these allow for mutability in the category membership of a single entity: a hunk of metal may be used as a paperweight or a projectile, and plausibly has membership in multiple artifact categories. They also have reduced inductive potential: learning that one chair is made of wood does not necessarily tell you another chair is made of wood.

Social Categories

While it is more accurate to think of social categories as artifactual (i.e., defined based on how humans interact with them) than natural (i.e., culturally invariant), psychological essentialism in the social domain studies when, how, and why people *act as though* and *believe* that there is underlying reality to these social groupings (e.g., Gelman, 2003; Hirschfeld, 1996; Prentice & Miller, 2007; Rothbart & Taylor, 1992). Consider systems of racial stratification. Though many cultures sort people based on race, specific organizational schemas are culturally and historically variant: the number of categories, who falls into which category, what is

expected of category members, and the salience of a given category system are all flexible (e.g., Brewster & Padavic, 2000; Bronfenbrenner & Ceci, 1994; Joel et al., 2014; Nelkin & Lindee, 1995). However, many people *believe* that racial categories represent an underlying truth, and *act as though* members of the same racial group are fundamentally the same. Cultural input appears to tune a general cognitive tendency towards psychological essentialism of social categories; though patterns of essentialist beliefs are widespread, there is variability in who essentializes whom in which ways (Chalik et al., 2017; Davoodi et al., 2019; Rhodes & Gelman, 2009; Smyth et al., 2017).

Social essentialism has been widely studied (a Google Scholar search for "social essentialism psychology" returns about 123,000 results), and scholars have taken different approaches to its definition. I rely primarily on work in which essences are considered fixed aspects of a person along a given social dimension, *independent* of other dimensions (a man may be seen as fundamentally different from a woman, but each could be either rich or poor; e.g., Haslam et al., 2000; Hegarty, 2002)[3]. Other scholars focus on essentialism as *infrahumanization*, where the ingroup is assigned more human essence than the outgroup (e.g., Leyens et al., 2001, 2003), or as *determinism*, where the life outcomes of people in general are defined by some underlying essence that varies between individuals (e.g., Keller, 2005; Rangel & Keller, 2011). Implicitly, these approaches offer different solutions to human multidimensionality, wherein individual people can be categorized in multiple social groups at the same level of specificity

[3] There is a burgeoning area of research about the intersections and mutual dependence of social category representation (e.g., intersectional invisibility, Purdie-Vaughns & Eibach, 2008; the race-is-gendered effect; Johnson et al., 2012; the stereotypic association between wealth and race, Lei & Bodenhausen, 2017). As a matter of measurement and theory, the psychological essentialism literature has not addressed such dependencies of category representation.

(e.g., an individual can be simultaneously Asian, middle-class, and heterosexual). I rely primarily on research with the *independence* framework because this approach, like mine, focuses on beliefs about particular social groups and allows for inter-category comparison.

What *Is* the Essence?

Essentialist beliefs do not require explicit knowledge of the category's supposed essence; such knowledge may be inaccessible. Rather than a fully elucidated causal theory, people hold "placeholders" (Medin & Ortony, 1989). In general, beliefs about animal categories are assumed to be based in biological essences; studying causal placeholders has provided evidence of cultural variation in the specific biological essence. Consider a pig whose blood is switched out with the blood of a cow. Brazilian children, who grow up in a culture where blood is tightly linked to identity, are more likely to consider that animal a cow than American children, who do not grow up in such a culture (Sousa et al., 2002; a similar asymmetry was documented in Waxman et al., 2007 with Native American and majority culture children in the US). In both cultures, pigs and cows are biologically essentialized; the cultural variation occurs in defining whether blood or body holds the essence.

The bulk of work on essentialism in social categories has carried forward this focus on *biogenetic* essentialist beliefs (e.g., Heine et al., 2017); for a similar argument, see (Bailey et al., In Press). In some cases, this focus is explicit. In a very common experimental manipulation called the switched-at-birth task, participants learn of a baby born to parents of one social group (e.g., Black parents), who was immediately adopted by parents of another group (e.g., White parents). With no further information, participants are asked to imagine the baby as an adult and judge which group membership and group-linked attributes they will have (e.g., a Black or White adult with stereotypically Black attributes such as athleticism or stereotypically White

attributes such as intelligence). In the absence of individuating information about the person, participants must rely on their existing ideas about the social domain (e.g., race). Traditionally, responding that group membership and attributes are transmitted genetically from the birth parents is defined as essentialist; (e.g., Gil-White, 2001; Hirschfeld, 1995; Mahalingam, 1998; Rad & Ginges, 2018). Many scale measures of social category essentialism specifically refer to biological forces (e.g., Andreychik & Gill, 2015; Bastian & Haslam, 2006; Coleman & Hong, 2008; Keller, 2005; M. J. Williams & Eberhardt, 2008), or exclude non-biological items from their initial scale development (Ho et al., 2015)[4].

In other cases, this biogenetic focus is more implicit. For instance, researchers use scales which do not specifically reference biology in measuring essentialist beliefs (e.g., Haslam et al., 2000). However, the literature leans strongly towards studying beliefs about race and gender – two social dimensions about which laypeople in the US often hold biological beliefs. The suggestion that visual cues to category membership (e.g., skin color) are most likely to produce essentialized categories has been used to explain why ethnic categories are essentialized, even though a wide variety of visual cues exist in any situation (e.g., hair length, wearing a headscarf or valuables; cf. Gil-White, 2001). Yzerbyt and colleagues further suggest that "social observers may often react as if they were adopting some sort of 'biological framework' when in fact they are referring to influences such as education, religion, culture, climate, and so on" (2004, p. 80). Thus, responses to questions with ambiguous causal forces may be interpreted as biological even when they reflect a different belief.

---

[4] No and colleagues (2008) omit such items even as they acknowledge the possibility of non-biological essentialist belief.

This explicit and implicit focus on biological causes runs counter to the theoretical stance that causal placeholders can vary (Medin & Ortony, 1989), and evidence which suggests laypeople can have complex and multiply-causal mental representations of the world (e.g., Condit, 2019; Kanovsky, 2007; Martin & Parker, 1995; Ryazanov & Christenfeld, 2018). More recently, researchers have been considering the potential heterogeneity of essentialism, distinguishing essentialism based on the perception of shared biogenetic factors from essentialism based on the perception of shared values (Bailey et al., In Press; Newman & Knobe, 2019), shared childhood socialization (Rangel & Keller, 2011), and shared culture (Segev et al., 2012; Yalcinkaya et al., 2017). Rangel & Keller (2011) developed a measure focused on sociocultural human determinism (i.e., not about categories). This scale predicted belief in the inductive potential and group homogeneity afforded by social experiences (profession, neighborhood, parenting style of parents, working class background, aristocratic background, rich family), while belief in genetic determinism predicted perceived homogeneity and inductive potential of physical attributes (genes, brain structure, skin color, blood type, body height; Rangel & Keller, 2011). Perceptions of biogenetic and sociocultural causality can both support essentialism, but that the basis for shared inference differs. These studies point towards the possibility that the perception of *group-linked experiences* could form the foundation for essentialist beliefs just as much as the perception of *group-linked biogenetics* (e.g., genes, chromosomes).

The Perception of Shared Experience

The perception that members of a certain group share a lot of experiences (such that those experiences might function as an essence) is closely connected to the social group concept of *entitativity* (e.g., Yzerbyt, Judd, et al., 2004), or the degree to which a group is seen as a real and

coherent whole (Campbell, 1958). Shared social experience is most closely connected to *common fate*, a component of entitativity that describes the extent to which different entities experience the same things over time. Perception of common fate with other ingroup members can contribute to ingroup cohesion and solidarity among marginalized groups (Craig & Richeson, 2012; Dawson, 1994), and appears in measures of collective identity (Lowery et al., 2007; Sellers et al., 1997).

There is ongoing debate about the exact nature of the close relationship between entitativity and essentialism (Yzerbyt, Judd, et al., 2004). Some scholars in the US, China, and Northern Ireland have found that scale measures of essentialism break down into two factors: natural kind beliefs (perceived discreteness, naturalness, immutability, stability, and necessary attributes) was uncorrelated with the degree to which it was rated as cohesive entity (perceived inductive utility, uniformity, inherence, and exclusivity; Coley et al., 2019; N. Haslam et al., 2000, 2002; Toosi & Ambady, 2011). These two factors have been interpreted (respectively) as essentialism and entitativity, driving theorizing that the two are independent facets of group perception (Prentice & Miller, 2007). However, this two-facet model would place inductive utility – a well-established affordance of psychologically essentialized categories – and inherence – the perception that group members share an underlying sameness – within entitativity and outside of essentialism (Demoulin et al., 2006). Other scholars have documented an apparently bidirectional relationship, where entitativity can lead to essentialism, and essentialism can lead to entitativity (Yzerbyt et al., 2001). In line with this, I propose that the perception of shared experiences – closely linked to entitativity – and the perception of shared biogenetic factors can both drive essentialist beliefs.

A Multiplicity of Causal Forces

Consider this question: why is it that there are more men than women in science and technology fields? A biogenetic essentialist response holds that this difference is an inevitable product of category-linked biological differences between men and women (e.g., XY vs. XX chromosomes; panel 1 of Figure 13). This is deterministic and inevitable, at least within an individual's lifetime. An individualist response says that only individuals affect their own outcomes (panel 2); this response has no way to account for large group disparities. To the extent this belief is based on uncontrollable factors specific to the individual, rather than individual agency, it may be considered essentialist under the human determinism version of essentialism (e.g., Keller, 2005; see Social Categories). However, it is unrelated to social group membership and therefore outside of essentialism as considered here. To the extent this belief is based on the perceived free will of the individual, it is more likely to be common in cultures that emphasize individual agency (Menon et al., 1999).

An environmental response might emphasize the microaggressions faced by women in STEM spaces (panel 3); because microaggressions are negative, women leave. This describes environmental barriers (or affordances) and how they affect different groups of people, but the locus of causality remains outside of the individual group members. In contrast, a socioculturally essentialist response could focus on childhood socialization (panel 4), and how the different ways parents interact with boys and girls can lead to gendered preferences in adulthood (e.g., Pruden & Levine, 2017). While the cause of the difference (parental socialization practices) is external to the group member, the causal pathway goes through them; boys are fundamentally more spatial because of those socialization experiences. As with essentialism in general, people may implicitly hold these beliefs without fully worked-out causal theories. The critical difference

between biogenetic and socioculturally essentialist perspectives is that socioculturally essentialist perspectives admit greater potential for change.

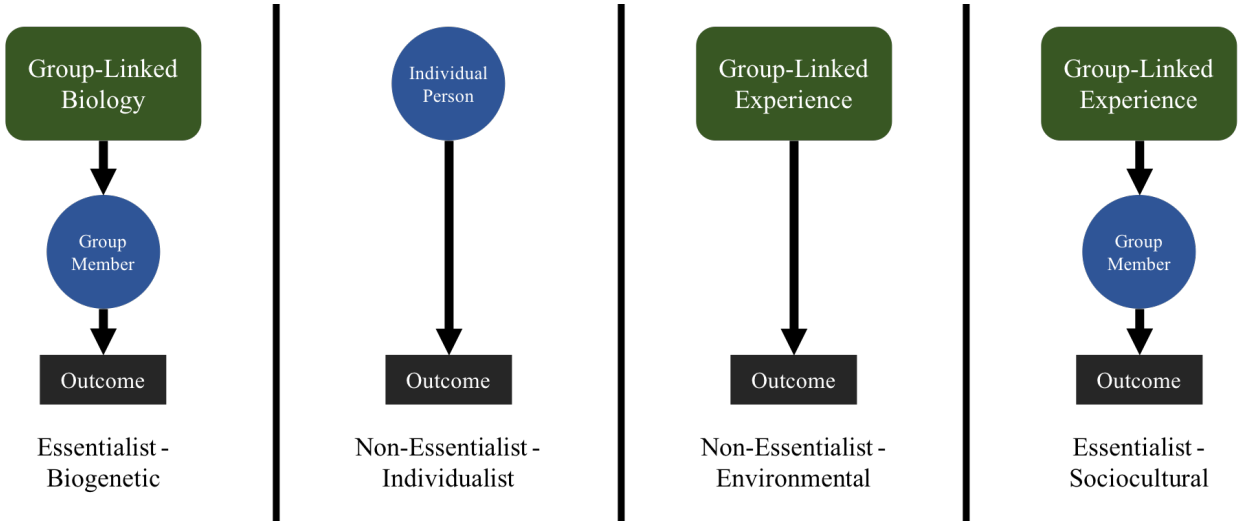

Figure 13. Different kinds of explanations, a non-exhaustive set. In essentialist explanations the group member is fundamentally affected by their group membership (biologically or socioculturally). In individualist explanations, group-linked experiences and biology are irrelevant to outcomes. In environmental explanations, group-linked experiences are relevant, but do not fundamentally change the group member.

From a theoretical perspective, these different beliefs need not be mutually exclusive. They likely coexist or trade off with each other, and with others not described here. People in the US consider gender differences and sexual orientation to be the product of both genetic and environmental factors (Coleman & Hong, 2008; Sheldon et al., 2007), and hold multi-causal theories to explain individual differences (Jayaratne et al., 2009). Biogenetic essentialist beliefs about race have correlated positively with sociocultural essentialist beliefs about race among White, Black, and Latinx participants (r ranged from 0.32 to to 0.51; Yalcinkaya et al., 2017), while biogenetic and sociocultural essentialist beliefs about individual differences are not inversely related ($r = 0.06$ to $r = 0.29$; Rangel & Keller, 2011). It is also likely that these beliefs are differentially applied based on the type of group in question. When someone believes group members share genes, they may rely on a biogenetic essentialism (e.g., race, Hirschfeld, 1995). When someone believes group members share social experiences (e.g., religious groups; Bailey

et al., In Press), they may rely on sociocultural essentialism. This also means that different people may hold different kinds of essentialist beliefs about a particular social domain (e.g., nationality, Rad & Ginges, 2018), or the same person may hold different kinds of essentialist belief in different contexts.

Boundary Conditions

However, not all sets of people are likely to be essentialized; patterns of essentialist thought generally track with salient dimensions of society (e.g., religious groups are more essentialized in Northern Ireland than the US; Smyth et al., 2017). Consider two sets of people: Rich People and Soccer Fans. Though neither is likely to be judged as biologically-based, the former is likely to be more essentialized than the latter, because members are expected to share more social experience in this entitative group (e.g., privilege; Lickel et al., 2000). If a category is not seen as socially relevant enough to provide meaningful shared social experience among group members, it should not be essentialized; this is the *shared experiences hypothesis*. This boundary condition is hinted at in the finding that American children believe that clothing with an occupational category (e.g., firefighter) is more likely to stay consistent over an individual's lifespan than clothing with a particular color (Hirschfeld, 1995), and the suggestion that common treatment of (e.g., discrimination towards) ingroup members should cue essentialist beliefs (Schmitt et al., 2003; Yzerbyt, Estrada, et al., 2004).

This intuitive distinction can also be understood in terms of category membership requirements to be assigned the category essence. First, the duration of membership should be important (*membership duration hypothesis*). Somebody who is a Rich Person for a day, but a Middle-Class Person all other days, has experienced far more of their life as a Middle-Class Person. Therefore, they should be attributed the essence of – and traits linked to – Middle-Class

People. Someone may be a Soccer Fan for a few hours, then a Football Fan, and a Figure Skating Fan during the Winter Olympics – on each of these days, their interest in a particular sport may affect how people treat them, but they are not consistently sharing experiences with the same set of people (on the other hand, Everton Football Club fans may be essentialized). The idea that more essence is imparted by long-term group membership is related to the finding that people in the US consider younger people more malleable than older people (Neel & Lassetter, 2015).

The influence of membership duration, however, is likely to be altered by the period of life in which the person was a category member (*formative period hypothesis*). In particular, I suspect that an understanding of developmental periods will offer extra importance to non-biological categories in which one was a member as a child[5]. For example, if somebody was a Middle-Class Person from ages 1 to 25, and then a Rich Person from 25 to 50, they will have had equal duration membership in each category. However, since the Middle-Class years were during youth and young adulthood, they should be perceived as leaving a greater impression on that person. Some evidence for this boundary condition can be found among the Wichí of Argentina, who believe that a baby is born with many potential ethnic identities, goes through a growth and socialization period, and after that point has a fixed ethnic identity (Erut, 2017, p. 50).

In This Paper

I use an experimental approach to test whether groups which are perceived to share social experience are essentialized. I first gather norming data to select a set of groups that are defined (by people on Amazon Mechanical Turk) as varying in their level of shared biology and shared social experience (e.g., racial and religious groups).

---

[5] This is conceptually similar to the distinction between ascribed and acquired categories, where the former are consistent over the lifespan and the latter are temporally variable (Lei, 2017).

In the two studies, I use a modified switched-at-birth task as a strict test of sociocultural essentialism. The classic task is focused on traits which are retained from biogenetic inheritance even after environmental change (e.g., a baby born to a Kazakh couple who is adopted very young by a Mongol couple; Gil-White, 2001). Responses are conventionally considered essentialist if the baby grows up to have group-linked attributes and group membership of the birth parents (but see Segev et al., 2012 for an alternative interpretation of results from the classic task). However, this leaves no space for sociocultural essentialism – what if the child had been adopted at age 15? Would they be judged to grow up even more typical of the birth parents' group, given that time for shared socialization? We don't know.

To incorporate the possibility of shared social experience, I modified the switched-at-birth task into a switched-category task. Participants learn about somebody who switched from one social group to another of the same dimension (e.g., was rich, is now poor), and are asked about the group membership and attributes of that person after the switch. In Study 1, I used a version of this task which emphasizes that the person was a member of the initial category for a long time. This was done to avoid participants making a switched-at-birth inference (e.g., that the person was rich for a year or two and poor for many years). In Study 2, I specify the age at which the person switched groups, to test whether the duration of membership and person's age at the time of membership affect their retention of the initial group's attributes.

I take two approaches to defining whether the group is essentialized. First, a participant saying that the person remains a "true" member of the initial group reflects immutability beliefs. Second, participant responses on the stereotyping measure can be understood as a less-explicit version of essence attribution. If participants' trait choices are predicted by trait association with the initial group ($A_{initial}$), this suggests the individual retains the initial group essence.

Category Norming

As a first step, I needed to select a set of groups which varied in the extent to which they are perceived to share common experiences, and the extent to which they are perceived to share biogenetic material. In order to use the switched category task, I needed to have at least two groups per social dimension. I conducted a norming process with 40 groups that I expected to vary in their perceived shared social experience and perceived shared biogenetic material: Vegans, Vegetarians, Men, Women, Transgender People, Republicans, Democrats, Liberals, Conservatives, Black People, White People, Asian People, Hispanic People, Native American People, Middle Class People, Poor People, Rich People, Unemployed People, Working Class People, Heterosexual People, Homosexual People, Bisexual People, Construction Workers, Students, Christians, Muslims, Jewish People, Hindus, Left-Handed People, Right-Handed People, Tall People, Short People, People Who Play Hockey, People Who Play Soccer, Southerners, Midwesterners, East Coasters, and West Coasters, English-Speakers, and Spanish-Speakers.

I recruited 204 participants from Amazon Mechanical Turk to complete the norming process. Four were excluded: 1 failed the attention check, and 3 said they did not take the study seriously. This leaves a final N of 200 (M_{age} = 38.88, SD_{age} = 10.08; 62.50% men, 37.00% women, 0.50% gender not listed; 72.00% White, 18.00% Black, 7.50% Asian, 6.50% Latinx, 1.0% Middle Eastern or Arab; $M_{duration}$ = 17 minutes, $SD_{duration}$ = 9 minutes). Each participant rated 10 groups, randomly selected from all the groups and in a random order; each group was

rated by 36 to 63 participants[6]. After completing the screening process and indicating their consent to participate, participants answered four questions about each group on a 0-100 scale (*Not At All – Completely*; based on Kang et al., 2015), and were asked to define the group. Participants finished all questions about one group before proceeding to the next:

Table 10. Group norming questions about each group.

(1) Out of all the [group members] in the world, if you were to choose two of them at random, how **genetically** similar would they be?
(2) Out of all the [group members] in the world, if you were to choose two of them at random, how **biologically** similar would they be?
(3) Out of all the [group members] in the world, if you were to choose two of them at random, how similar would their **social experiences** be?
(4) Out of all the [group members] in the world, if you were to choose two of them at random, how similarly would they be **treated by society**?
(5) What is the definition of being a [GROUP MEMBER]?

Participants then reported their demographics and answered the attention check and seriousness check questions. A group's normatively-perceived shared biogenetics was defined as the mean of all genetic and biological similarity judgments (in a multi-level model with random intercepts for participant and target group, genetic ratings strongly predicted biological ratings, $t(1617) = 38.93$, $p < 0.0001$). A group's normatively-perceived common fate was defined as the mean of all social experience and societal treatment questions (in a multi-level model with random intercepts for participant and target group, social experience ratings strongly predicted social treatment ratings, $t(1856) = 29.26$, $p < 0.0001$). Based on these two judgments, I chose 12 groups in 6 social domains for use in the studies: Black people, White people, rich people, poor people, Democrats, Republicans, women, men, vegetarians, vegans, Jewish people, and Muslim

[6] The first three participants rated 20 groups each; as it was taking them longer to complete 20 ratings than is recommended for the duration of Mturk studies, the number of groups rated by each participant was reduced and the number of total participants recruited was increased.

people. Figure 12 shows the ratings of these groups, along with the ratings of the other 40 groups. I selected groups where two groups from the same social domain (e.g., religion) were evaluated similarly on the two dimensions. Gender is notable exception, where women were judged substantially more biogenetically similar than men; this may be a result of linguistic androcentrism where "men" often means people overall. Women and men were still included as stimuli to put this study in conversation with the existing literature on gender essentialism.

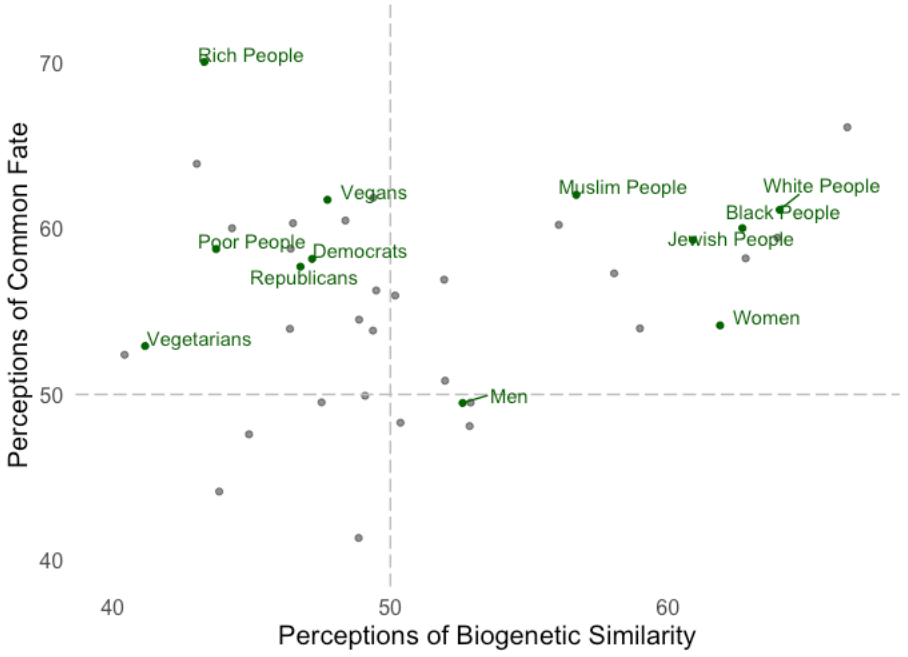

Figure 14. Norming ratings of biogenetic similarity and common fate for 40 social groups. Groups selected for use in the next studies are highlighted in green.

For the selected groups, I read all the definitions that participants had offered of what it means to be a group member, using those to generate definitions to use as stimuli in the two studies (see Table 11). I used this approach to maximize the extent to which participants in the studies found the definitions meaningful, while also ensuring parallel structure in the definitions of each group in a given social dimension. I specifically excluded any definition content that was

(a) physical (e.g., skin color), or (b) an attribute included in the Adjective Checklist from Chapter I (e.g., masculine, feminine).

Table 11. A definition of each social group to be used as stimuli in the two studies, based on the free-response data of pilot participants.

Social Group	Definition Based on Free-Response Data
Black people	A person who identifies with their African heritage and is recognized as Black.
White people	A person who identifies with their European heritage and is recognized as White.
Poor people	A person who has little money, status, or opportunity.
Rich people	A person who has a lot of money, status, and opportunity.
Democrats	A person who holds socially and fiscally liberal views.
Republicans	A person who holds socially and fiscally conservative views.
Men	A person who identifies as a man and other people see as a man.
Women	A person who identifies as a woman and other people see as a woman.
Vegans	A person who does not eat any food derived from animals and doesn't use products that exploit animals.
Vegetarians	A person who does not eat meat.
Jewish people	A person who follows the religion of Judaism and believes in Yahweh.
Muslim people	A person who follows the religion of Islam and believes in Allah.

Study 1: The Shared Experiences Hypothesis

In Study 1, I tested the shared experiences hypothesis: that social categories are essentialized based on the perception of common fate, even if they are not perceived to be biological. Participants read brief paragraphs about individual people identified only by their initials, in which the individual had switched from one social group to another social group within a domain. They then described the person using the adjective checklist measure developed in Chapter I, judged the group membership of the individual, and wrote a brief explanation for how the individual happened to change social groups.

Method

Participants. I recruited 198 participants from Amazon Mechanical Turk to complete the study. One was excluded for missing the attention check, for a final *N* of 197 (M_{age} = 38.39, SD_{age} = 11.52; 54.82% men, 44.67% women, 0.51% gender not listed; 80.71% White, 10.66% Black, 8.63% Latinx, 6.60% Asian, 1.52% Native American, 1.02% Middle Eastern, 0.51% Pacific Islander, 1.02% race/ethnicity not listed; $M_{duration}$ = 13 minutes, $SD_{duration}$ = 7 minutes).

Procedure. This study was completed as a 6 (social domain) x 2 (direction of change) design, partly within and partly between participants. Each participant was randomly assigned to three of the six social domains; within each social domain, they were randomly assigned to one of the two directions of group change (each of the 12 conditions was completed by 43 to 55 participants). For each condition, participants saw a brief description of an individual, identified only by a pair of initials. In the description, participants read that the person was part of an initial group for a long time, but that the person was part of a different group now. Definitions of each group, shown in Table 11, were included in the description.

The description of the individual is minimal by design. First, providing any more information might bias the results in some way towards one group or the other; additional detail added to the description would need to be entirely unrelated to any of the groups included in the study (and even if it was, it could undermine the focal category identity; Nisbett et al., 1981). Second, providing more information might influence the perception of the group-transition process. If a person is described in seemingly neutral terms (e.g., as an okay person), this may seem to preclude possible means of group switch (e.g., if J.M. is just okay, it may seem that they achieved some great feat to accrue substantial wealth, and so perhaps participants would be more likely to attribute the group-change to luck). Finally, this level of detail is similar to the

switched-at-birth design, making it as comparable as possible to the existing literature. This is a sample description for an individual who was a poor person and is now a rich person:

Table 12. Example description of a person in the switched-group task.

J.M. was a poor person for a long time, they had little money, status, or opportunity. **J.M. is a rich person now**, they have a lot of money, status, and opportunity.

After completing the screener (see Appendix D: Online Screening) and consenting to participate, participants completed the adjective checklist measure developed in Chapter I, with the instructions included in Appendix E: Adjective Checklist Measure Instructions (once per condition, in a random order). Using this measure, I can test for patterns of essentialism by seeing whether the choice of a particular adjective in a particular condition is predicted by the adjective's association with the initial group ($A_{initial}$), and its association with the latter group (A_{latter}). This is an advantage because it allows the same measure to be used across conditions and incorporates a terms' relative association with both groups, rather than having domain-specific stimuli with only strongly-differentiated terms (e.g., terms which differentiate women and men but not poor people and rich people).

Once they were done with the adjective checklist measure, participants moved on to a block of questions about the group membership of each individual. These questions are focused on having the participant consider the fundamental identity of the individual (Newman & Knobe, 2019). If the participant recognizes that J.M. now fits the definition of a rich person, but thinks J.M. is – in J.M.'s central self – a poor person, these questions are intended to capture that intuition. This is a sample set of questions, from the same condition as Table 12:

Table 13. Example questions about group membership from the switched-group task.

(1) Is J.M. **truly** a rich person?
 Yes
 No
(2) Is J.M. **truly** a poor person?
 Yes
 No
(3) Is J.M. **more of** a rich person or **more of** a poor person?
 J.M. is more of a rich person
 J.M. is more of a poor person

Participants then generated a brief causal narrative to explain why or how the individual changed between groups ("What do you think is the most likely reason that J.M. used to be poor and now is rich?"). For each condition, participants saw the same group description as in the prior two blocks, and had an open text box to answer the causal question. After completing all measures about the groups, participants reported whether they were a member of any the groups about which they responded. If so, they reported the identity centrality of that self-aspect, using the same measure as Chapter I. Finally, they reported their demographics, including the attention and seriousness checks (see Appendix F: Attention Check and Seriousness Check).

Results

The central question in this study was whether social groups believed to have substantial shared social experience would be essentialized; I tested this in four stages. First, I looked at when participants said the described individual was *more of* an initial group member than a latter group member, relying on this as an explicit measure of essentialism. Next, I tested the relationship between this operationalization of essentialism and essentialism as indexed by the adjective checklist: whether participants' adjective choices were guided by the adjective's association with the initial group ($A_{initial}$ from Chapter I). After this, I tested whether participants reliably essentialized the different social domains, according to the adjective checklist data. In

these analyses, I whether the individual is seen to have *retained* initial-group attributes as an index of essentialism; I also tested whether they were seen to have *gained* latter-group attributes. Finally, I noted adjective choices that were poorly predicted by the prior analyses, to explore how participants were interpreting the scenario above and beyond stereotypes of the initial and latter groups.

"**True" Group Membership.** Across all conditions, when forced to make a binary choice, participants denied the change 35.03% of the time, saying the individual was more of a member of the initial group (question 3 from Table 13). This differed by condition (see Figure 15); participants were significantly more likely to deny the change about Race than Gender (z = 2.870, p = 0.004), and Gender than Religion (z = 2.655, p = 0.008). However, this was not more likely for Religion than Political Ideology, Wealth Class, or Eating Habits (ps > 0.25). These results suggest that Race and Gender are more essentialized than the other social domains. However, it is worth noting that even for Eating Habits, 14.29% of participants denied that the change had occurred – Race and Gender were more essentialized on this explicit measure, but some participants considered group membership immutable for every social domain.

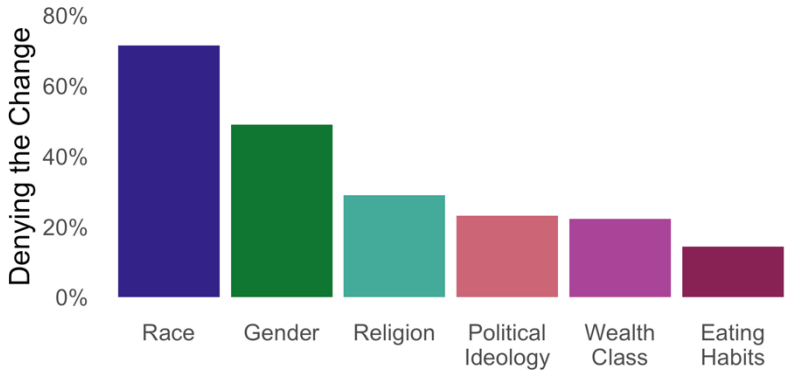

Figure 15. Proportion of participants who said the individual was more of an initial group member.

I next considered how my two indicators of essentialism were related. One straightforward possibility was that change-deniers would choose adjectives associated with the initial category. If a participant says someone who was Jewish and is now Muslim is *really* still Jewish, that participant could describe them with adjectives associated with Jewish people (based on $A_{initial}$), and without adjectives associated with Muslim people (based on A_{latter}). In this case, the adjective choice of change-deniers would rely more on $A_{initial}$ and less on A_{latter} than the adjective choices of change-accepters. This would mean that both operationalizations of essentialism point in the same direction – participants who explicitly essentialized the person also stereotyped them more according to their initial group membership.

However, another possibility is that change-deniers – people who explicitly reject the possibility of the group transition – would simply dislike a person who challenged their beliefs about the immutability of group membership. In this case, the participant who denied that a Jewish person could become a Muslim person would describe that person with negative adjectives, independent of whether these adjectives otherwise had any particular association with these religious groups. This would reduce the extent to which $A_{initial}$ or A_{latter} could predict their adjective choice, such that the two operationalizations of essentialism would counteract each other. Participants who explicitly essentialized the person would be stereotyping them less in terms of their initial group membership.

I used two sets of binomial multi-level regression models predicting adjective choice to test this (Bates et al., 2015; Kuznetsova et al., 2017; Raudenbush & Bryk, 2002). In the first set, I tested whether change-deniers relied more or less on $A_{initial}$ and A_{latter} than change-accepters, controlling for A_{latter}. These models included a random effect for participant, accounting for individual differences in likelihood of choosing adjectives overall. They also included a fixed

effect of membership judgement (contrasting change-deniers with change-accepters), a fixed effect of $A_{initial}$, a fixed effect of A_{latter}, and the interactions of $A_{initial}$ and A_{latter} with membership judgement. The interactions are of central interest: are change-deniers more or less likely than change-accepters to choose adjectives based on $A_{initial}$ and A_{latter}? For Gender, Race, and Religion, change-denying participants were less likely than change-accepting participants to associate the individual with adjectives based on $A_{initial}$ (see Table 14 for all results and Figure 16 for sample graphs for the Race domain). This means that, for these domains, change-denying participants were less likely than change-accepting participants to essentialize the individual in terms of initial-group attributes.

To see whether the adjective choice of change-denying participants was instead being guided by valence, I conducted a second set of binomial multi-level regression models. These models included a random effect of participant, a fixed effect of membership judgement, a fixed effect of adjective valence, and a fixed effect of the interaction between valence and membership judgment. Once again, the interaction is of central interest. For Gender, Race, and Religion, change-denying participants used more negative terms to describe the individual than change-accepting participants (see Table 14 for all results and Figure 16 for sample graph for the Race domain).

Table 14. Participants who say the described individual is more likely to be an initial group member chose adjectives that were less typical of the initial group, less typical of the latter group, and more negative.

Social Dimension	**$A_{initial}$**	**A_{latter}**	**Valence**
Eating Habits	$z = 1.128$ $p = 0.26$	$z = 1.081$ $p = 0.28$	$z = –0.020$ $p = 0.98$
Gender	$z = –2.781$ $p = 0.005$	$z = –4.719$ $p < 0.0001$	$z = –8.691$ $p < 0.0001$
Political Ideology	$z = 2.081$ $p = 0.04$	$z = –3.400$ $p = 0.0007$	$z = 1.065$ $p = 0.29$
Race	$z = –3.096$ $p = 0.002$	$z = –3.256$ $p = 0.001$	$z = –5.385$ $p < 0.0001$

Social Dimension	$A_{initial}$	A_{latter}	Valence
Religion	$z = -2.418$ $p = 0.02$	$z = -1.085$ $p = 0.28$	$z = -3.194$ $p = 0.001$
Wealth Class	$z = 1.057$ $p = 0.29$	$z = -4.560$ $p < 0.0001$	$z = 2.387$ $p = 0.02$

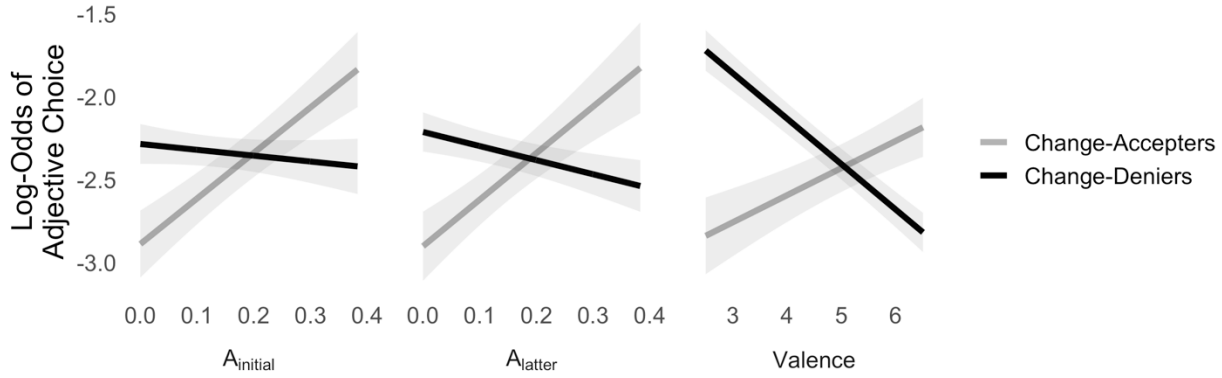

Figure 16. Change-deniers in the Race condition were less likely than change-accepters to describe the individual using terms associated with the initial group or terms associated with the latter group. They were more likely than change-accepters to use more negative terms.

For Race, Gender, and Religion, the two operationalizations of essentialism in this data do not line up: participants who deny that the individual changed groups are less likely than participants who accept that the individual changed groups to assign the individual adjectives associated with the initial group. Instead of affirming the individual's initial group membership through stereotypical adjectives, these participants choose more negative adjectives to describe that individual. To focus on the second measure of essentialism – the attribution of traits based on $A_{initial}$ – I therefore excluded change-denying participants in these conditions.

Adjective Choice. After looking at how the two operationalizations of essentialism aligned, I turned to a full analysis of the adjective checklist data. The measure appeared to have good coverage, with no more than 15% of participants in any condition using the option to add new adjectives (for further descriptive analysis of adjective choice, see Appendix H: Descriptive Analyses of Adjective Choice for Chapter II). Having determined the measure was appropriate

for use in this context, I turned to seeing whether participants' adjective choice showed essentialist responding. If change-accepting participants still chose adjectives based on their stereotypic association with the initial group ($A_{initial}$), I consider that evidence of essentialism.

However, traits can be associated with more than one group, and some may be similarly associated with the individual's initial group and their second group (e.g., hardworking has $A_{RichPeople} = 0.23$ and $A_{PoorPeople} = 0.21$). I therefore predict the participant's binary choice on a given adjective both by $A_{initial}$ and A_{latter}. If the adjective's association with the initial group predicts adjective choice above and beyond the adjective's association with the second group, the individual is retaining essence of the initial group. I used a multi-level binomial model to predict adjective choice in each social domain. Each model contains a random intercept of participant and five fixed effects: condition, $A_{initial}$, A_{latter}, and the interactions of $A_{initial}$ and A_{latter} with condition. For each model, I report the main effect of $A_{initial}$ and A_{latter} in Table 15 and display the main effect of $A_{initial}$ in Figure 17.

For Race, Religion, Eating Habits, and Political Ideology, $A_{initial}$ predicted adjective choice, above and beyond A_{latter}. Participants were more likely to pick adjectives associated with the initial group, and less likely to pick adjectives un-associated with the initial group. While Race and Religion were judged (in Category Norming) to share common biogenetics and common fate, Eating Habits and Political Ideology were judged only to share common fate. This pattern shows that, as predicted, both kinds of groups can be essentialized, such that individuals who have left the group retain its traits. However, this straightforward essentialist pattern did not emerge for Gender or Wealth Class. I explore these two domains further in the next sections.

Table 15. Effect of adjective association with initial and second group predicting adjective choice.

Social Dimension	$A_{initial}$	A_{latter}
Eating Habits	$z = 2.739$ $p = 0.006$	$z = 6.673$ $p < 0.0001$
Gender	$z = -1.760$ $p = 0.08$	$z = 4.416$ $p < 0.0001$
Political Ideology	$z = 3.893$ $p < 0.0001$	$z = 12.044$ $p < 0.0001$
Race	$z = 2.051$ $p = 0.04$	$z = 0.593$ $p = 0.55$
Religion	$z = 6.245$ $p < 0.0001$	$z = 4.348$ $p < 0.0001$
Wealth Class	$z = -2.317$ $p = 0.02$	$z = 14.668$ $p < 0.0001$

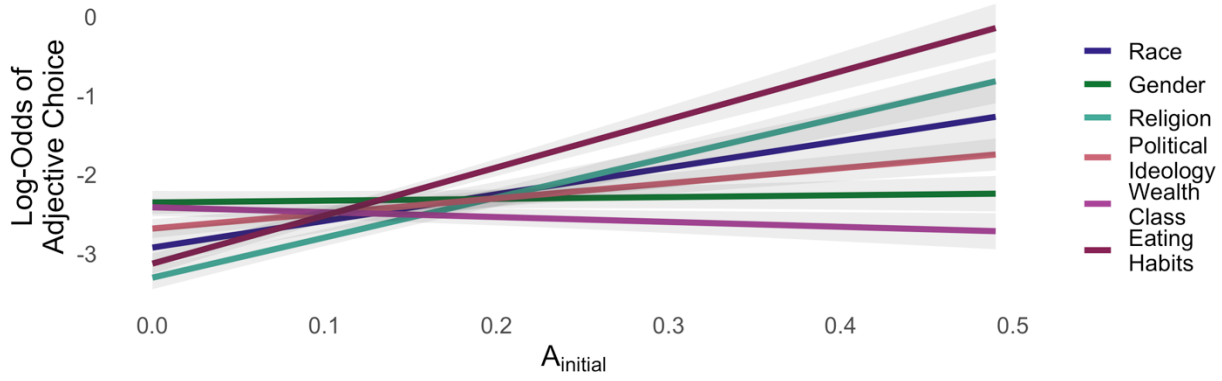

Figure 17. For Eating Habits, Political Ideology, Race, and Religion, the strength of an adjective's association with the individual's initial group predicted greater likelihood of selecting the adjective. For Gender and Wealth Class, this effect did not emerge.

Gender and Transgender Stereotypes. Of the six social domains, Gender may be unique because of growing cultural awareness about transgender people. Participants may have interpreted the description as referencing a transgender person going through the process of transition. In this view, change-accepting participants were affirming trans* identity and so it is not particularly surprising that these participants emphasized the current gender identity of the described individual. This interpretation is partially supported by the free-response data –

20.83% of change-accepting participants included the word *trans* or *transgender* in their causal description. Further, it fits with the most-chosen adjectives (see Table 32), where *unique* and *expressive* rank highly for both conditions (these adjectives were also chosen more frequently than would be predicted by their association with women or men; see Table 16). To test this interpretation more stringently, I used $A_{transgender}$ as defined for each adjective from the Chapter I norming data to see whether an adjective's stereotypical association with transgender people predicted adjective choice in the Gender conditions. In this multi-level regression, $A_{transgender}$ predicted adjective choice ($z = 6.970$, $p < 0.0001$), an effect which did not differ by condition. In the Gender domain, the description of someone switching gender groups seems to have activated ideas about transgender people, such that the individual was assigned adjectives based on their association with that group.

Wealth Class and Valence. Despite the normative perception that members of the same wealth class share substantial social experience, $A_{initial}$ negatively predicted adjective choice in this social domain. In the Wealth Class condition, the process of switching between groups may activate approval (of those who gained wealth, by whatever mechanism) and disapproval (of those who have lost wealth, by whatever mechanism). That is, because there is open and socially acceptable approval of rich people and disapproval of poor people in a nominally meritocratic society (e.g., Lei & Bodenhausen, 2017), participants may have had an affective reaction about the process of going between the two groups that obscured the effects of $A_{initial}$ and A_{latter}.

To test this explanation, I first used the norming data gathered in Chapter I and looked at correlations between A and adjective valence. For eleven of the twelve target groups, A correlated with valence (ps ≤ 0.01; for Republicans, $r(68) = 0.10$, $p = 0.39$). Poor people was the only target group for which A was negatively correlated with valence ($r(68) = -0.55$, $p <$

0.0001); the higher $A_{PoorPeople}$, the more negative the adjective. This means that Wealth Class was the only social domain in which the described individual was transferring between a normatively positive group and a normatively negative group. This does not mean there was no valence differentiation between groups in other domains; for example, $A_{WhitePeople}$ correlated more highly with valence than did $A_{BlackPeople}$ (White: $r(68) = 0.69$, $p < 0.0001$; Black: $r(68) = 0.40$, $p = 0.0007$), indicating that the norming participants endorsed relatively more positive adjectives for the White than Black target group. However, Wealth Class was the only domain in which the person transferred between groups that had an absolute difference in valence.

Given this, it seemed plausible that valence might play an extra role in this domain, relative to the other domains. I tested this by controlling for valence in the regressions. For each condition, I used a multi-level binomial model predicting adjective choice, including a random intercept for participant, a fixed effect for $A_{initial}$, a fixed effect for A_{latter}, and a fixed effect for adjective valence. As expected, valence strongly predicted adjective choice in each condition, in opposite directions (rich-to-poor: $z = -7.919$, $p < 0.0001$; poor-to-rich: $z = 12.20$, $p < 0.0001$). Beyond $A_{initial}$ and A_{latter}, participants chose approving adjectives for the poor-to-rich individual and disapproving adjectives for the rich-to-poor person. When controlling for valence, Wealth Class was essentialized in the same manner as Race, Religion, Political Ideology, and Eating Habits: $A_{initial}$ positively predicted adjective choice alongside A_{latter} (rich-to-poor $A_{initial}$: $z = 2.340$, $p = 0.02$; rich-to-poor A_{latter}: $z = 9.693$, $p < 0.0001$; poor-to-rich $A_{initial}$: $z = 2.12$, $p = 0.03$; poor-to-rich A_{latter}: $z = 11.98$, $p < 0.0001$). In the case of Wealth Class, valence starkly differentiates the two groups; controlling for valence reveals a pattern of essentialist response where $A_{initial}$ predicts adjective choice.

Emergent Content. In the prior sections of adjective checklist analysis, I have focused on whether and when an adjective's association with an initial or latter group predicts participants choosing that adjective. For Gender and Wealth Class, I have also assessed how participants' ideas about the process of group-switch seems to have impacted the role of $A_{initial}$ and A_{latter}. In this last section of results, I follow up on this across domains, asking: when were participants not choosing adjectives based on $A_{initial}$ and A_{latter}? What else were they doing in attributing traits to the individuals in these sparse descriptions? This is by nature exploratory and intended to contextualize the results of the analyses up to this point.

For each condition, I used binomial models similar to the prior section to calculate the model-based likelihood that an adjective would be selected in a given condition (based on the overall intercept, $A_{initial,}$ and A_{latter}). I compared the model-based value with the actual proportion of participants who chose the adjective in that condition. In Table 16, I show all adjectives chosen 15% more often or 15% less often than would have been predicted by the model – these are adjectives whose selection was not well-predicted according to their stereotypic association with either group. In general, we see that participants included adjectives more than would be expected based simply on $A_{initial}$ and A_{latter}.

Looking more closely at the adjectives, they seem connected to ideas about the process of transferring between groups; for Religion, Political Ideology, and Eating Habits, *curious* and *thoughtful* were chosen more than would be expected, suggesting a process of exploration before switching groups. For Gender, the under-predicted adjectives were similar to those gathered in other research assessing stereotypes of transgender people (Gallagher & Bodenhausen, Under Review). For Wealth Class they were connected to approval and disapproval (as would be expected from the role of valence in the Wealth Class and Valence analyses). This shows that,

while $A_{initial}$ and A_{latter} may have guided adjective attribution, change-accepting participants did not simply evaluate the person as a member of either group, but as someone who had gone through some process to change groups.

Table 16. Adjectives chosen 15% more often or 15% less often than would be predicted by their strength of association with the initial or latter group.

Social Dimension	Condition	Terms Chosen More Frequently Than Would Be Expected from Model	Terms Chosen Less Frequently Than Would Be Expected from Model
Eating Habits	Vegan to Vegetarian	thoughtful, healthy	important, happy
	Vegetarian to Vegan	thoughtful, healthy	important
Gender	Men to Women	expressive, feminine, unique, comfortable, adult, happy, brave, misunderstood	fashionable
	Women to Men	brave, unique, interesting, expressive, misunderstood, masculine, proud	--
Political Ideology	Democrat to Republican	responsible	--
	Republican to Democrat	responsible, thoughtful, curious	powerful
Race	Black to White	selfish, motivated, interesting	--
	White to Black	expressive, misunderstood, unique, proud, cultural, uncomfortable, struggling	--
Religion	Jewish to Muslim	unique, curious, interesting, thoughtful, religious, dedicated, adult, holy	hardworking
	Muslim to Jewish	curious, thoughtful, adult, unique, proud, cultural	traditional, protective
Wealth Class	Poor to Rich	successful, motivated, hardworking, employed, skilled, dedicated, responsible, lucky, interesting, adult, comfortable, brave	powerful, privileged, selfish, aggressive, gentle
	Rich to Poor	struggling, alone, unfortunate, anxious, stressed, uncomfortable	homeless, misunderstood

Study 1 Discussion

Participants in Study 1 showed essentialist patterns of thought for Religion, Race, Eating Habits, and Political Ideology – social domains normatively considered to have both common genes and common fate, and the ones only considered to have common fate. Though research on essentialism of human categories has emphasized the perception of group-linked biology giving rise to essentialist beliefs, this shows that the perception of group-linked social experience can also drive essentialist responding.

There are also more nuanced conclusions to be made from the results. First, though change-denial and choosing adjectives associated with the initial group can both be considered evidence of essentialist reasoning in the switched-group task, they did not function in parallel. Specifically, for Race, Gender, and Religion, participants who denied that the individual had changed groups picked more negative adjectives to describe the individual, *not* adjectives more associated with the initial group.

For Gender and Wealth Class, the pattern of results was more complex. This seems likely to be because of cultural ideas surrounding the process of switching between the groups in these domains. In the Gender case, participants seem to have interpreted the description as representing a transgender person. As transgender people tend to be stereotyped more as members of a "third" gender group than as women or men (e.g., Gallagher & Bodenhausen, Under Review), interpreting these results with an essentialist framework may be inappropriate. In the Wealth Class case, participants had overwhelmingly positive responses to the person who went from poor-to-rich, and overwhelmingly negative responses to the person who went from rich-to-poor. This may be because of meritocratic beliefs that the accrual (or loss) of wealth reflects an individual's character (e.g., Lei & Bodenhausen, 2017) above and beyond category

membership, or it may be because Wealth Class was the only social domain in which the described individual switched between a normatively-negatively-stereotyped group and a normatively-positively-stereotyped group.

In general, participants chose more adjectives than would be predicted by the adjective's level of association with the initial and latter groups (see Table 16). Looking qualitatively at the adjectives that were selected, they seem again to point out that participants are not simply adding together the initial and latter groups, but rather engaging with ideas about how and why the individual had switched groups. It will be important in future work to consider in more detail how ideas about the process of switching groups impacts the strength and frequency of both membership-based and attribute-based patterns of essentialist thought, and when such narratives do – or don't – reify essentialist beliefs about the initial or latter groups.

Study 2: The Membership Duration and Formative Years Hypotheses

In Study 2, I tested the membership duration and formative years hypotheses: that an individual will retain a social group's essence more if they have been a group member longer, and particularly if they were a member during their youth or adolescence. I focus on two social domains – Religion and Political Ideology – that showed essentialist reasoning in Study 1 and differ in their perceived biogenetic basis. I measure and define essentialism in the same way as in Study 1, and use a similar paradigm to study it. However, while Study 1 emphasized that the person had been a member of the initial group for a long time, the paradigm now investigates the effect of duration by stating the group membership of the family into which the person was born, the life-stage at which the person switched social groups, and that the person has been a member of the latter group for 20 years. I used five age labels: child (age 5-12), teenager (age 13-18),

young adult (age 19-29), adult (age 30-44), and middle-aged (age 45-60; for similar use of age-labels, see Neel & Lassetter, 2015).

Method

Participants. I recruited 600 participants from Amazon Mechanical Turk to complete this study. Six were excluded: 4 failed the attention check, and 2 said they did not take the study seriously. This leaves a final *N* of 594 (M_{age} = 38.41, SD_{age} = 14.80; 56.40% women, 43.27% men, 0.34% gender not listed; 71.55% White, 14.65% Asian, 12.12% Black, 6.73% Latinx, 1.68% Middle Eastern, 1.01% Native American, 0.67% Pacific Islander, 1.01% race/ethnicity not listed; $M_{duration}$ = 12 minutes, $SD_{duration}$ = 10 minutes).

Procedure. The design for this study is 2 (social domains) x 2 (direction of switch) x 5 (age of switch). Each participant completed one condition per social domain and was randomly assigned to direction and age of switch (each of the 20 condition-age of switch combinations was completed by 49 to 76 participants). Within a given age label (e.g., teenager, young adult), the participant was randomly assigned to a year-value according to the age range for that label.

Participants completed the study in the same procedure as Study 1, except that the description of was slightly modified to state that (a) the person was born to a family of a given social group, (b) the age at which the person switched groups, and (c) that the person had been part of the latter group for 20 years (sample description in Table 17). Participants completed all the same measures as in Study 1 for each condition to which they were assigned. Participants then completed measures of group membership, centrality, demographics, and quality checks as in Study 1.

Table 17. Sample description for the switched-group task for Religion with teenager as age-of-switch.

J.M. was born into a Jewish family, they followed the religion of Judaism and believed in Yahweh. Twenty years ago, **when J.M. was a 14-year-old teenager, they became a Muslim person.** Ever since then, they have followed the religion of Islam and believed in Allah.

Results

The central questions in this study were (a) whether duration of initial group membership would increase the perception of the individual retaining initial-group essence and (b) whether this would be nonlinear such that membership during adolescent years was especially important. I used the same two indicators of essentialism as in Study 1, and followed a similar analytic process.

"True" Group Membership. Across conditions, participants said the person was more of an initial group member 13.89% of the time (question 3 from Table 13). This was more common in the Religion than Political Ideology domain ($z = -7.293$, $p < 0.0001$), and more common when evaluating individuals who had switched groups later ($z = 3.240$, $p = 0.001$; see Figure 18). However, the curvilinear effect of age was not significant ($p = 0.53$).

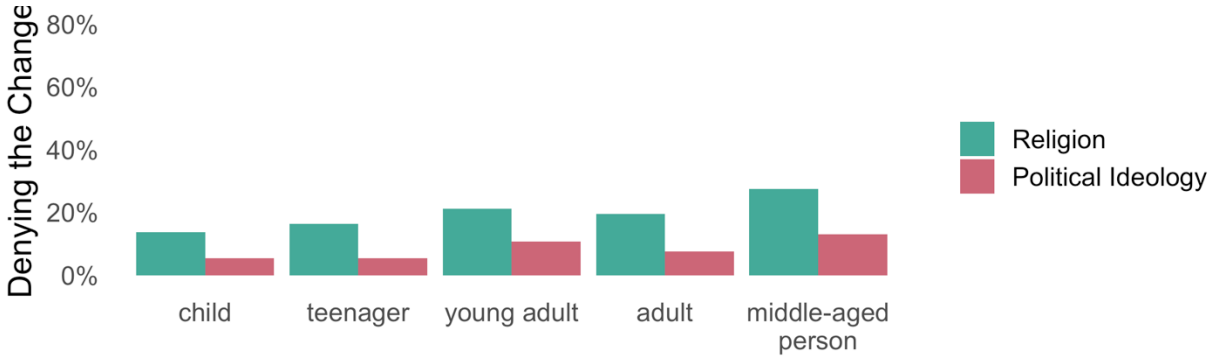

Figure 18. Participants were more likely to think the person was still an initial group member for Religion than Political Ideology, and for people who switched groups later than people who switched groups earlier.

In the Religion domain in Study 1, change-denying participants attributed the individual less initial-group adjectives and less latter-group adjectives than change-accepting participants. Change-denying participants also chose more negative terms to describe the individual. In the

Political Ideology domain in Study 1, change-denying participants attributed the individual more initial-group adjectives and less latter-group adjectives than change-accepting participants. I tested whether these patterns were replicated in Study 2. I analyzed each Age Group within each domain, using the same modeling approach as Study 1 (for results of interactions of interest, see Table 18).

For Religion, regardless of Age Group, change-denying participants ascribed more negative adjectives to the individual. However, change-denying participants only relied on $A_{initial}$ less than change-accepting participants when the described individual changed groups in childhood or as a middle-aged person. This partially replicates the pattern from Study 1, and suggests that change-denying participants broadly dislike individuals who change religious groups. As in Study 1, I therefore excluded adjective selection data from these participants.

For Political Ideology, change-denying participants relied more on $A_{initial}$ than change-accepting participants, but only when the individual switched groups as a child, a young adult, or a middle-aged person. This partially replicates the Study 1 finding.

Table 18. In the Religion domain, change-denying participants assigned more negative adjectives to the described individual than change-accepting participants.

Social Dimension	Age Group	$A_{initial}$	A_{latter}	Valence
Political Ideology	Child	$z = 3.762$ $p = 0.0002$	$z = -2.163$ $p = 0.03$	$z = 2.508$ $p = 0.01$
	Teenager	$z = -0.772$ $p = 0.44$	$z = -1.341$ $p = 0.18$	$z = -1.597$ $p = 0.11$
	Young Adult	$z = 2.114$ $p = 0.03$	$z = -3.675$ $p = 0.0002$	[7] $z = -0.136$ $p = 0.89$
	Adult	$z = -0.014$ $p = 0.989$	$z = -0.357$ $p = 0.72$	$z = 0.22$ $p = 0.83$

[7] This model did not converge as a multi-level model, and these results are from a simple binomial regression. I include them for consistency with the other conditions, but they should be taken with a grain of salt because they do not account for the repeated-measures nature of the data.

Social Dimension	Age Group	$A_{initial}$	A_{latter}	Valence
	Middle-Aged Person	$z = 2.571$ $p = 0.01$	$z = -1.653$ $p = 0.09$	$z = -3.920$ $p < 0.0001$
Religion	Child	$z = -2.850$ $p = 0.004$	$z = 1.064$ $p = 0.29$	$z = -3.245$ $p = 0.0001$
	Teenager	$z = 1.615$ $p = 0.11$	$z = -3.802$ $p = 0.0001$	$z = -2.772$ $p = 0.006$
	Young Adult	$z = -0.127$ $p = 0.90$	$z = -1.613$ $p = 0.11$	$z = -2.487$ $p = 0.01$
	Adult	$z = -1.072$ $p = 0.28$	$z = -1.166$ $p = 0.24$	$z = -3.886$ $p = 0.0001$
	Middle-Aged Person	$z = -2.540$ $p = 0.011$	$z = -1.597$ $p = 0.11$	$z = -5.378$ $p < 0.0001$

Adjective Choice. In general, the measure again appeared to have good coverage, with participants only adding adjectives 9.05% of the time (for further descriptive analysis of adjective choice, see Appendix H: Descriptive Analyses of Adjective Choice for Chapter II).

Next, I tested the membership duration hypothesis separately for each condition, with binomial multi-level models predicting adjective choice. Each model contained a random effect of participant, fixed effects of $A_{initial}$ and A_{latter}, a fixed linear effect of the individual's age, a fixed curvilinear effect of the individual's age. Each model also contained two interactions: $A_{initial}$ with linear age, and $A_{initial}$ with curvilinear age.

The two interactions were of central interest. If the interaction between $A_{initial}$ and linear age is positive and significant, that supports the *membership duration* hypothesis. If the interaction between $A_{initial}$ and the curvilinear effect of age is positive and significant, that supports the *formative years* hypothesis. These effects are reported in Table 19; I also include the main effect of $A_{initial}$ to see whether the pattern from Study 1 replicated. For three of the four conditions, participant replies were more essentialist when the described individual changed groups at an older age, supporting the membership duration hypothesis. The formative years hypothesis was not supported in these analyses.

Table 19. Effect of adjective association with initial group and age predicting adjective choice.

Condition	$A_{initial}$	Linear Age and $A_{initial}$	Curvilinear Age and $A_{initial}$
Democrat to Republican	$z = 6.499$ $p < 0.0001$	$z = 2.243$ $p = 0.02$	$z = 1.100$ $p = 0.27$
Republican to Democrat	$z = 5.459$ $p < 0.0001$	$z = 2.433$ $p = 0.01$	$z = –1.547$ $p = 0.12$
Jewish to Muslim	$z = 9.902$ $p < 0.0001$	$z = 2.237$ $p = 0.03$	$z = –0.094$ $p = 0.9$
Muslim to Jewish	$z = 7.492$ $p < 0.0001$	$z = –0.730$ $p = 0.47$	$z = 0.155$ $p = 0.88$

Emergent Content. As in Study 1, I looked at which adjectives were chosen substantially more – or less – than would have been predicted by $A_{initial}$ and A_{latter}? In this exploratory analysis, I was particularly interested in how age might have affected participants' ideas about the process of transitioning between groups. For each Age Group in each condition, I used the same approach as in Study 1 to identify the adjectives that were chosen 15% more or 15% less than would have been predicted according to a model using $A_{initial}$ and A_{latter} to pick adjective choice for that age group in that condition.

As in Study 1, participants generally chose adjectives more often than would be predicted by the model. Some appeared consistently across age groups – for instance, participants in general were likely to think the individuals switching groups were *adult*. As in Study 1, some terms that over-appeared seemed to be connected to ideas about what kind of group-switch was occurring; *curious* appears often for both social domains, while *responsible* appears often in the Political Ideology domain and *thoughtful* is frequent within the Religion domain. It is somewhat surprising that the over-predicted and under-predicted adjectives did not qualitatively differ by the individual's age group – this may be a result of all described individuals being adults at the time of their description (albeit ranging in age from 25 – 80).

Table 20. Adjectives chosen 15% more often or 15% less often than would be predicted by their strength of association with the initial or latter group.

Condition	Age Group	Terms Chosen More Frequently Than Would Be Expected from Model	Terms Chosen Less Frequently Than Would Be Expected from Model
Democrat to Republican	Child	proud, responsible, traditional	important
	Teenager	privileged, traditional	important
	Young Adult	adult, privileged, responsible, traditional	important
	Adult	adult, responsible, traditional	important, nationalist, trustworthy
	Middle-aged Person	adult, masculine, privileged, respectful, responsible, selfish	important
Republican to Democrat	Child	democratic, responsible, thoughtful	powerful
	Teenager	curious, dedicated, democratic, responsible, smart, thoughtful, young	powerful
	Young Adult	adult, curious, democratic, feminist, thoughtful	hardworking, powerful
	Adult	adult, curious, democratic, responsible, thoughtful	expressive, important, powerful
	Middle-aged Person	adult, curious, democratic, privileged, responsible, thoughtful	expressive, hardworking, powerful
Jewish to Muslim	Child	adult, cultural, curious, holy, misunderstood, religious	--
	Teenager	adult, curious, dedicated, motivated, religious, thoughtful	protective, successful, traditional
	Young Adult	adult, cultural, curious, holy, misunderstood, religious, unique	happy
	Adult	adult, cultural, curious, dedicated, holy, motivated, thoughtful, unique	powerful, successful
	Middle-aged Person	adult, cultural, curious, holy, thoughtful	--
Muslim to Jewish	Child	brave, curious, dedicated, holy, responsible, thoughtful, unique	nationalist, successful, traditional
	Teenager	adult, cultural, curious, holy, misunderstood, religious, thoughtful	nationalist, powerful, protective, traditional
	Young Adult	adult, cultural, curious, dedicated, religious, unique	traditional

Condition	Age Group	Terms Chosen More Frequently Than Would Be Expected from Model	Terms Chosen Less Frequently Than Would Be Expected from Model
	Adult	adult, brave, cultural, curious, interesting, motivated, respectful, thoughtful, unique	powerful, traditional
	Middle-aged Person	adult, brave, cultural, curious, unique	nationalist, powerful, protective

Study 2 Discussion

As in Study 1, participants in Study 2 showed essentialist patterns of thought for Religion and Political Ideology: some participants denied that switching between groups was possible, and those who accepted the switch still attributed initial-group adjectives to the individual. Study 2 expanded on Study 1 by providing partial support for the membership duration hypothesis. In three of the four conditions, participants described the group-switching individual more in terms of the initial group when that individual had switched groups later in life. However, this pattern did not appear in the Muslim to Jewish condition. The data from Study 2 did not support the formative years hypothesis – though the age at which the person switched impacted the extent to which they retained initial-group attributes, there was no significant curvilinear effect of age such that membership during adolescence or young adulthood would especially impart initial-group attributes.

General Discussion

The two studies in this paper show that groups perceived to share social experience – but not biogenetic factors – can be essentialized. For groups perceived to share common fate, whether or not they were perceived to share biogenetic factors, participants attributed traits to an individual based on their initial group membership. This was limited by membership duration – participants were more likely to attribute traits based on the initial group when the person

changed groups later in life. This means that participants were attentive to the individual's experience, and incorporated information about their life course. However, the formative years hypothesis was not supported; the effect of membership duration was not greater at younger ages (e.g., teenage and young adult years).

Though I set out to test essentialist patterns of reasoning about groups defined by (the perception of) shared social experiences, in contrast with groups defined by (the perception of) shared biogenetic factors, these two dimensions were not neatly separable (see Figure 14). There were groups perceived to share common fate and not biogenetic factors, but no groups perceived to share biogenetic factors without sharing common fate. It is possible this is an artifact of the groups included in the norming process (but consider that even *Tall People* were not considered to share biogenetic factors!). Thus, though this chapter demonstrates that essentialist reasoning can occur for groups *only* perceived to share common fate, it is unclear that there is any evidence for essentialist reasoning about social groups *only* perceived to share biogenetic factors.

Both studies used an adjective checklist which included a wide array of attributes – not only attributes strongly associated with the initial or latter group – in order to let participants holistically evaluate the individual. This allowed for direct comparison between two essentialism measures: the attribute-based measure from the adjective checklist and an explicit membership-based measure of essentialism. Both studies also used on a modified switched-at-birth task, incorporating the effect of socialization into an experimental paradigm often used to study biogenetic essentialism. Having summarized the results, I now turn to more detailed consideration of (a) the relationship between the two operationalizations of essentialism, (b) the lack of support for the formative years hypothesis, and (c) participant reasoning about the process of group transition before briefly outlining future directions in this program of research.

Membership-Based and Attribute-Based Essentialism

These studies used two measures of essentialism, both similar to measures used in prior work relying on the adjective checklist measure (e.g., Eidson & Coley, 2014). In one, participants had to explicitly define the group membership of the individual – they had to accept or deny that the person had switched groups. In the other, participants had to choose adjectives from a list. I tested to see whether an adjective's association with the described individual's initial group or latter group predicted adjective choice. Using the adjective checklist, I was able to parse out whether the individual *retained* attributes of the initial group or *gained* attributes of the latter group; retention of initial-group attributes was considered an attribute-based measure of essentialism. However, these two measures of essentialism did not show the same pattern for all of the domains.

According to the membership-based measure, though some participants essentialized each of the six social domains, participants overall essentialized Race and Gender the most (see Figure 15). This is in line with other research suggesting that these social domains tend to be especially essentialized (e.g., Prentice & Miller, 2007), and by this metric it appears that groups are most essentialized when they are perceived to share both social experience and biogenetic factors. However, participants who essentialized Race, Gender, or Religion groups according to this metric – change-denying participants – did not essentialize it according to the attribute-based metric. For these domains, change-denying participants showed *less* attribute-based essentialism than change-accepting participants, instead choosing negative adjectives that were not particularly associated with either the original or the new group. The membership-based measure thus seemed to primarily index biogenetic essentialism, with biogenetically essentialist

participants expressing prejudice against individuals who did not conform to their ideas about reality (e.g., Keller, 2005).

So how should we think about the attribute-based measure of essentialism? Because I used the adjective checklist rather than a measure containing only adjectives that strongly distinguished the two groups, it was clear that both the initial group and the latter group influenced adjective selection. In five of the six social domains, the initial group indelibly impacted the individual, even as they gained latter-group attributes; they carried their old selves with them. Attribute-based essentialism appeared in the domains defined by only shared experience, and among the change-accepting participants in the Religion, Race, and Gender domains. This suggests that – at least as measured here – attribute-based essentialism primarily indexed socioculturally essentialist reasoning.

The Formative Years Hypothesis

Though the formative years hypothesis was not supported in Study 2, I do not consider this strong evidence that the hypothesis is incorrect. Instead, methodological choices of this initial test probably limited its ability to demonstrative a formative years effect. From a statistical perspective, this hypothesis relied on testing a curvilinear effect of the age at which the described individual switched groups – it is possible that the effect is simply small and was not detected by this analysis. A more critical issue may have been the way in which the sparse descriptions in Study 2 emphasized linear age, rather than specific formative experiences. Learning that someone studied Hebrew and had a b'nai mitzvah before leaving the Jewish faith, or prayed five times a day and studied the Qur'an before leaving the Islamic faith, might activate essentialism to a greater extent than simple membership duration. Such formative experiences would vary by group, and may be orthogonal to age – for instance, participating in protests may be seen as a

formative political experience, regardless of the age at which the person does this. Studies focused on the effect of specific experiences could shed more light on the formative years hypothesis.

Reasoning Beyond the Description

These studies were designed to test specific hypotheses about essentialism, and the analysis process reflects this. However, it also became clear that participants were doing more than adding together attributes of the initial and latter groups. As discussed in the Emergent Content sections of both studies, participants attributed content to the individual based on their ideas about the *switch* between groups, as well as both initial and latter groups. For Wealth Class, this meant approving of individuals moving up in socioeconomic status and disapproving of those moving down. For Gender, this meant stereotyping the individual primarily as a transgender person. For Religion, Political Ideology, and Eating Habits, this meant describing the individual as *curious* or *thoughtful* – someone who had considered their group membership before changing it. Further insight into participants ideas about the process of transition, and how these ideas are related to essentialist reasoning, could be garnered by developing and conducting a thorough qualitative analysis on the causal narratives provided by participants in both studies. This would offer insight into how participants interpreted or elaborated on the sparse descriptions, and could be connected to research on perceived locus of control (e.g., Henry et al., 2004; Weiner, 1985) and concept combination (e.g., Asch & Zukier, 1984).

Future Directions

Two methodological choices in this chapter may be worthy of further exploration: the definitions of the groups provided to participants and the labelling of the groups. Though I based the definitions on norming data, I intentionally excluded all physical markers (e.g.,

chromosomes, skin color). This was intended to create similar styles of definitions across domains. However, those definitions may have fit some domains (i.e., ones normatively defined only by common fate) better than others (i.e., ones normatively defined by common fate and shared biogenetic factors), thereby skewing participant responses. Second, the labelling of groups may have impacted the results, especially because of evidence that noun labels are connected to essentialist reasoning (e.g., Gelman & Heyman, 1999). Both of these methodological points would be worth teasing apart in future studies.

Further, though I have discussed essentialist beliefs as connected to a specific group (e.g., White people), I have primarily tested it as connected to a social domain (e.g., Race). I took this approach in order to test a wide array of social domains while using an experimental measure. However, this leaves questions about specific groups unanswered. Does sociocultural essentialism differ if one group is high-status and the other is low-status, or if two groups are defined in opposition to each other? Other literature, such as that documenting hypodescent in perception of biracial Americans, shows that perception of two groups – even two groups in the same social domain – needs to be considered in terms of their particular history and status. Living in a hierarchical society means that the (historical and current) relation between groups within a domain is important in considering psychological outcomes such as social categorization and stereotyping. Though the social domains included in these studies can be considered case studies – Wealth Class representing a status differential, and Political Ideology representing oppositional categories – research using different sets of groups would be needed to fully understand how sociocultural essentialism applies to specific groups rather than social domains. Such work should also incorporate group membership, to see whether this affects socioculturally essentialist reasoning.

Finally, it will be important to consider sociocultural essentialism outside of the experimental context. The approach and analysis in this paper focused on central tendencies among participants on Amazon Mechanical Turk – how, on average, did these participants respond to group-change scenarios, based on normative perceptions about what group members shared? In this way, I have provided evidence that groups perceived to share *only* social experience can be essentialized, at least in terms of their attributes. However, this approach cannot stand alone. Further work is needed, looking at who engages in sociocultural essentialist reasoning when, and to what ends. This will require different kinds of work (e.g., qualitative, dialectical, and observational as well as experimental) in a variety of contexts and with diverse participants (Medin et al., 2017; Nzinga et al., 2018). Chapter III is a step in this direction, considering the function sociocultural essentialism may serve among LGB+ people.

Chapter III: Sociocultural Essentialism as System-Rejection

Social essentialism – the belief that members of a social group are fundamentally the same (Gelman, 2004; Medin & Ortony, 1989) – predicts negative intergroup outcomes such as prejudice and avoidance of outgroup members (e.g., Bastian & Haslam, 2006; Kung et al., 2018; M. J. Williams & Eberhardt, 2008). However, as discussed in Chapter II, existing research has focused on biogenetic essentialism – the belief that *group-linked genes* determine group membership and cause group-typical attributes. This emphasis means that we know far less about the intergroup impact of sociocultural essentialist beliefs, and about essentialist beliefs among members of marginalized populations. In this chapter, I will argue that sociocultural essentialism may play a particular role for marginalized group members in supporting a coherent sense of the ingroup while also rejecting restrictive hegemonic stereotypes. I will then take an experimental approach to test this theory within the LGB+ population.

What Do We Know About the Essentialism-Prejudice Link?

There is a sizeable body of literature arguing that essentialism is – per se – connected to prejudicial attitudes. For example, biogenetic beliefs about race have been shown to predict lower trust and cooperation with a racial outgroup member (Kung et al., 2018), increased face-memory for own-race faces (Gaither et al., 2014), and less-diverse personal social networks (M. J. Williams & Eberhardt, 2008), as well as negative attitudes towards Black, Aboriginal, and Asian people (Jayaratne et al., 2006; Mandalaywala et al., 2018; Morton, Hornsey, et al., 2009). Biogenetically essentialist views also predict more negative attitudes towards immigrants and people with psychiatric disorders (Howell et al., 2011; Pehrson et al., 2009; Rad & Ginges, 2018; Rüsch et al., 2010; Walker & Read, 2002; Zagefka et al., 2013). Biogenetically essentialist beliefs about gender predict gender stereotypical beliefs (Brescoll & LaFrance, 2004), stereotype

threat among Canadian women (Dar-Nimrod & Heine, 2006), anti-transgender prejudice (Rad et al., 2019), sexism (Coleman & Hong, 2008; Skewes et al., 2018), and endorsements of violence against women (Mahalingam et al., 2007).

Though such evidence shows that essentialist beliefs and prejudice *can* be linked, the studies generally (a) measure biogenetic essentialist beliefs, (b) do so among members of socially advantaged groups, and (c) do so about restrictively stereotyped groups. There is some evidence that these conditions matter. For instance, Andreychik and Gill (2015) found that the essentialism-prejudice link only appeared when the essentialized group was negatively stereotyped, while Morton and Postmes (2009) found that LGB+ people's endorsement of biogenetic essentialism about sexual orientation differed based on social experience. However, we know less about the essentialism-prejudice relationship for (a) sociocultural essentialism, (b) essentialist beliefs among members of marginalized groups, and (c) ingroup essentialism.

When Does Essentialism Support the Status Quo?

Social categories can function as tools of system justification for between-category differences in the world (Jost et al., 2004; Tajfel & Forgas, 2000), especially when disadvantaged groups are biogenetically essentialized. Among members of advantaged groups, essentialist beliefs serve a system-justifying purpose: if the features of groups (including their positions in the social hierarchy) can be attributed to causal essences of the groups themselves, then there is little threat to the status quo or motivation for change (Brescoll et al., 2013). Essentialist beliefs and prejudice both support the privileged status of advantaged group members. A variety of findings support this general principle: White Americans who endorse biogenetic essentialist beliefs about race are more likely to rely on hypodescent in racial classification (Chao et al., 2013; Ho et al., 2015), accept race-based inequities (M. J. Williams & Eberhardt, 2008), and

have weaker support for social justice policies (Yalcinkaya et al., 2017); people of higher socioeconomic status who endorse biogenetic essentialist beliefs about class are more likely to supportive retributive (not restorative) educational policies (Kraus & Keltner, 2013); men who hold biologically essentialist beliefs about gender contribute less to childcare (Gaunt, 2006). Experimentally priming biogenetic essentialism about race and gender can also directly increase system-supporting ideologies (Kray et al., 2017; Mandalaywala et al., 2018).

It is not merely that essentialism can produce system-justification; system threat can also increase essentialist thinking. For example, after learning that gender inequality was declining (and thus that their comparative advantage was decreasing), Australian and British men expressed greater biologically essentialist beliefs about gender (Morton, Postmes, et al., 2009). More general system threat has also increased gender essentialist beliefs among both men and women (Brescoll et al., 2013). Among mostly White participants, Wilton et al. (2019) showed that reading about how multiculturalism is the solution to racial conflict in the US increases biogenetic essentialism. When everybody in the system is threatened, essentialist beliefs go up overall; when a group with comparative advantage feels that their advantage is threatened, members of that group become more essentialist (for more on White Americans feeling threatened by diversity, see Craig et al., 2018).

Who Does Essentialism Serve?

The system justification-essentialism link suggests that members of advantaged groups can concretize their advantage by essentializing that social dimension (Yzerbyt et al., 1997). In contrast, marginalized groups can disrupt hierarchy by arguing that the category system (and their current place in it) does not reflect a fundamental truth about category members. There is some evidence of a general tendency for social dimensions to be essentialized more by members

of advantaged than disadvantaged groups. For instance, empirical studies have shown that men (Morton, Postmes, et al., 2009; Rad et al., 2019), American boys (Smith & Russell, 1984), higher-SES children (Davoodi, 2018; del Río & Strasser, 2011), higher social-status American adults (Kraus & Keltner, 2013), higher-caste adults in India (Mahalingam, 1998, 2003), White Americans (Jayaratne et al., 2009), and German students in high-status classes (Keller, 2005) were more likely to hold biogenetic essentialist theories of the relevant social dimension than women, American girls, lower-SES children, lower social-status American adults, lower-caste adults in India, Black Americans, and German students in low-status classes, respectively. Studies using non-biological measures of essentialism have not found the same effect (see Haslam et al., 2002, for gender and Toosi & Ambady, 2011, for religion).

However, strategically de-essentializing one's marginalized ingroup may have its own costs. It could mean invalidating the meaningfulness of one's identity group, counter to the motivation for positive ingroup distinctiveness (Cross, 1995; Tajfel, 1982; for a similar argument about ingroup homogeneity see Simon, 1992). Reducing ingroup coherence can also make it harder to organize collective action and push for change to group-level inequities (Fanon, 1963; Simon & Klandermans, 2001). Thus, members of marginalized groups may maintain essentialist beliefs while rejecting negative cultural stereotypes about their group (for discussion, see Schor & Weed, 1994; Spivak, 1988). Having a coherent sense of ingroup doesn't necessarily mean passively accepting cultural stereotypes; when we are part of a group about which there are hegemonically negative or restrictive stereotypes, rejecting or resisting cultural stereotypes can be part of identity development and predict positive individual outcomes (e.g., well-being, happiness, academic success; Anyon, 1984; Hammack & Toolis, 2015; Howarth, 2002; Rogers, 2020; Rogers & Way, 2018; Way et al., 2014). Essentialist beliefs about a marginalized ingroup

may thus be "a means to assert a valued identity in the face of a majority that denies it" (Morton & Postmes, 2009, p. 656). In this sense, asserting essentialist beliefs about an ingroup may be similar identity centrality, in that both mean a given group is fundamental to how one sees the self (e.g., Leach et al., 2008). This idea is further buttressed by a documented relationship between ingroup essentialism and ingroup identification with respect to ethnicity (Verkuyten & Brug, 2004), sexual orientation (Morton & Postmes, 2009), and religion (Chalik et al., 2017), as well as the extra boost provided by a sense of belonging in an essentialized group (Bernstein et al., 2010).

How Can Essentialism Be System-Disrupting?

Though the bulk of social essentialism research has focused on biogenetic essentialism, there is growing attention to essentialist beliefs based on the perception that groups share values (Bailey et al., In Press; Newman & Knobe, 2019), childhood socialization (Rangel & Keller, 2011), or culture (Segev et al., 2012; Yalcinkaya et al., 2017). The common thread of these ideas is that groups can be essentialized if their members share social experience, or *common fate* (as discussed in Chapter II). This kind of essentialist beliefs allow for a cohesive and centrally unified understanding of a social group, including how being a group member affects individuals, while making space for change over time. Shared experiences provide rich ground for understanding people as fundamentally shaped by similar experiences, but may just as easily be system-disrupting as system-supporting.

This possibility – that sociocultural essentialism can be especially important for members of marginalized groups as a tool for system-disruption – is hinted at in existing research. For instance, members of ethnic minority groups in the Netherlands use cultural essentialism to argue against assimilationist policy (Verkuyten, 2003). Among Black and Latinx Americans,

socioculturally essentialist views about race have predicted support for affirmative action and cultural inclusion, two policies that are system-disrupting (i.e., socially progressive; Yalcinkaya et al., 2017). Within gender, some researchers contrasted Canadian women's experience of stereotype threat after reading either an article that advocated a biological explanation for gender differences in math ability or an article that advocated a sociological explanation for the same (Dar-Nimrod & Heine, 2006). Women in the latter condition experienced *less* stereotype threat, even though the general pattern of group-level differences had been reinforced. Women with more socioculturally determinist views of gender describe themselves with more positive characteristics and reject benevolent sexism (Coleman & Hong, 2008). These examples suggest that socioculturally essentialist beliefs can be specifically useful to marginalized groups.

Though people may have a baseline level of social essentialism, there is also some evidence that essentialism can be used strategically to advantage the ingroup. A study with Australian and British White populations demonstrated that when racial exclusion favored participants, prejudiced participants endorsed essentialist views; when it disadvantaged them, they did not (Morton, Hornsey, et al., 2009). When discussing the need to respect multiculturalism, members of ethnic minority groups in Rotterdam (Turkish, Moroccan, and Hindustani) advocated for cultural essentialism: expecting minorities to adopt the hegemonic culture is inappropriate because their culture fundamentally shapes who they are (Verkuyten, 2003). Ethnic Dutch individuals demonstrated the reverse pattern, where they used de-essentializing ideas to explain why ethnic minority individuals should assimilate, but essentialist ideas to explain why social inequity persisted. Essentialist arguments can be leveraged to support the interests of the ingroup; for members of marginalized groups in particular, sociocultural

essentialism allows for the rejection of current social hierarchy while maintaining positive distinctiveness about the ingroup.

Does the Kind of Marginalization Matter?

The type of marginalization a group experiences may be relevant to the affordances of biogenetic or sociocultural essentialism. Reminding LGBQ+ participants of identity-denying treatment (e.g., situations where the interests of LGBQ+ people are not considered) increased their beliefs that sexuality was immutable (including the belief that sexuality is based in biology), though reminding them of identity-devaluing treatments (e.g., situations where someone is derogated for their LGBQ+ identity) did not (Morton & Postmes, 2009). In cases of hegemonic narratives that *deny* a group's existence (e.g., LGBQ+ people, biracial people), biogenetic essentialism may serve as a means of defending the existence of one's group (e.g., "born this way"; Hegarty, 2002). In cases of hegemonic narratives that *devalue* a group's existence (e.g., Black Americans, women), socioculturally essentialist beliefs may be especially useful in allowing for a coherent and important sense of group membership without also reinforcing the "naturalness" or "rightness" of their devalued status (for a similar argument, see page 3 footnote in Yalcinkaya et al., 2017). This is because biogenetic essentialist beliefs suggest that the current state of social inequity is an inevitable product of biological forces (regardless of who holds the beliefs), while sociocultural essentialist beliefs instead affirm the coherence of a group without subscribing to inevitability.

Denial and devaluation are not neatly separable experiences. When a bisexual person hears that bisexuality is a stop on the way to homosexuality, that is denial. When they hear that bisexual people are promiscuous, that is devaluation. When they hear that bisexuality is a mental illness (as in the DSM until 1973), that includes both. While chronic marginalization of one kind

or another may produce chronic essentialist beliefs in one pattern, this is likely to occur through micro-level activation of strategic essentialism – the use of essentialist arguments to reject the present marginalization experience.

Study 3: Strategic Essentialism Among LGB+ People

In this study, I tested whether LGB+ people responded to reminders of group devaluation and group denial by increasing their levels of sociocultural and biogenetic essentialist beliefs, and whether those beliefs predicted ingroup attitudes and support for policy change. I chose this population because there is evidence that non-heterosexual people experience both devaluation and denial, such they might have strategic responses to both kinds of discrimination. This inclusion criteria omits the "T" of LGBTQ+ – though transgender people who identified as non-heterosexual were eligible to participate, heterosexual transgender people were excluded in order to keep the design specifically focused on sexual orientation.

I expected participants reminded of devaluation experiences to have a higher level of sociocultural essentialist beliefs about sexual orientation, and for that belief to strongly correlate with positive ingroup attitudes and support for policy change. I expected participants reminded of denial experiences, in contrast, to have higher levels of both sociocultural and biogenetic essentialism, with both of these predicting ingroup attitudes and support for policy change. Given the close theoretical relationship between centrality and sociocultural essentialism, and prior research suggesting that group members with higher identity centrality are more responsive to group threats (e.g., Cobb et al., 2019; Holman et al., 2021; Morton & Postmes, 2009), I expected these effects to be stronger for participants higher in pre-manipulation identity centrality.

Method

Participants. I recruited 328 non-heterosexual participants from Amazon Mechanical Turk; potential participants who were heterosexual were turned away at the screener (see Appendix D: Online Screening). Two were excluded: one indicated they did not take the study seriously, and the other failed the attention check. This leaves a final *N* of 326 (M_{age} = 34.44, SD_{age} = 10.66; 63.19% women, 27.91% men, 8.90% gender not listed; 16.56% identified as transgender; 78.53% White, 9.51% Black, 9.51% Latinx, 7.97% Asian, 3.07% Native American, 1.23% Middle Eastern, 0.92% Pacific Islander, 3.37% race/ethnicity not listed; $M_{duration}$ = 12 minutes, $SD_{duration}$ = 7 minutes). The majority of participants identified as bisexual (64.11%), with a sizable minority of gay or lesbian participants (26.07%), and smaller numbers of questioning participants (1.53%) and participants whose sexuality was not listed in the question (8.28%). Of those who noted their sexuality was not listed, 15 wrote in that they were asexual, 10 that they were pansexual, and 3 that they were queer (one participant indicated they identified as both pansexual and queer).

Measures. After completing the screener (see Appendix D: Online Screening) and consent process, I measured participants' sexual orientation identity centrality and connection with the LGBTQ+ community. Identity centrality was measured as in prior chapters, with a 0-100 slider ("Being a [specific sexual orientation group] person is an important part of how I see myself"; Leach et al., 2008). Connection with the LGBTQ+ community was measured separately to incorporate participants' sense of belonging within the broader non-heterosexual community (three questions on a 0-100 scale; α = 0.94, drawn from Frost & Meyer, 2012; see Table 21; as this correlates highly with centrality at $r(324) = 0.60$, $p < 0.0001$, I focus on centrality in my analyses for consistency with prior chapters).

Table 21. Community connection scale.

(1) I feel I am part of the LGBTQ+ community. (2) Participating in the LGBTQ+ community is a positive thing for me. (3) I feel a bond with the LGBTQ+ community.

Participants were then randomly assigned to one of three conditions: group devaluation, group denial, and control (these are closely based on those used by Morton and Postmes, 2009). For each experimental condition, the participant read examples of negative treatment towards LGB+ people in general, and were invited to list any examples they could think of which represented similar treatment (see Table 22 and Table 23; specific examples inspired by: Garr-Schultz & Gardner, 2021; Gruberg et al., 2020). On the screen after the manipulation, participants in the devaluation and denial conditions were asked to summarize what they had read on the prior screen ("On the last screen, you read some information about how people are treated in society. What did you read about?"). This was intended to encourage participants to engage with the prompt longer.

Table 22. Group Devaluation Manipulation

In this study, we are interested in how people experience discrimination, or **situations in which people receive negative treatment from others because of who they are**. Some examples of this are: (1) LGB+ people being told they're disgusting. (2) Employers preferring to hire heterosexual people rather than LGB+ people. (3) Doctors refusing to treat LGB+ people. (4) LGB+ people being excluded from religious communities. Can you think of any examples that reflect this kind of treatment? Please write about them below.

Table 23. Group Denial Manipulation

In this study, we are interested in how people experience marginalization, or **situations in which people are ignored by others or treated as if they are invisible**. Some examples of this are: (1) LGB+ people being told that their sexual orientation is just a phase. (2) Employers telling LGB+ people not to discuss their romantic lives at work, even though heterosexual people can do so. (3) Birth certificates only having space to list one mother and one father. (4) LGB+ people being told their sexual orientation is a lifestyle choice. Can you think of any examples that reflect this kind of treatment? Please write about them below.

After the manipulation, participants completed three subscales of the Sexual Orientation Beliefs Scale, Form 1: naturalness, entitativity, and discreteness (SOBS, Arseneau et al., 2013). The naturalness and entitativity subscales map closely onto biogenetic and sociocultural essentialist beliefs as I have been discussing them up to this point, and the discreteness scale was included for potential exploratory analyses. SOBS was developed to measure the beliefs of LGBTQ+ people about sexual orientation, making it better-suited for this study than other related measures (e.g., N. Haslam & Levy, 2006; Hegarty, 2002; Keller, 2005; Rangel & Keller, 2011; Yalcinkaya et al., 2017) Participants indicated their agreement or disagreement with each statement using 1 – 9 scale (Strongly Disagree – Strongly Agree), and all questions were presented in a random order ($\alpha_{\text{naturalness}} = 0.79$; $\alpha_{\text{entitativity}} = 0.84$; $\alpha_{\text{discreteness}} = 0.78$; see Table 24).

Table 24. Sexual Orientation Beliefs Scale (Arseneau et al., 2013). These were presented to participants as part of the same measure, with question order randomized. Asterisk indicates the item is reverse-scored.

Naturalness
(1) Sexual orientation is innate
(2) Individuals choose their sexual orientation*
(3) Biology is the main basis of an individual's sexual orientation
(4) Social and environmental factors are the main basis of an individual's sexual orientation*
(5) People have control over changing or keeping their sexual orientation*
(6) Something deep inside of a person determines their sexual orientation
(7) The existence of different sexual orientations is natural

(8) If someone comes out as gay or lesbian they were probably attracted to the same sex all along
(9) The percentages of people in different sexual orientation groups are roughly the same all over the world
(10) It is impossible to truly change one's sexual orientation
(11) The idea that individuals have a "sexual orientation" is a social invention*
(12) Sexual orientation is set early on in life

Entitativity

(1) Individuals with the same sexual orientation seem to be connected to one another by some invisible link
(2) People who have the same sexual orientation are very similar to one another.
(3) There are more similarities than differences among people who have the same sexual orientation
(4) It is usually possible to know about many aspects of a person once you know their sexual orientation
(5) It is usually possible to know a person's sexual orientation even without being told
(6) People tend to have a sense of group belonging based on their sexual orientation
(7) People who share the same sexual orientation pursue common goals
(8) Knowing a person's sexual orientation tells you a lot about them
(9) People who have the same sexual orientation interact frequently with one another
(10) People with the same sexual orientation share a common fate

Discreteness

(1) Sexual orientation is a category with distinct boundaries: a person is either gay/lesbian or heterosexual
(2) Sexual orientation is a category with clear boundaries: a person is either gay/lesbian, bisexual, or heterosexual
(3) People who identify as bisexual are confused about their true sexual orientation
(4) A person has only one true sexual orientation
(5) It is possible to be "partially" or "somewhat" gay or straight*
(6) People may reasonably identify as two sexual orientations at the same time*

Participants then completed three measures in a randomized order. Participants reported their affect towards a variety of sexual orientation groups (asexual, bisexual, gay, heterosexual, homosexual, lesbian, pansexual, queer, straight) using the feeling thermometer; this included attitudes towards heterosexual people for exploratory analysis. Participants indicated whether they supported eight federal policies intended to advance equity around sexual orientation, currently supported by the Human Rights Campaign (see Table 25; HRC, 2021), as well as their

support for governmental action in general (1-7 scale, “The government should spend resources and time protecting the rights and well-being of LGB+ people”). These ratings of ingroup warmth and support for federal action were the central dependent variables for the study. Participants also indicated their belief in the existence of sexual-orientation-based inequity, on a 0-100 scale measure (α = 0.87; based on Kelly & Breinlinger, 1995; see Table 26) and by making numerical estimates (based on Kraus et al., 2017; see Table 27). These measures were included to test if the experimental manipulation impacted participants’ overall sense of group-level disadvantage in society.

Table 25. Policy support items drawn from the Human Rights Campaign (HRC, 2021).

You will now see short descriptions of federal policy proposals that are intended to reduce inequity based on sexual orientation. For each one, please say whether you support or do not support the proposal.

(1) The Equality Act would clarify that prohibitions against sex discrimination include discrimination based on sexual orientation, and further prohibit such discrimination in public spaces, public services, and all federally funded activities.
□ I support this policy
□ I do not support this policy

(2) The Every Child Deserves a Family Act would prohibit any child welfare agency receiving federal financial assistance from discriminating against any potential foster or adoptive family on the basis of sexual orientation, as well as discriminating against any foster youth on the basis of their sexual orientation.
□ I support this policy
□ I do not support this policy

(3) The Fair and Equal Housing Act would provide consistent and explicit non-discrimination protections for LGBQ+ people by adding “sexual orientation” as a protected characteristic under the Fair Housing Act.
□ I support this policy
□ I do not support this policy

(4) The Jabara-Heyer NO HATE Act would promote better hate crime data collection for prevention purposes across the federal, state, and local levels.
□ I support this policy
□ I do not support this policy

(5) The Jury Non-Discrimination Act would prohibit attorneys from removing potential federal jurors based on their sexual orientation.
□ I support this policy
□ I do not support this policy

(6) The PrEP Access and Coverage Act would require all private and public insurance plans to cover the HIV prevention pill and related services with no out-of-pocket costs for patients.
□ I support this policy
□ I do not support this policy

(7) The Safe Schools Improvement Act would require school districts in states that receive federal funds to adopt codes of conduct prohibiting bullying and harassment on the basis of sexual orientation.
□ I support this policy
□ I do not support this policy

(8) The Paycheck Fairness Act would require employers to disclose worker compensation information, and ban employers from asking potential employees about their salary history.
□ I support this policy
□ I do not support this policy

Table 26. Scale measure of perception of social-orientation-based inequity.

(1) In terms of power and status in society, heterosexual people get a better deal than bisexual, gay, lesbian, pansexual, and queer people.
(2) LGB+ people as a group deserve a better deal in society.
(3) It makes me feel angry that heterosexual people in general have higher status than bisexual, gay, lesbian, pansexual, and queer people.

Table 27. Numerical estimates of social-orientation-based inequity.

(1) Of the 535 members of the US Congress, how many do you think are bisexual, gay, lesbian, pansexual, or queer?
(2) In the US, for every $1.00 earned by a heterosexual person, how much do you think a bisexual, gay, lesbian, pansexual, or queer person earns?

Finally, participants reported their demographics; there was an attention check embedded in the demographics section, and after completing their demographics participants completed the seriousness check (see Appendix F: Attention Check and Seriousness Check).

Primary Results

The central question of this study was whether asking LGB+ participants to engage with ideas about group *denial* and group *devaluation* would produce distinct patterns of biogenetic and sociocultural essentialism. I expected group devaluation to increase sociocultural essentialism relative to the control, and for sociocultural essentialism in this condition to predict positive ingroup attitudes and support for federal policy change. I expected group denial to increase both sociocultural and biogenetic essentialism relative to control, and for both types of essentialist beliefs to predict positive ingroup attitudes and support for policy change.

All results were analyzed using a regression approach, with two orthogonal contrast codes for conditions. The first contrast code compared the control condition with both experimental conditions, while the second compared the devaluation and denial conditions to each other. In analyses with significant effects of condition, I used follow-up tests to determine which of the experimental conditions differed significantly from the control condition. I report standardized betas as the effect size, along with *p*-values for the specific predictor or interaction. All graphs display estimated marginal means and standard errors from model analyses.

Levels of Biogenetic and Sociocultural Essentialism. Contrary to expectation, the experimental manipulation did not shift participants' levels of biogenetic or sociocultural essentialist beliefs (ps > 0.2). In Morton & Postmes (2009), the prompt's effect on biogenetic essentialism depended on (pre-assessed) participant identity centrality. I did not replicate this; participants' pre-assessed identity centrality did not impact the effect of condition on either kind of essentialism (ps > 0.1). I did replicate a general tendency for identity centrality to predict higher biogenetic essentialism across conditions ($\beta = 0.18$, $p = 0.002$), and found that centrality also predicted sociocultural essentialism ($\beta = 0.23$, $p < 0.0001$). Participants endorsed biogenetic

essentialism more than sociocultural essentialism ($t(325) = -25.789$, $p < 0.0001$; $M_{biogenetic} = 6.48$, $M_{sociocultural} = 3.98$).

Policy Support and Ingroup Attitudes. I next tested whether the manipulation had impacted support for specific federal policies, federal action in general, and ingroup attitudes. Participants overwhelmingly supported all eight federal policy proposals; the Safe Schools Improvement Act had the highest support (96.93%), and the Paycheck Fairness Act had the lowest (90.18%). For each participant, I define their specific policy support as the percentage of policies they indicated they support ($M_{policysupport} = 94.70\%$, $SD_{policysupport} = 12.30\%$). Participants also expressed high levels of support for federal action on LGB+ rights ($M_{federal} = 6.18$, $SD_{federal} = 1.16$; significantly higher than the scale midpoint, $t(325) = 33.908$, $p < 0.0001$), and very positive attitudes towards their sexual orientation ingroups ($M_{ingroup} = 80.25$, $SD_{ingroup} = 18.71$; significantly higher than the scale midpoint, $t(288) = 27.479$, $p < 0.0001$). The high values of these dependent variables mean that there will be ceiling effects in the remaining analysis.

Specific policy support differed by condition ($M_{denial} = 91.93\%$; $M_{devaluation} = 97.22\%$; $M_{control} = 94.95\%$); participants in the devaluation condition supported a greater proportion of policies than participants in the denial condition ($\beta = 0.18$, $p = 0.001$; see Figure 19). Support for general federal action, ingroup warmth, and all three inequality perceptions did not differ by condition ($ps > 0.14$). This suggests that specific policy support was more responsive to the manipulation than the other measures, though this should be interpreted with caution because of ceiling effects.

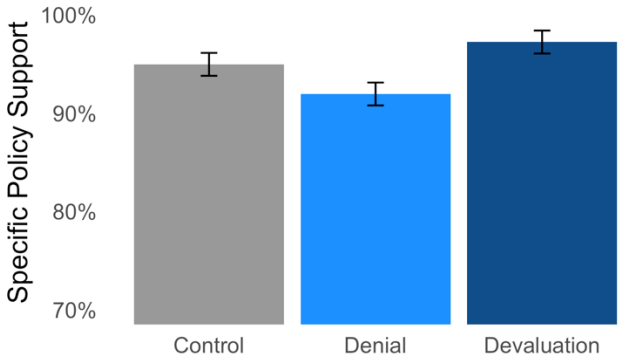

Figure 19. Participants in the devaluation condition supported more specific policies than participants in the denial condition.

Essentialism Predicting Policy Support and Ingroup Attitudes. Though the experimental manipulation did not directly influence participants' levels of biogenetic and sociocultural essentialism, I had also predicted that these types of essentialism would differ by condition in their connection to the dependent variables. I expected sociocultural essentialism to predict all three dependent variables in the devaluation condition, and both biogenetic and sociocultural essentialism to predict them in the denial condition. As it is possible that the manipulation affected the interpretation or meaning of essentialist beliefs without shifting their absolute levels, I proceeded to test whether biogenetic and sociocultural essentialism predicted specific policy support, general policy support[8], and ingroup attitudes, including whether those relationships differed by condition. For each of the dependent variables, I tested a regression

[8] As both indices of support for change are about federal policy in the US, one possibility was that political ideology would affect the pattern of results. I measured political ideology with two items in the demographics section, asking participants to endorse their overall agreement with conservative and liberal political ideology on a 1-7 scale. These are strongly negatively correlated ($r(323) = -0.67$, $p < 0.0001$), so I collapsed them into a single 1-7 metric in which higher values indicate more liberal views. In general, the sample was more liberal than conservative ($M_{politicalideology} = 5.73$; significantly higher than the scale midpoint, $t(324) = 22.305$, $p < 0.0001$). Political ideology predicted specific policy support ($\beta = 0.44$, $p < 0.0001$) and general policy support ($\beta = 0.62$, $p < 0.0001$), such that more liberal participants were more supportive. However, controlling for political ideology did not change the overall pattern of results.

model with both biogenetic and sociocultural essentialism independently predicting the DV and interacting with condition in predicting the DV.

Biogenetic essentialist beliefs predicted specific policy support ($\beta = 0.22$, $p < 0.0001$; see Figure 20), though this relationship was blunted by high overall support in the devaluation condition (denial: $\beta = 0.30$, $p = 0.002$; control: $\beta = 0.23$, $p = 0.02$; devaluation: $p = 0.57$). Across conditions, biogenetic essentialism predicted general support for federal action ($\beta = 0.26$, $p < 0.0001$) and positive ingroup attitudes ($\beta = 0.12$, $p = 0.04$). The effect of sociocultural essentialism did not differ by condition – it predicted positive ingroup attitudes ($\beta = 0.17$, $p = 0.003$) and did not predict specific either measure of policy support ($ps > 0.25$, see Figure 20).

Thus, biogenetic and sociocultural essentialism seemed to play different roles in how LGB+ people related to their ingroup, though this is a correlational rather than experimental finding. Regardless of being prompted to consider ingroup denial or devaluation experiences, sociocultural essentialism was linked to positive ingroup attitudes and biogenetic essentialism was linked to stronger support for federal action.

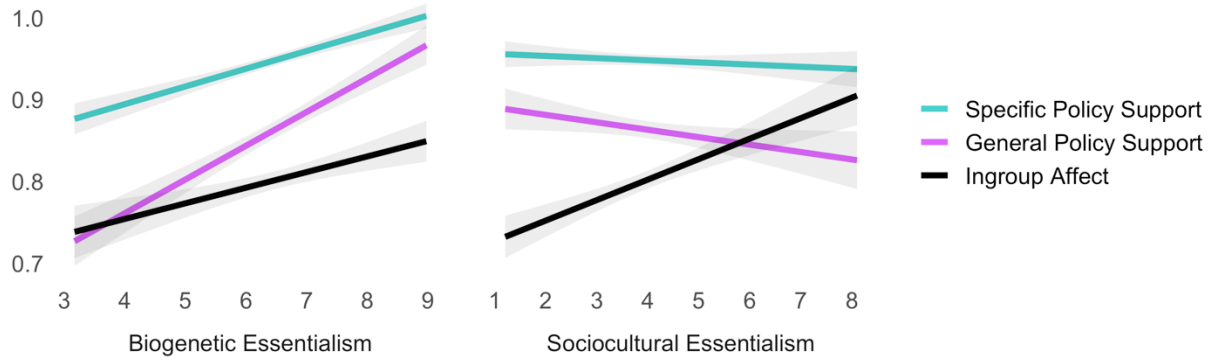

Figure 20. Biogenetic essentialism strongly predicted policy support and weakly predicted ingroup attitudes. Sociocultural essentialism predicted ingroup attitudes but not support for policy. For presentation, I used a linear conversion to put every variable on a 0-1 scale.

Exploratory Results

Not all LGB+ people experience (or have experienced) the same treatment in American society. To that end, I conducted exploratory analyses by comparing gay and lesbian participants with all other participants (bisexual, questioning, or not listed; I use the term "queer" to refer this group in analyses), and using participant age as a linear predictor (ranging from 18 – 71). I report all cases where these demographic factors affected the pattern of results in the Primary Results section. I also compared men with women and trans* with non-trans* participants, but these group memberships did not significantly alter the pattern of results in the prior section. As I did not recruit for balance on these demographic factors, all of these analyses should be recognized as exploratory and I was unable to meaningfully test for interactions (e.g., older queer participants compared to younger gay ones).

Gay/Lesbian and Queer Participants. Gay/lesbian people endorsed sociocultural essentialism more than queer people ($\beta = -0.17$, $p = 0.002$, $M_{gay} = 4.35$, $M_{queer} = 3.84$; the groups did not differ in biogenetic essentialism, $p = 0.14$). Identity centrality also predicted sociocultural essentialism more strongly for gay/lesbian participants than queer participants (gay/lesbian: $\beta = 0.43$, $p < 0.0001$; queer: $\beta = 0.15$, $p < 0.01$); this was especially pronounced in the control condition relative to the experimental conditions ($\beta = -0.77$, $p = 0.02$; see Figure 21). These patterns may suggest that the measure of sociocultural essentialism mapped more clearly onto how gay/lesbian participants were thinking about sexual orientation than how queer participants were.

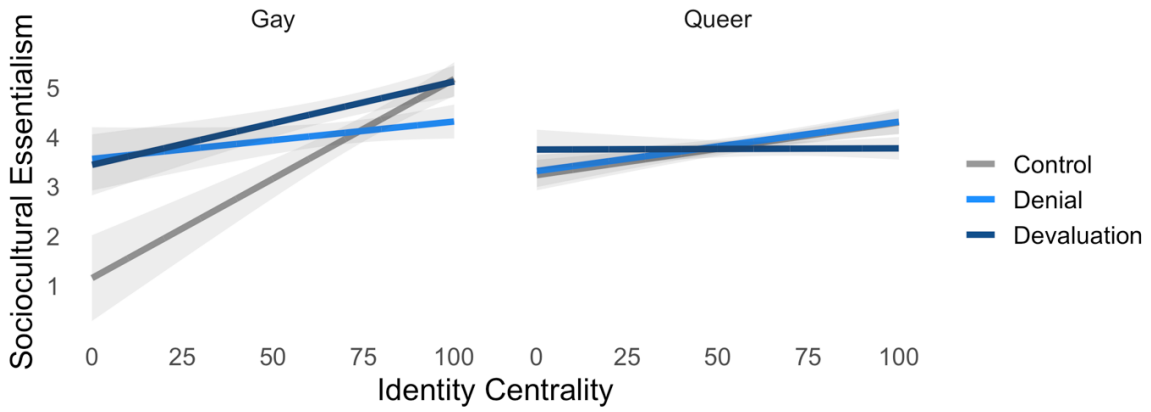

Figure 21. For gay/lesbian participants, identity centrality predicted sociocultural essentialism more strongly in the control than experimental conditions; this pattern did not emerge for queer participants.

Participant Age. Older participants endorsed biogenetic essentialism more than younger people ($\beta = 0.20$, $p = 0.0005$; age did not impact sociocultural essentialism, $p > 0.2$). Older participants perceived less sexual-orientation based inequity and thought the wage gap was smaller than younger participants (inequity perception: $\beta = -0.11$, $p < 0.05$; wage gap: $\beta = 0.12$; $p < 0.05$). Age did not impact estimates of congressional representation ($p = 0.7$) or general support for federal policy ($p > 0.3$).

In several cases, the denial and devaluation conditions produced different effects for older and younger participants. First, these two conditions predicted different levels of sociocultural essentialism according to the interaction of age and identity centrality ($\beta = 1.41$, $p = 0.003$). Among younger participants, identity centrality predicted sociocultural essentialism in the denial condition, but not in the devaluation condition. Among older participants, this pattern was reversed – identity centrality positively predicted sociocultural essentialism in the devaluation condition, but not in the denial condition (see Figure 22).

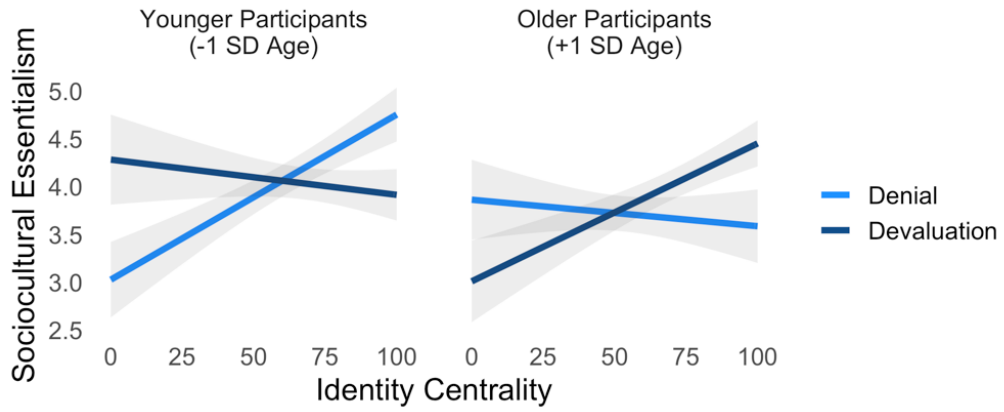

Figure 22. The effect of condition on level of sociocultural essentialism depended on the interaction between age and identity centrality. For younger participants, centrality predicted sociocultural essentialism in the denial condition; for older participants, centrality predicted sociocultural essentialism in the devaluation condition.

Second, younger participants supported more specific federal policies than younger participants ($\beta = -0.12$; $p = 0.03$), but this differed significantly between the denial and devaluation conditions. In the devaluation condition, participants strongly endorsed specific federal policies regardless of age. In the denial condition, younger participants endorsed specific federal policies more than older participants ($\beta = -0.19$, $p < 0.05$; see left panel Figure 23).

Third, while age did not have an overall effect on ingroup warmth ($p = 0.9$), this differed between the denial and devaluation conditions ($\beta = 0.54$, $p = 0.005$). In the denial condition, younger participants expressed more ingroup warmth than older participants; in the devaluation condition, younger participants expressed less ingroup warmth than younger participants (see right panel Figure 23).

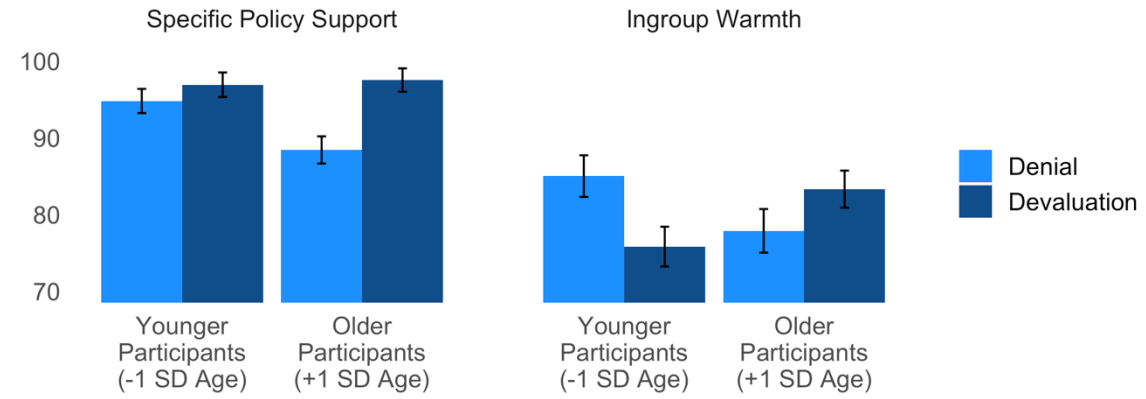

Figure 23. Older and younger participants reacted differently to the denial and devaluation conditions. In the denial condition, age predicted lower specific policy support and lower ingroup warmth. In the devaluation condition, age predicted greater ingroup warmth.

Finally, age played a role in how condition impacted the relationship between sociocultural essentialism and ingroup warmth ($\beta = -1.40$, $p = 0.04$). For younger participants, sociocultural essentialism predicted ingroup warmth only in the experimental conditions; for older participants, sociocultural essentialism predicted ingroup warmth regardless of condition (see Figure 24).

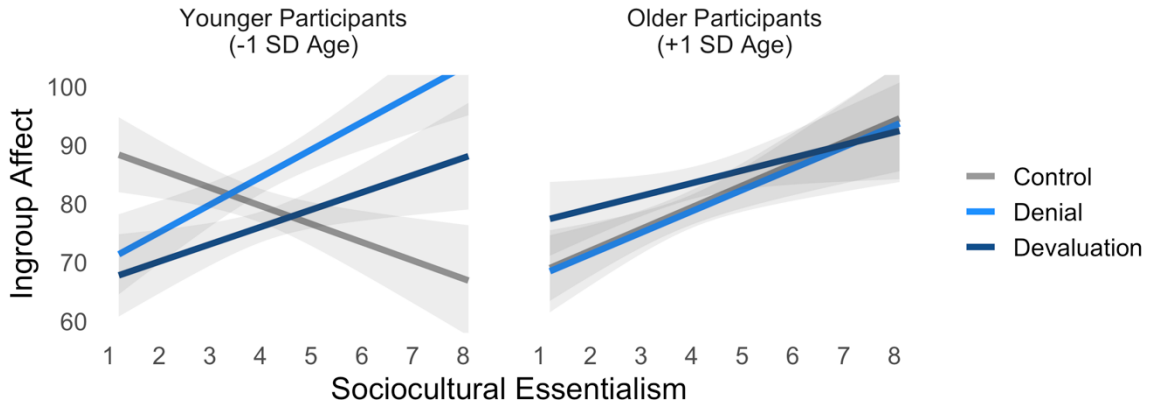

Figure 24. For older participants, sociocultural essentialism consistently predicted ingroup warmth. For younger participants, it did so in the experimental conditions but not the control condition.

Discussion

The central aim of this experiment was to activate strategic essentialism by asking participants to engage with prompts about group denial and group devaluation. However, the experimental manipulations did not alter participants' endorsement of either biogenetic or sociocultural essentialism. Identity centrality predicted both kinds of essentialist belief, regardless of condition. In general, participants strongly supported both specific and general federal action to support LGB+ rights, and felt warmly about their sexual orientation ingroups. Across conditions, biogenetic essentialism predicted support for policy change and ingroup warmth, while sociocultural essentialism predicted ingroup warmth without predicting support for policy change. Exploratory analyses suggested that younger and older participants experienced the denial and devaluation conditions differently from each other. In the next several sections, I explore the implications of these findings in greater detail.

Activating Essentialism – Strategic and Sociocultural

Though many studies have successfully manipulated essentialist beliefs, the manipulation used in this study did not produce a corresponding difference in either sociocultural or biogenetic essentialism. This could mean that feelings of devaluation or denial occurred without provoking strategically essentialist responses, or that this manipulation was too weak. I did not directly assess participants' feelings of denial or devaluation – either about themselves or about the broad LGB+ ingroup. However, as the manipulation did not change participants' perceptions of inequity, it seems likely that the manipulation was too weak. A thorough qualitative analysis of the written prompt responses could offer post hoc insight into how participants were understanding the prompts, but would not offer a clear manipulation check as there was no such question in the control condition.

Another possibility is that essentialist beliefs in this population were too stable for the manipulation to have an effect. In Morton & Postmes (2009), the same manipulation did shift biogenetic essentialism among high-centrality LGB+ people in Europe. As there is some evidence that essentialist beliefs are more connected to pro-LGBTQ+ attitudes in the US than elsewhere (Hegarty, 2002), US participants' essentialism may be more stable. Our sample may also have differed in that I recruited from Amazon Mechanical Turk while they relied on community mailing lists and bulletin boards – their population may be more enmeshed in an LGB+ community, and thereby more likely to be affected by reminders of group-level discrimination. More broadly, published manipulations of essentialist belief have mostly focused on biogenetic essentialism (e.g., Dar-Nimrod & Heine, 2006). Though there is non-experimental work suggesting that individuals can strategically use sociocultural essentialism during conversation (Verkuyten, 2003), it is unclear whether short experimental manipulations can have the same effect.

Ingroup Warmth and Support for Change

Though the experimental manipulation did not affect levels of biogenetic and sociocultural essentialism, I did find correlational patterns showing differential impact of the two kinds of belief among LGB+ people. Biogenetic essentialism predicted both policy support and ingroup warmth, and sociocultural essentialism uniquely predicted ingroup warmth. This difference emerged even though both kinds of essentialism were strongly predicted by identity centrality. Essentialism may thus moderate the relationship between identity centrality and collective action more broadly; feeling connected to the group matters, but your ontological understanding of the group explains what you want to *do* with that feeling.

The role of biogenetic essentialism in predicting support for policy change is consistent with other research showing that biogenetic essentialism about sexual orientation in particular can be connected to pro-LGB+ attitudes (e.g., Hegarty, 2002; Hegarty & Pratto, 2001; Huic et al., 2018). However, it may be a product of the "born this way" argument for LGBTQ+ rights that is prevalent in the US, wherein emphasizing the biological source of sexual orientation is used as a basis for claiming civil rights. In cultures where this argument is less prevalent (e.g., the UK, Hegarty, 2002), or social domains where it is less prominent (e.g., race in the US; see Yalcinkaya et al., 2017), biogenetic essentialism is unlikely to play this role.

The lack of correlational relationship between sociocultural essentialism and support for policy change in this LGB+ sample directly contrasts evidence that such beliefs predict support for policy change among members of minoritized racial groups (Yalcinkaya et al., 2017). This may reflect a domain-level tradeoff between the two kinds of essentialist belief; at a cultural level, if endorsement of biogenetic essentialism indicates support for collective action and change in the LGB+ domain, then sociocultural essentialism does not. It is worth considering how such beliefs function in other types of social hierarchy (e.g., gender, nationality, class), given that sexual orientation is relatively unique in having a documented positive relationship between biogenetic essentialism and support for minority-group rights.

Population Heterogeneity

Exploratory analyses focused on demographic subgroups of participants revealed remarkable consistency – trans* and non-trans* participants showed the same patterns of results, as did men and women. There were a few differences between gay/lesbian participants and bisexual, questioning, pansexual, queer, and asexual participants. Gay/lesbian participants endorsed sociocultural essentialism more than other participants, and sociocultural essentialism

was connected to identity centrality for gay/lesbian participants (especially in the control condition), but not for other participants. Though these were the only cases where these groups differed, it may be that the SOBS entitativity items (e.g., "People tend to have a sense of group belonging based on their sexual orientation", see Table 24) were more accessible to gay/lesbian participants than bisexual, questioning, pansexual, queer, and asexual participants. These participants are in less-visible subgroups of the LGB+ population and more likely to experience identity denial (Garr-Schultz & Gardner, 2021). This could reduce their sense of personally sharing experience with other members of their ingroup, or could reduce their belief that sexual orientation structures individuals' social experiences even beyond their own ingroup.

The other demographic factor that meaningfully impacted results – across an array of analyses – was age. Older participants endorsed biogenetic essentialism more highly and perceived less sexual-orientation-based inequity than younger participants. Older and younger participants also reacted differently to the denial and devaluation conditions. Among younger participants, the *denial* condition increased ingroup warmth and activated an identity centrality-sociocultural essentialism link. Among older participants, the *devaluation* condition increased ingroup warmth, increased specific policy support, and activated the identity centrality-sociocultural essentialism link (see Figure 23 and Figure 24). That is, the denial condition seemed to be a potent threat for the younger participants, while devaluation was a potent threat for the older participants.

There are possible interpretations of these findings, which I discuss below. However, as this study was not designed to focus on age, it offers little opportunity to distinguish between them and they remain speculative. Broadly speaking, these results could be framed in terms of

participants' current developmental stages or participants' sociohistorical experiences as members of the LGB+ community (e.g., greater general acceptance of LGB+ people over time).

If we consider the results in terms of developmental stages, the denial and devaluation threats may carry differential weight for older and younger participants because of what they care most about given their ages. Younger participants are more likely to be still sorting out their sense of self, establishing their own understanding of their sexual orientation and how it affects their relationship to the rest of society (e.g., Hammack et al., 2021). For this reason, denial may be a particularly potent reminder for younger participants, increasing policy support and leading high-centrality participants to greater endorsement of sociocultural essentialism. However, identity development does not occur at the same rate for everybody; older and younger participants alike may be learning more about themselves. For groups with concealable stigmatized identities, like LGB+ people, this internal process may also be affected by when and how people "come out" to close (and distant) others.

Another difference between the younger and older groups is their experience of LGB+ rights and acceptance in America (Fetner, 2016). The first pride parade, building on decades of LGB+ activism, was in 1970. Homosexuality as a psychiatric disorder was only removed from the DSM in 1973 – 13.41% of my participants were alive for that. Anti-sodomy laws were judged unconstitutional in 2003. Same-sex marriage was legalized throughout the US in 2015. This trajectory means that older participants have likely experienced – at the level of national culture – more straightforwardly negative attitudes towards LGB+ people than younger participants. This could explain why reminders of devaluation were particularly potent for older participants, and activated sociocultural essentialism for high-centrality participants. However, progress is not homogeneous (e.g., Kaufman & Compton, 2021; Kite & Bryant-Lees, 2016).

LGB+ acceptance varies by location (as big as region, as small as household), by intersection with other social identities (e.g., queer people who are also Black or Hispanic or Asian), by class, by specific environment (e.g., a church vs. a bar), and by subgroup (e.g., gay/lesbian people vs. bisexual people). Without more information about participants' personal experiences, I cannot make nuanced conclusions about participants' personal experiences with LGB+ discrimination over their lifespans to date.

In all likelihood, the effects obtained here are a combination of all these general frameworks (e.g., Hammack, 2005), and may further reflect the narrative processes participants use to understand themselves and their own sexual orientations with respect to cultural narratives about what those identities mean (e.g., Cohler & Hammack, 2006). Further work, especially explicitly developmental work that digs into how sociocultural essentialism develops alongside LGB+ identity, would be needed to make strong conclusions about these exploratory findings.

Future Directions

Beyond the failure of the manipulation, several other methodological choices may have hampered this study. First, the extremely high support for federal policy change introduced ceiling effects into the analysis, hampering interpretation. Future research using dependent variables that are less subject to ceiling effects would be informative – for instance, asking participants to allot 100% of the federal government budget to different areas, with one of those areas being LGB+ rights and well-being, would probably reduce ceiling effects and allow for a more holistic sense of participants' priorities. Second, though other studies have shown evidence of strategic essentialism using scale measures (e.g., Morton, Postmes, et al., 2009), measures focused on the communication of ontological beliefs may be more fruitful, emphasizing the social *use* of those beliefs rather a participant's personal endorsement of them. The dialect

analysis in Verkuyten (2003) is one example of analysis focused on social interaction; this could also be approached in survey form (i.e., “what do you wish the general public knew about LGB+ people?”).

Chapter IV: What's Learned, What's Left

In this chapter, I will review the findings of the previous three chapters and discuss ongoing questions and future directions raised by the described research. In this dissertation, I set out to explore *whether* and *why* people would use essentialist reasoning about groups perceived to share social experience. This aim grew out of two apparent paradoxes in the literature. First, though the literature on social essentialism has often focused on groups perceived to be biological (i.e., race and gender) or on the biological transmission of group essence (e.g., Gil-White, 2001), this causal specificity was more limited than the "placeholders" described in the original theory (Medin & Ortony, 1989) and seemed contrary to both developmental findings that essentialism is tuned by cultural input (e.g., Davoodi et al., 2019; Smyth et al., 2017) and work in other areas showing that laypeople can have complex and multiply-causal mental representations of the world (e.g., Condit, 2019; Kanovsky, 2007; Martin & Parker, 1995; Ryazanov & Christenfeld, 2018). If essentialism is a basic cognitive tool, why would it be limited to the perception of nominally biogenetic social groups?

Second, though some essentialism researchers suggested that members of marginalized groups should de-essentialize their ingroup in order to reject social hierarchy (e.g., Yzerbyt et al., 1997), this ran counter to identity development models emphasizing the importance of maintaining a coherent and positively distinctive sense of one's ingroup, especially for members of marginalized groups (e.g., Cross, 1995). It seemed possible that a *sociocultural* essentialism – grounded in the causal lay theory that people are indelibly shaped by their group-linked experiences – could square these circles. This was buttressed by a recently growing body of literature demonstrating causal heterogeneity in essentialist reasoning about social categories

(e.g., Bailey et al., In Press; Newman & Knobe, 2019; Rangel & Keller, 2011; Segev et al., 2012; Yalcinkaya et al., 2017).

Brief Summary of Chapters

As a starting point, I needed a measure of stereotype content that was equally useful across a wide array of social domains – groups perceived to share social experience and groups perceived to share biogenetic factors. Existing studies tend to measure the retention of essentialized traits in a single social domain (e.g., gender), and in terms of traits that strongly differentiate two groups in a domain (e.g., traits that are uniquely feminine or uniquely masculine). I wanted to use the same measure across a variety of domains, incorporate the possibility that individuals are perceived both to *retain* essentialized trains and *gain* new ones, and enable my participants to respond relatively holistically. The Adjective Checklist measure was well-suited to these aims, except that the most common set of adjectives had been developed specifically to measure racial stereotypes and only slightly updated since 1933 (Galinsky et al., 2013; Katz & Braly, 1933). In Chapter I, I therefore developed an updated version of the measure that accommodates groups from an array of social domains. I combined online and lab-based adjective sources and mathematically defined the characteristics an adjective checklist should have. I conducted a norming study and used an optimization process to select a set of 70 adjectives for inclusion in a modern checklist measure.

Having developed a suitable measure, I used it in to experimentally test sociocultural essentialism in Chapter II. In that chapter, I fully developed a response to the first paradox, describing how the perception of common fate – conceptually close to the perception of entitativity – could ground essentialist reasoning. I further described two expected boundary conditions: (a) membership duration should predict the extent to which a sociocultural essence is

imparted, and (b) membership during developmentally formative years should impart essence to a greater extent than membership during non-formative years. I used a norming process to define a set of six social domains of interest – Gender, Race, Religion, Political Ideology, Wealth Class, and Eating Habits. The first three were perceived to share both social experience and biogenetic factors, while the latter three were perceived to share social experience but not meaningful biogenetics. Interestingly, there were no groups in the norming data perceived to meaningfully share biogenetic factors *without* sharing social experience. This may imply that studies interpreted as demonstrating biogenetic essentialism actually demonstrate sociocultural essentialism, or essentialism grounded in multiply causal reasoning (similar to a suggestion in Yzerbyt, Estrada, et al., 2004).

Study 1 used an updated switched-at-birth task in which participants read about an individual who, after a long time as a member of an initial group, were now a member of a latter group. Participants were asked to choose traits to describe that individual, and judge their group membership. Saying the person was still an initial-group-member was one index of essentialism, and it appeared more strongly in the social domains defined by the perception of both shared biology and shared social experience. Attributing the person traits based on their association with the initial group was a second index of essentialism, and it appeared in domains perceived to share only social experience as well as domains perceived to share both social experience and biogenetic factors. Because I used the adjective checklist to measure stereotype content, I could see that the switched-group individual was perceived to retain initial group traits while also being attributed latter-group traits. Gender and Wealth Class complicated the results, with ad hoc analyses suggesting that participants' ideas about the *process* of switching groups (e.g., by identifying as a transgender person, or by being praiseworthy or blameworthy) obscured

essentialist responding. Study 2 slightly modified the Study 1 paradigm, incorporating information about when the individual switched social groups. For both Religion and Wealth Class, switching groups at an older age predicted higher perceived retention of initial-group membership and attributes – the membership duration hypothesis was supported. However, the results did not support the formative years hypothesis. Together, these studies offered support for the theory of sociocultural essentialism; on average, participants considered initial-group membership an important determiner of an individual's characteristics after that person had left the group, even if the group was defined only by shared social experience. In response to the first paradox, this suggests that the field has overemphasized biogenetic essentialism and joins the chorus of recent work on causal heterogeneity in social essentialism.

In Chapter III, I moved to the second paradox, testing if members of marginalized groups would strategically recruit sociocultural essentialism in response to ingroup devaluation, or biogenetic essentialism in response to ingroup denial. I conducted an experimental study within the LGB+ population in the US, because LGB+ people face both devaluation and denial in society today, and thus seemed more likely than other marginalized group members to have strategic responses to both kinds of discrimination. I did not find the expected effects, failing to replicate prior research where the manipulation's altered endorsement of biogenetic essentialism among non-heterosexual people. Despite this, correlational analyses indicated that biogenetic essentialism uniquely predicted support for federal policy change while sociocultural essentialism uniquely predicted ingroup warmth. Exploratory analyses on demographic factors also suggested that sociocultural essentialism might be a strategic response to devaluation for older LGB+ people, but to denial for younger LGB+ people. Though this chapter did not provide conclusive evidence to resolve the second paradox, it opened up new questions about how the

role of sociocultural essentialism could differ by individual life-course and historical eras as well as social domain.

Future Directions

In this section, I discuss four possible future directions growing from this program of research. I focus on conceptual questions relevant across multiple chapters rather than issues of clarification such as were discussed in individual chapters (e.g., using a different measure of policy support to avoid ceiling effects). What is yet to be learned?

Group Membership and Group-Linked Attributes

The relationship between a social group label and that group's stereotyped attributes was central to the first two chapters. In Chapter I, I noted that the final list of adjectives in the modern checklist measure (see Table 8) contained situational attributes infrequently measured in stereotype content research (e.g., *homeless*) alongside more commonly measured dispositional attributes (e.g., *masculine*, *feminine*). In future work with this measure (or other measures that include both dispositional and situational attributes) it would be interesting to test whether participants are more likely to essentialize one or the other, and whether this differs by target group (similar to the patterns of biogenetic and sociocultural essentialism in Rangel & Keller, 2011). It would also be interesting to take a causal network approach, seeing whether (and when) participants think that dispositional traits cause situational ones and vice-versa (e.g., Ahn, 1999). In particular, it would be worth knowing if the two types of attributes are seen as mutually causal when assessing groups perceived to share social experience and biogenetic factors, but situational attributes have greater causal importance for groups perceived to share only social experiences. Such studies would offer insight into the relationship between group labels and

group attributes and dig deeper into how causal lay theories (biogenetic and otherwise) influence stereotypical trait attribution.

In Chapter II, attributes and membership judgments diverged in two interesting ways. First, for Gender, Race, and Religion, participants who judged the individual to still be an initial group member – an essentialist response – attributed the individual less initial-group-stereotypic traits – a non-essentialist response. Second, participants were making attributions based not only on the individual's initial and latter group memberships, but also on the perceived process of transiting between the groups. In both cases, the relatively open-ended nature of the adjective checklist allowed participants to express holistic stereotypes, beyond binary choices of adjectives that strongly distinguished the initial and latter groups. These patterns suggest that essentialist reasoning – whether biogenetic or sociocultural – does not operate in clean isolation, but rather as part of larger patterns of beliefs. Carrying this forward, it would be interesting to know when membership-based and attribute-based operationalizations of essentialism do and don't align, and when essentialist beliefs rise to the fore in social reasoning (e.g., the patterns of results for Religion, Race, Eating Habits, and Political Ideology) or are weaker than other components (e.g., the pattern of results for Wealth Class).

Normativity and Participant Recruitment

In this dissertation, I relied heavily on recruitment from Amazon Mechanical Turk. This method of recruitment is convenient, and common across a lot of different research within psychological science (Paolacci & Chandler, 2014). It is also particular – it means I recruited from a set of people who do tasks online for compensation. Though this sample is more diverse in most respects than undergraduate psychology students, a majority of my participants were White. Participants bring their lived experiences with them into studies, and generalizing beyond

the study population, when possible, is complex. How would data from mostly non-White participants affect the final adjective checklist? How would sociocultural essentialism in Studies 1 and 2 look among participants who live in more- and less-diverse parts of the country or world? How would Chapter III's manipulation look if the participants had been interacting directly with a researcher, instead of anonymously online?

As with recruitment on Amazon Mechanical Turk, the problem of generalization from relatively small or particular samples is widespread across the social sciences (e.g., Henrich, 2010). Within this dissertation, I have attempted to be cautious in my language, stating that a pattern *can* occur or *did* occur – acknowledging potential conditionality and being specific to my sample and their conditions of participation – rather than that it *does* occur (which implies broad generalization). This research resides within a broader body of research suggesting that stereotype content matters, that essentialist belief is widespread, and that discrimination experiences can activate ingroup-defensive mechanisms. More work, looking at these questions among different participants and with population-appropriate methods (see Medin et al., 2010), is critical to understanding how, when, and if the findings generalize beyond a default group in psychological research.

Sociocultural Essentialism and Power

The relationship between essentialism and maintenance of the status quo was a starting point for this dissertation work. For members of advantaged groups, biogenetic essentialism could reinforce the supposed naturalness of the status quo, offering reassurance about the stability of one's social position. For members of marginalized groups, de-essentializing one's ingroup could run counter to the motivation to be a member of a positively distinctive ingroup, but biogenetically essentializing it could justify the existing inequitable system. Sociocultural

essentialism, on the other hand, may allow us to recognize the lasting impacts our social groups leave on us without reinforcing existing hierarchy.

After establishing that sociocultural essentialism could be detected as a general pattern of thought in Chapter II, this question – of whether members of marginalized groups would strategically alter their endorsement of sociocultural essentialism – was the focus of Chapter III. I conducted an experiment with LGB+ people to see whether members of this group, which experiences both devaluation and denial in contemporary society, would engage in strategic essentialism in response to reminders of those experiences. As Study 3 failed to replicate the manipulation's effect on levels of biogenetic essentialism, its failure to change participants' level of sociocultural essentialism is hard to interpret. Perhaps LGB+ people never strategically activate essentialist reasoning. Perhaps LGB+ people in the US have more stable essentialist beliefs than LGB+ people in Europe. Perhaps, as the exploratory results indicated, discrimination experiences and sociocultural essentialism hold different meaning for older and younger LGB+ people. Among LGB+ people, and in other social domains, there is much yet to be learned about the relationship between sociocultural essentialism and maintenance of social hierarchy.

Though I was centrally interested in how sociocultural essentialism might disrupt hierarchy, it is also easy to bring to mind socioculturally essentialist arguments that justify current inequities. Consider arguments that trace all of modern society to "Western culture," or that attribute the perpetuation of wealth inequity to the "culture of poverty" (the 1965 Moynihan report is one notorious example of such an argument; for a thorough treatment of the topic, see Small et al., 2010). Consider also how trans-exclusionary radical feminists exclude transwomen from the superordinate social group *women* by arguing that transwomen lack the formative experiences (e.g., gendered socialization) that define womanhood, or how bisexual people in

opposite-gender relationships get told they're "not gay enough" because their relationships are straight-passing. These are only examples, but they suggest that sociocultural essentialism can be used to defend the status quo, as well as challenge it – depending on who is using it when and how. How sociocultural essentialism is *used*, and how it functions within the systems of power and hierarchy that characterize modern society, deserves further exploration.

Beyond the Individual

People's ideas about the world and social organization do not arise from the ether. Like us, they are products of the sociohistorical moment, emerging from everyday mixtures of social interaction, cultural learning, and thought. If sociocultural essentialism has a role to play – and I believe it does – it is worth tracking it down beyond the individual mind. This could involve observing the interactive process by which such ideas spread, in text analysis of online communication, with controlled experiments, or via interviews with participants reflecting on the source of their ontological convictions. This could also take an institutional lens, assessing when diversity statements or diversity policies (e.g., at academic institutions, at companies, in federal legislation) implicitly endorse biogenetically essentialist, socioculturally essentialist, environmental, or individually-based explanations (see Figure 13) – and how statements with these contents impact members of different social groups. Given conceptual shared ground between sociocultural essentialism and multiculturalism, this could offer a new perspective in ongoing conversations about diversity ideologies (e.g., Plaut et al., 2018). Finally, this could take a historical lens – as with stereotyping research that has used historical text corpora (e.g., Bhatia & Bhatia, 2020), it would be informative to see how cultural narratives featuring sociocultural essentialism rise and fall over historical time, with respect to particular events or movements (e.g., the Black Power movement, the removal of homosexuality from the DSM, changes in

immigration laws) and public opinion (e.g., as indexed by the General Social Survey or other national polling). These studies would go beyond whether sociocultural essentialism exists as a pattern of thought within an individual mind, and beyond whether individual people can use it strategically, to assess whether and how it travels across people, institutions, and time.

References

Abele, A. E., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2021). Navigating the social world: Toward an integrated framework for evaluating self, individuals, and groups. *Psychological Review*, *128*(2), 290–314.

Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y. (2016). Facets of the fundamental content dimensions: Agency with competence and assertiveness - Communion with warmth and morality. *Frontiers in Psychology*, *7*, 1810.

Advancing Academic Research with Twitter Data. (2021). Twitter.Com. https://developer.twitter.com/en/solutions/academic-research

Ahler, D. J., Roush, C. E., & Sood, G. (2019). *Micro-task market for lemons*. Meeting of the Midwest Political Science Association, Chicago, IL.

Ahn, W.-K. (1999). Effect of causal structure on category construction. *Memory & Cognition*, *27*(6), 1008–1023.

Alghamdi, R., & Alfalqi, K. (2015). A survey of topic modeling in text mining. *International Journal of Advanced Computer Science and Applications*, *6*(1).

Andreychik, M. R., & Gill, M. J. (2015). Do natural kind beliefs about social groups contribute to prejudice? Distinguishing bio-somatic essentialism from bio-behavioral essentialism, and both of these from entitativity. *Group Processes & Intergroup Relations*, *18*(4), 454–474.

Anyon, J. (1984). Intersections of gender and class: Accommodation and resistance by working-class and affluent females to contradictory sex role ideologies. *Journal of Education*, *166*(1), 25–48.

Apostolou, M. (2019). Why men stay single? Evidence from Reddit. *Evolutionary Psychological Science*, *5*(1), 87–97.

Arseneau, J. R., Grzanka, P. R., Miles, J. R., & Fassinger, R. E. (2013). Development and initial validation of the Sexual Orientation Beliefs Scale (SOBS). *Journal of Counseling Psychology*, *60*(3), 407–420.

Asch, S. E., & Zukier, H. (1984). Thinking about persons. *Journal of Personality and Social Psychology*, *46*(6), 1230–1240.

Atran, S., Medin, D., Lynch, E., Vapnarsky, V., Ucan Ek', E., & Sousa, P. (2001). Folkbiology doesn't come from folkpsychology: Evidence from Yukatek Maya in cross-cultural perspective. *Journal of Cognition and Culture*, *1*(1), 3–42.

Auguostinos, M., & Walker, I. (1998). The construction of stereotypes within social psychology: From social cognition to ideology. *Theory & Psychology*, *8*(5), 629–652.

Baars, B. J. (1986). *The Cognitive Revolution in Psychology*. Guilford Press.

Bailey, A. H., Knobe, J., & Newman, G. E. (In Press). Value-based essentialism: Essentialist beliefs about social groups with shared values. *Journal of Experimental Psychology: General*.

Bailey, A. H., LaFrance, M., & Dovidio, J. F. (2019). Is man the measure of all things? A social cognitive account of androcentrism. *Personality and Social Psychology Review*, *23*(4), 307–331.

Bailey, A. H., LaFrance, M., & Dovidio, J. F. (2020). Implicit androcentrism: Men are human, women are gendered. *Journal of Experimental Social Psychology*, *89*.

Bastian, B., & Haslam, N. (2006). Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology*, *42*(2), 228–235.
Bastide, R., & Van Den Berghe, P. (1957). Stereotypes, norms and interracial behavior in Sao Paulo, Brazil. *American Sociological Review*, *22*(6), 689.
Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using **lme4**. *Journal of Statistical Software*, *67*(1).
Baumgartner, J., Zannettou, S., Keegan, B., Squire, M., & Blackburn, J. (2020). The Pushshift Reddit dataset. *ArXiv:2001.08435 [Cs]*. http://arxiv.org/abs/2001.08435
Bayton, J. A., & Byoune, E. F. (1947). Racio-national stereotypes held by Negroes. *The Journal of Negro Education*, *16*(1), 49.
Beaty, R. E., & Johnson, D. R. (2021). Automating creativity assessment with SemDis: An open platform for computing semantic distance. *Behavior Research Methods*, *53*(2), 757–780.
Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). *On the dangers of stochastic parrots: Can language models be too big?* Association for Computing Machinery Conference on Fairness, Accountability, and Transparency, Online Conference. http://faculty.washington.edu/ebender/papers/Stochastic_Parrots.pdf
Bendig, A. W., & Hountras, P. T. (1958). College student stereotypes of the personality traits of research scientists. *Journal of Educational Psychology*, *49*(6), 309–314.
Bernstein, M. J., Sacco, D. F., Young, S. G., Hugenberg, K., & Cook, E. (2010). Being "in" with the in-crowd: The effects of social exclusion and inclusion are enhanced by the perceived essentialism of ingroups and outgroups. *Personality and Social Psychology Bulletin*, *36*(8), 999–1009.
Berreman, J. V. (1958). Filipino stereotypes of racial and national minorities. *The Pacific Sociological Review*, *1*(1), 7–12.
Bhatia, N., & Bhatia, S. (2020). Changes in gender stereotypes over time: A computational analysis. *Psychology of Women Quarterly*.
Bishop, C., Kiss, M., Morrison, T. G., Rushe, D. M., & Specht, J. (2014). The association between gay men's stereotypic beliefs about drag queens and their endorsement of hypermasculinity. *Journal of Homosexuality*, *61*(4), 554–567.
Bjerstedt, Å. (1960). "Ego-involved world-mindedness," nationality images, and methods of research: A methodological note. *Journal of Conflict Resolution*, *4*(2), 185–192.
Blake, R., & Dennis, W. (1943). The development of stereotypes concerning the Negro. *The Journal of Abnormal and Social Psychology*, *38*(4), 525–531.
Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition*, *60*, 1–29.
Bodenhausen, G. V., Macrae, C. N., Garst, J., Sedikides, C., Schopler, J., & Insko, C. A. (1998). Stereotypes in thought and deed: Social-cognitive origins of intergroup discrimination. In *Intergroup cognition and intergroup behavior* (pp. 311–335). Lawrence Erlbaum Associates, Inc.
Bodenhausen, G. V., & Richeson, J. A. (2010). Prejudice, stereotyping, and discrimination. In R. F. Baumeister & E. J. Finkel (Eds.), *Advanced Social Psychology: The State of the Science* (pp. 341–383). Oxford University Press.
Bohannon, J. (2016). Mechanical Turk upends social sciences. *Science*, *352*(6291), 1263–1264.
Bourne, E. (1977). Can we describe an individual's personality? Agreement on stereotype versus individual attributes. *Journal of Personality and Social Psychology*, *35*(12), 863–872.

Boyer, P. (1996). What makes anthropomorphism natural: Intuitive ontology and cultural representations. *The Journal of the Royal Anthropological Institute*, *2*(1), 83.
Brekhus, W. (1998). A sociology of the unmarked: Redirecting our focus. *Sociological Theory*, *16*(1), 34–51.
Brescoll, V. L., & LaFrance, M. (2004). The correlates and consequences of newspaper reports of research on sex differences. *Psychological Science*, *15*(8), 515–520.
Brescoll, V. L., Uhlmann, E. L., & Newman, G. E. (2013). The effects of system-justifying motives on endorsement of essentialist explanations for gender differences. *Journal of Personality and Social Psychology*, *105*(6), 891–908.
Brewster, K. L., & Padavic, I. (2000). Change in gender-ideology, 1977-1996: The contributions of intracohort change and population turnover. *Journal of Marriage and Family*, *62*(2), 477–487.
Brigham, J. C. (1969). *Ethnic stereotypes, attitudes, and treatment of ethnic group members* [Doctoral Dissertation]. University of Colorado.
Brigham, J. C. (1971). Ethnic stereotypes. *Psychological Bulletin*, *76*(1), 15–38.
Bronfenbrenner, U., & Ceci, S. J. (1994). Nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychological Review*, *101*(4), 568–586.
Brown, P. M., & Turner, J. C. (2002). The role of theories in the formation of stereotype content. In C. McGarty, V. Y. Yzerbyt, & R. Spears (Eds.), *Stereotypes as Explanations* (1st ed., pp. 67–89). Cambridge University Press.
Buchanan, W. (1951). Stereotypes and tensions as revealed by the UNESCO international poll. *International Social Science Bulletin*, *3*, 515–528.
Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science*, *3*(1), 14–25.
Cantril, H., & Strunk, M. (Eds.). (1951). National characteristics. In *Public Opinion 1935-1946* (pp. 496–502). Princeton University Press.
Carlson, R., & Levy, N. (1968). Brief method for assessing social-personal orientation. *Psychological Reports*, *23*(3), 911–914.
Carneggie, D. M. (1990). *The effect of AIDS education on students' knowledge about AIDS and attitudes towards persons with AIDS* [Master's Thesis]. University of Alaska - Anchorage.
Causadias, J. M., Vitriol, J. A., & Atkin, A. L. (2018). Do we overemphasize the role of culture in the behavior of racial/ethnic minorities? Evidence of a cultural (mis)attribution bias in American psychology. *American Psychologist*, *73*(3), 243–255.
Chalik, L., Leslie, S.-J., & Rhodes, M. (2017). Cultural context shapes essentialist beliefs about religion. *Developmental Psychology*, *53*(6), 1178–1187.
Chao, M. M., Hong, Y., & Chiu, C. (2013). Essentializing race: Its implications on racial categorization. *Journal of Personality and Social Psychology*, *104*(4), 619–634.
Cinelli, M., Quattrociocchi, W., Galeazzi, A., Valensise, C. M., Brugnoli, E., Schmidt, A. L., Zola, P., Zollo, F., & Scala, A. (2020). The COVID-19 social media infodemic. *Scientific Reports*, *10*(1), 16598.
Cobb, C. L., Meca, A., Branscombe, N. R., Schwartz, S. J., Xie, D., Zea, M. C., Fernandez, C. A., & Sanders, G. L. (2019). Perceived discrimination and well-being among unauthorized Hispanic immigrants: The moderating role of ethnic/racial group identity centrality. *Cultural Diversity and Ethnic Minority Psychology*, *25*(2), 280–287.

Cohler, B. J., & Hammack, P. L. (2006). Making a gay identity: Life story and the construction of a coherent self. In D. P. McAdams, R. Josselson, & A. Lieblich (Eds.), *Identity and story: Creating self in narrative.* (pp. 151–172). American Psychological Association.

Coleman, J. M., & Hong, Y.-Y. (2008). Beyond nature and nurture: The influence of lay gender theories on self-stereotyping. *Self and Identity*, *7*(1), 34–53.

Coley, J., Feeney, A., Xu, Y., Cohen-Pilat, M., Eidson, R. C., Smyth, K., Fangfang, W., & Bin, Z. (2019). *A two-component framework captures cross-cultural similarities and differences in essentialist thinking about social categories* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/jbg4r

Condit, C. M. (2019). Laypeople are strategic essentialists, Not genetic essentialists. *Hastings Center Report*, *49*, S27–S37.

Correll, J., Judd, C. M., Park, B., & Wittenbrink, B. (2010). Measuring prejudice, stereotypes, and discrimination. In J. F. Dovidio, M. Hewstone, P. Glick, & V. M. Esses (Eds.), *The SAGE handbook of prejudice, stereotyping, and discrimination* (pp. 45–62). SAGE Publications.

Craig, M. A., & Richeson, J. A. (2012). Coalition or derogation? How perceived discrimination influences intraminority intergroup relations. *Journal of Personality and Social Psychology*, *102*(4), 759–777.

Craig, M. A., Rucker, J. M., & Richeson, J. A. (2018). Racial and political dynamics of an approaching "majority-minority" United States. *The ANNALS of the American Academy of Political and Social Science*, *677*(1), 204–214.

Cross, W. E. (1995). The psychology of nigrescence: Revising the Cross model. In J. G. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander (Eds.), *Handbook of Multicultural Counseling*. Sage Publications Inc.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, 18.

Cuddy, A. J. C., Norton, M. I., & Fiske, S. T. (2005). This old stereotype: The pervasiveness and persistence of the elderly stereotype. *Journal of Social Issues*, *61*(2), 267–285.

Dar-Nimrod, I., & Heine, S. J. (2006). Exposure to scientific theories affects women's math performance. *Science*, *314*(5798), 435–435.

Dastin, J. (2018, October 10). Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

Datta, S., & Adar, E. (2019). *Extracting inter-community conflicts in reddit*. Thirteenth International AAAI Conference on Web and Social Media.

Davoodi, T. (2018). *Essentialization of social categories and links to moral development*. 144.

Davoodi, T., Soley, G., Harris, P. L., & Blake, P. R. (2019). Essentialization of social categories across development in two cultures. *Child Development*, *91*(1), 289–306.

Dawson, M. (1994). *Behind the mule: Race and class in African-American politics*. Princeton University Press.

del Río, M. F., & Strasser, K. (2011). Chilean children's essentialist reasoning about poverty: Children's reasoning about poverty. *British Journal of Developmental Psychology*, *29*(4), 722–743.

Demoulin, S., Leyens, J.-P., & Yzerbyt, V. (2006). Lay theories of essentialism. *Group Processes & Intergroup Relations*, *9*(1), 25–42.

Dempsey, J. M. (1992). *Stereotypes of six clinical nursing specialties* [Doctoral Dissertation]. Teachers College, Columbia University.
Der-Karabetian, A., & Smith, A. J. (1977). Sex-role stereotyping in the United States: Is it changing? *Sex Roles*, *3*(2), 193–198.
Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton Trilogy revisited. *Personality and Social Psychology Bulletin*, *21*, 1139–1150.
Dolan, S. (2020). *Jq* [Linux]. https://stedolan.github.io/jq/
Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological and Personality Science*, *3*(5), 562–571.
Duggan, M., & Smith, A. (2013). *6% of Online Adults are reddit Users*. Pew Research Center.
Duijker, H. C. J., & Frijda, N. H. (1960). *National character and national stereotypes: A trend report*. International Union of Scientific Psychology.
Eagly, A. H., & Kite, M. E. (1987). Are stereotypes of nationalities applied to both women and men? *Journal of Personality and Social Psychology*, *53*(3), 451–462.
Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2020). Gender stereotypes have changed: A cross-temporal meta-analysis of U.S. public opinion polls from 1946 to 2018. *American Psychologist*, *75*(3), 301–315.
Edelstein, M. L. (1972). *What do young Africans think? An attitude survey of urban African matric pupils in Soweto with special reference to stereotyping and social distance: A sociological study*. South African Institute of Race Relations.
Ehrlich, H. J., & Rinehart, J. W. (1965a). On checklists and open-ended questions: A reply. *Social Forces*, *44*(3), 420–421.
Ehrlich, H. J., & Rinehart, J. W. (1965b). A brief report on the methodology of stereotype research. *Social Forces*, *43*(4), 564–575.
Eidson, R. C., & Coley, J. D. (2014). Not so fast: Reassessing gender essentialism in young adults. *Journal of Cognition and Development*, *15*(2), 382–392.
Eisele, G., Vachon, H., Lafit, G., Kuppens, P., Houben, M., Myin-Germeys, I., & Viechtbauer, W. (2020). The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population. *Assessment*, 1016.
Embrick, D. G., & Henricks, K. (2013). Discursive colorlines at work: How epithets and stereotypes are racially unequal. *Symbolic Interaction*, *36*(2), 197–215.
Erut, A. S. (2017). *Ethnicity, essentialism, and folk sociology among the Wichï of Northern Argentina* [Master of Arts]. University of California, Los Angeles.
Ewens, W. L. (1969). *Reference, other support, ethnic attitudes, and perceived influence of others in the performance of overt acts* [Doctoral Dissertation]. The University of Iowa.
Fanon, F. (1963). *The Wretched of the Earth*. Grove Press.
Fetner, T. (2016). U.S. attitudes toward lesbian and gay people are better than ever. *Contexts*, *15*(2), 20–27.
Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83.
Fowlie, G. M., & Cooper, E. B. (1978). Traits attributed to stuttering and nonstuttering children by their mothers. *Journal of Fluency Disorders*, *3*(4), 233–246.
Freeman, D. (1969). The survival value of the beard. *Psychology Today*, *3*, 36–39.

Frost, D. M., & Meyer, I. H. (2012). Measuring community connectedness among diverse sexual minority populations. *Journal of Sex Research*, *49*(1), 36–49.
Fryer, J. H., & Cohen, L. (1988). Effects of labeling patients "Psychiatric" or "Medical": Favorability of traits ascribed by hospital staff. *Psychological Reports*, *62*(3), 779–793.
Gaither, S. E., Schultz, J. R., Pauker, K., Sommers, S. R., Maddox, K. B., & Ambady, N. (2014). Essentialist thinking predicts decrements in children's memory for racially ambiguous faces. *Developmental Psychology*, *50*(2), 482–488.
Galinsky, A. D., Hall, E. V., & Cuddy, A. J. C. (2013). Gendered races: Implications for interracial marriage, leadership selection, and athletic participation. *Psychological Science*, *24*(4), 498–506.
Gallagher, N. M., & Bodenhausen, G. V. (Under Review). *Gender essentialism and the mental representation of transgender women and men: A multimethod investigation of stereotype content*.
Ganong, L. H., & Coleman, M. (1995). The content of mother stereotypes. *Sex Roles*, *32*(7–8), 495–512.
Garcia-Marques, L., Santos, A. S. C., & Mackie, D. M. (2006). Stereotypes: Static abstractions or dynamic knowledge structures? *Journal of Personality and Social Psychology*, *91*(5), 814–831.
Gardner, R. C., Kirby, D. M., Gorospe, F. H., & Villamin, A. C. (1972). Ethnic stereotypes: An alternative assessment technique, the stereotype differential. *The Journal of Social Psychology*, *87*(2), 259–267.
Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences*, *115*(16), E3635–E3644.
Garr-Schultz, A., & Gardner, W. (2021). "It's just a phase": Identity denial experiences, self-concept clarity, and emotional well-being in bisexual individuals. *Self and Identity*, *20*(4), 528–544.
Gaunt, R. (2006). Biological essentialism, gender ideologies, and role attitudes: What determines parents' involvement in child care. *Sex Roles*, *55*(7–8), 523–533.
Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford University Press.
Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences*, *8*(9), 404–409.
Gelman, S. A. (2013). Artifacts and essentialism. *Review of Philosophy and Psychology*, *4*(3), 449–463.
Gelman, S. A., & Heyman, G. D. (1999). Carrot-eaters and creature-believers: The effects of lexicalization on children's inferences about social categories. *Psychological Science*, *10*(6), 489–493.
Gelman, S. A., & Rhodes, M. (2012). "Two-thousand years of stasis": How psychological essentialism impedes evolutionary understanding. In K. S. Rosengren, S. K. Brem, E. M. Evans, & G. M. Sinatra (Eds.), *Evolution Challenges: Integrating Research and Practice in Teaching and Learning about Evolution* (pp. 3–21). Oxford University Press.
Gilbert, G. M. (1951). Stereotype persistence and change among college students. *The Journal of Abnormal and Social Psychology*, *46*(2), 245–254.

Gil-White, F. J. (2001). Are ethnic groups biological "species" to the human brain?: Essentialism in our cognition of some social categories. *Current Anthropology*, *42*(4).
Gimenes, M., & New, B. (2016). Worldlex: Twitter and blog word frequencies for 66 languages. *Behavior Research Methods*, *48*(3), 963–972.
Goodman, L. A. (1954). Kolmogorov-Smirnov tests for psychological research. *Psychological Bulletin*, *51*(2), 160–168.
Gordon, R. A., Rozelle, R. M., & Baxter, J. C. (1989). The effect of applicant age, job level, and accountability on perceptions of female job applicants. *The Journal of Psychology*, *123*(1), 59–68.
Gough, H. G. (1960). The Adjective Check List as a personality assessment research technique. *Psychological Reports*, *6*, 107–122.
Gough, H. G., & Heilbrun, A. B. (1983). *The Adjective Check List Manual*. Mind Garden, Inc.
Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216.
Gruberg, S., Mahowald, L., & Halpin, J. (2020). *The state of the LGBTQ community in 2020: A national public opinion study*. Center for American Progress. https://www.americanprogress.org/issues/lgbtq-rights/reports/2020/10/06/491052/state-lgbtq-community-2020/
Hammack, P. L. (2005). The life course development of human sexual orientation: An integrative paradigm. *Human Development*, *48*(5), 267–290.
Hammack, P. L., Hughes, S. D., Atwood, J. M., Cohen, E. M., & Clark, R. C. (2021). Gender and sexual identity in adolescence: A mixed-methods study of labeling in diverse community settings. *Journal of Adolescent Research*.
Hammack, P. L., & Toolis, E. E. (2015). Putting the social into personal identity: The master narrative as root metaphor for psychological and developmental science. *Human Development*, *58*(6), 350–364.
Haratani, T., Matsuyama, Y., & Minami, Y. (1960). Study on stereotypes and preferences among Japanese students toward themselves and other national and ethnic groups. *Japanese Journal of Education and Psychology*, *8*(1), 1–7.
Harper, M. (2007). The stereotyping of nonreligious people by religious students: Contents and subtypes. *Journal for the Scientific Study of Religion*, *46*(4), 539–552.
Haslam, N., & Levy, S. R. (2006). Essentialist beliefs about homosexuality: Structure and implications for prejudice. *Personality and Social Psychology Bulletin*, *32*(4), 471–485.
Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology*, *39*(1), 113–127.
Haslam, N., Rothschild, L., & Ernst, D. (2002). Are essentialist beliefs associated with prejudice? *British Journal of Social Psychology*, *41*(1), 87–100.
Haslam, S. A., & Turner, J. C. (1995). Context-dependent variation in social stereotyping 3: Extremism as a self-categorical basis for polarized judgement. *European Journal of Social Psychology*, *25*(3), 341–371.
Hassell, J., & Smith, E. W. L. (1975). Female homosexuals' concepts of self, men, and women. *Journal of Personality Assessment*, *39*(2), 154–159.

Hegarty, P. (2002). “It’s not a choice, it’s the way we’re built”: Symbolic beliefs about sexual orientation in the US and Britain. *Journal of Community & Applied Social Psychology*, *12*(3), 153–166.
Hegarty, P., & Pratto, F. (2001). Sexual orientation beliefs: Their relationship to anti-gay attitudes and biological determinist arguments. *Journal of Homosexuality*, *41*(1), 121–135.
Heine, S. J., Dar-Nimrod, I., Cheung, B. Y., & Proulx, T. (2017). Essentially biased. In *Advances in Experimental Social Psychology* (Vol. 55, pp. 137–192). Elsevier.
Henrich, J. (2010). The weirdest people in the world. *Behavioral and Brain Sciences*, *33*(2–3), 61–83.
Henry, P. J., Reyna, C., & Weiner, B. (2004). Hate welfare but help the poor: How the attributional content of stereotypes explains the paradox of reactions to the destitute in America. *Journal of Applied Social Psychology*, *34*(1), 34–58.
Hirschfeld, L. A. (1995). Do children have a theory of race? *Cognition*, *54*(2), 209–252.
Hirschfeld, L. A. (1996). *Race in the Making: Cognition, Culture, and the Child’s Construction of Human Kinds*. MIT Press.
Ho, A. K., Roberts, S. O., & Gelman, S. A. (2015). Essentialism and racial bias jointly contribute to the categorization of multiracial individuals. *Psychological Science*, *26*(10), 1639–1645.
Hocking, R. R., & Leslie, R. N. (1967). Selection of the best subset in regression analysis. *Technometrics*, *9*(4), 531–540.
Holman, E. G., Ogolsky, B. G., & Oswald, R. F. (2021). Concealment of a sexual minority identity in the workplace: The role of workplace climate and identity centrality. *Journal of Homosexuality*, 1–19.
Hopkins, N., & Murdoch, N. (1999). The role of the ‘other’ in national identity: Exploring the context-dependence of the national ingroup stereotype. *Journal of Community & Applied Social Psychology*, 321–338.
Hopkins, N., Regan, M., & Abell, J. (1997). On the context dependence of national stereotypes: Some Scottish data. *British Journal of Social Psychology*, *36*(4), 553–563.
Horch, J. D. (2011). *Problem gambling stigma: Stereotypes, labels, self-stigma, and treatment-seeking.* [Doctoral Dissertation]. University of Calgary.
Howarth, C. (2002). Identity in whose eyes? The role of representations in identity construction. *Journal for the Theory of Social Behaviour*, *32*(2), 145–162.
Howell, A. J., Weikum, B. A., & Dyck, H. L. (2011). Psychological essentialism and its association with stigmatization. *Personality and Individual Differences*, *50*(1), 95–100.
HRC. (2021). *Federal legislation and policies HRC supports*. HRC. https://www.hrc.org/resources/federal-legislation
Hruschka, D. J., Medin, D. L., Rogoff, B., & Henrich, J. (2018). Pressing questions in the study of psychological and behavioral diversity. *Proceedings of the National Academy of Sciences*, *115*(45), 11366–11368.
Hswen, Y., Xu, X., Hing, A., Hawkins, J. B., Brownstein, J. S., & Gee, G. C. (2021). Association of “#covid19” versus “#chinesevirus” with anti-Asian sentiments on Twitter: March 9–23, 2020. *American Journal of Public Health*, *111*(5), 956–964.

Huic, A., Jelic, M., & Kamenov, Z. (2018). Essentialist beliefs about homosexuality predict positive and negative behavioral intentions toward lesbian women and gay men. *Journal of Homosexuality*, *65*(12), 1631–1655.

Iliev, R., Dehghani, M., & Sagi, E. (2015). Automated text analysis in psychology: Methods, applications, and future developments. *Language and Cognition*, *7*(2), 265–290.

Iqbal, M. (2018, October 5). Twitter Revenue and Usage Statistics (2021). *Business of Apps*. https://www.businessofapps.com/data/twitter-statistics/

Jaidka, K., Giorgi, S., Schwartz, H. A., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2020). Estimating geographic subjective well-being from Twitter: A comparison of dictionary and data-driven language methods. *Proceedings of the National Academy of Sciences*.

Jayaratne, T. E., Gelman, S. A., Feldbaum, M., Sheldon, J. P., Petty, E. M., & Kardia, S. L. R. (2009). The perennial debate: Nature, nurture, or choice? Black and White Americans' explanations for individual differences. *Review of General Psychology*, *13*(1), 24–33.

Jayaratne, T. E., Ybarra, O., Sheldon, J. P., Brown, T. N., Feldbaum, M., Pfeffer, C. A., & Petty, E. M. (2006). White Americans' genetic lay theories of race differences and sexual orientation: Their relationship with prejudice toward Blacks, and gay men and lesbians. *Group Processes & Intergroup Relations*, *9*(1), 77–94.

Joel, D., Tarrasch, R., Berman, Z., Mukamel, M., & Ziv, E. (2014). Queering gender: Studying gender identity in 'normative' individuals. *Psychology & Sexuality*, *5*(4), 291–321.

Johnson, K. L., Freeman, J. B., & Pauker, K. (2012). Race is gendered: How covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology*, *102*(1), 116–131.

Jones, J., Amin, M., Kim, J., & Skiena, S. (2020). Stereotypical gender associations in language have decreased over time. *Sociological Science*, *7*, 1–35.

Joseph, K., Wei, W., & Carley, K. M. (2017). Girls rule, boys drool: Extracting semantic and affective stereotypes from Twitter. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 1362–1374.

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology*, *33*(1), 1–27.

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, *25*(6), 40.

Kalish, C. W. (2002). Essentialist to some degree: Beliefs about the structure of natural kind categories. *Memory & Cognition*, *30*(3), 340–352.

Kang, S. K., Plaks, J. E., & Remedios, J. D. (2015). Folk beliefs about genetic variation predict avoidance of biracial individuals. *Frontiers in Psychology*, *6*.

Kanovsky, M. (2007). Essentialism and folksociology: Ethnicity again. *Journal of Cognition and Culture*, *7*(3–4), 241–281.

Karlins, M., Coffman, T. L., & Walters, A. (1969). On the fading of social stereotypes: Studies in three generations of college students. *Journal of Personality and Social Psychology*, *13*(1), 1–16.

Katz, D., & Braly, K. W. (1933). Racial stereotypes of one hundred college students. *The Journal of Abnormal and Social Psychology*, *28*, 280–290.

Katz, D., & Braly, K. W. (1935). Racial prejudice and racial stereotypes. *The Journal of Abnormal and Social Psychology*, *30*(2), 175–193.

Kaufman, G., & Compton, D. (2021). Attitudes toward LGBT marriage and legal protections post-Obergefell. *Sexuality Research and Social Policy*, *18*(2), 321–330.
Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology*, *88*(4), 686–702.
Kelly, C., & Breinlinger, S. (1995). Identity and injustice: Exploring women's participation in collective action. *Journal of Community & Applied Social Psychology*, *5*, 41–57.
Kite, M. E., & Bryant-Lees, K. B. (2016). Historical and contemporary attitudes toward homosexuality. *Teaching of Psychology*, *43*(2), 164–170.
Koch, A., Imhoff, R., Dotsch, R., Unkelbach, C., & Alves, H. (2016). The ABC of stereotypes about groups: Agency/socioeconomic success, conservative–progressive beliefs, and communion. *Journal of Personality and Social Psychology*, *110*(5), 675–709.
Koeske, G. (1971). *Unpublished manuscript*.
Koestner, R., Franz, C., & Weinberger, J. (1990). The family origins of empathic concern: A 26-year longitudinal study. *Journal of Personality and Social Psychology*, *58*(4), 709–717.
Kotzur, P. F., Veit, S., Namyslo, A., Holthausen, M.-A., Wagner, U., & Yemane, R. (2020). 'Society thinks they are cold and/or incompetent, but I do not': Stereotype content ratings depend on instructions and the social group's location in the stereotype content space: Individual versus societal perspective in SCM research. *British Journal of Social Psychology*.
Kraus, M. W., & Keltner, D. (2013). Social class rank, essentialism, and punitive judgment. *Journal of Personality and Social Psychology*, *105*(2), 247–261.
Kraus, M. W., Rucker, J. M., & Richeson, J. A. (2017). Americans misperceive racial economic equality. *Proceedings of the National Academy of Sciences*, *114*(39), 10324–10331.
Kray, L. J., Howland, L., Russell, A. G., & Jackman, L. M. (2017). The effects of implicit gender role theories on gender system justification: Fixed beliefs strengthen masculinity to preserve the status quo. *Journal of Personality and Social Psychology*, *112*(1), 98–115.
Kunda, Z., Sinclair, L., & Griffin, D. (1997). Equal ratings but separate meanings: Stereotypes and the construal of traits. *Journal of Personality and Social Psychology*, *72*(4), 720–734.
Kung, F. Y. H., Chao, M. M., Yao, D. J., Adair, W. L., Fu, J. H., & Tasa, K. (2018). Bridging racial divides: Social constructionist (vs. essentialist) beliefs facilitate trust in intergroup contexts. *Journal of Experimental Social Psychology*, *74*, 121–134.
Kusunoti, K. (1936). Mental characteristics of the Japanese as seen by Japanese and American students. *Japanese Journal of Applied Psychology*, *4*, 232–237.
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13).
Leach, C. W., van Zomeren, M., Zebel, S., Vliek, M. L. W., Pennekamp, S. F., Doosje, B., Ouwerkerk, J. W., & Spears, R. (2008). Group-level self-definition and self-investment: A hierarchical (multicomponent) model of in-group identification. *Journal of Personality and Social Psychology*, *95*(1), 144–165.
Lei, R. (2017). *Ascribed Versus Acquired Representations of Social Class and Their Relation to Anti-Poor Prejudice* [Doctoral Dissertation]. Northwestern University.
Lei, R., & Bodenhausen, G. V. (2017). Racial assumptions color the mental representation of social class. *Frontiers in Psychology*, *8*.

Leslie, S.-J. (2013). Essence and natural kinds: When science meets preschooler intuition. *Oxford Studies in Epistemology*, *4*, 108–165.
Lewis, L. D., Darley, J. M., & Glucksberg, S. (1972). *Stereotype persistence and change among college students* [Unpublished Manuscript].
Leyens, J.-P., Cortes, B., Demoulin, S., Dovidio, J. F., Fiske, S. T., Gaunt, R., Paladino, M.-P., Rodriguez-Perez, A., Rodriguez-Torres, R., & Vaes, J. (2003). Emotional prejudice, essentialism, and nationalism: The 2002 Tajfel lecture. *European Journal of Social Psychology*, *33*(6), 703–717.
Leyens, J.-P., Rodriguez-Perez, A., Rodriguez-Torres, R., Gaunt, R., Paladino, M.-P., Vaes, J., & Demoulin, S. (2001). Psychological essentialism and the differential attribution of uniquely human emotions to ingroups and outgroups. *European Journal of Social Psychology*, *31*(4), 395–411.
Lickel, B., Hamilton, D. L., Wieczorkowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology*, *78*(2), 223–246.
Liu, M., & Wang, Y. (2015). Data collection mode effect on feeling thermometer questions: A comparison of face-to-face and Web surveys. *Computers in Human Behavior*, *48*, 212–218.
Lowery, B. S., Knowles, E. D., & Unzueta, M. M. (2007). Framing inequity safely: Whites' motivated perceptions of racial privilege. *Personality and Social Psychology Bulletin*, *33*(9), 1237–1250.
Maass, A., Salvi, D., Luciano, A., & Semin, G. (1989). Language use in intergroup contexts: The linguistic intergroup bias. *Journal of Personality and Social Psychology*, *57*(6), 981–993.
MacIntyre, R. D. (1981). *Modifying children's attitudes toward the handicapped through personal interaction* [Doctoral Dissertation]. The University of San Francisco.
Madon, S. (1997). What do people believe about gay males? A study of stereotype content and strength. *Sex Roles*, *37*(9–10), 663–685.
Madon, S., Guyll, M., Aboufadel, K., Montiel, E., Smith, A., Palumbo, P., & Jussim, L. (2001). Ethnic and national stereotypes: The Princeton Trilogy revisited and revised. *Personality and Social Psychology Bulletin*, *27*(8), 996–1010.
Mahalingam, R. (1998). *Essentialism, power and representation of caste: A developmental study* [Doctor of Philosophy]. University of Pittsburgh.
Mahalingam, R. (2003). Essentialism, culture, and power: Representations of social class. *Journal of Social Issues*, *59*(4), 733–749.
Mahalingam, R., Haritatos, J., & Jackson, B. (2007). Essentialism and the cultural psychology of gender in extreme son preference communities in India. *American Journal of Orthopsychiatry*, *77*(4), 598–609.
Mandalaywala, T. M., Amodio, D. M., & Rhodes, M. (2018). Essentialism promotes racial prejudice by increasing endorsement of social hierarchies. *Social Psychological and Personality Science*, *9*(4), 461–469.
Martin, C. L., & Parker, S. (1995). Folk theories about sex and race differences. *Personality and Social Psychology Bulletin*, *21*(1), 45–57.
Matan, A., & Carey, S. (2001). Developmental changes within the core of artifact concepts. *Cognition*, *78*(1), 1–26.

Matheson, K., & Strickland, L. (1986). The stereotype of the computer scientist. *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du Comportement*, *18*(1), 15–24.
Matthews, G., Jones, D. M., & Chamberlain, A. G. (1990). Refining the measurement of mood: The UWIST Mood Adjective Checklist. *British Journal of Psychology*, *81*(1), 17–42.
McCauley, C., & Stitt, C. L. (1978). An individual and quantitative measure of stereotypes. *Journal of Personality and Social Psychology*, *36*(9), 929–940.
McCauley, C., Stitt, C. L., & Segal, M. (1980). Stereotyping: From prejudice to prediction. *Psychological Bulletin*, *87*(1), 195–208.
Medin, D., Bennis, W., & Chandler, M. (2010). Culture and the home-field disadvantage. *Perspectives on Psychological Science*, *5*(6), 708–713.
Medin, D., Ojalehto, B., Marin, A., & Bang, M. (2017). Systems of (non-)diversity. *Nature Human Behaviour*, *1*(5), 0088.
Medin, D., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). Cambridge University Press.
Meenes, M. (1943). A comparison of racial stereotypes of 1935 and 1942. *The Journal of Social Psychology*, *17*(2), 327–336.
Menon, T., Morris, M. W., Chiu, C., & Hong, Y. (1999). Culture and the construal of agency: Attribution to individual versus group dispositions. *Journal of Personality and Social Psychology*, *76*(5), 707–717.
Miller, D. I., Nolla, K. M., Eagly, A. H., & Uttal, D. H. (2018). The development of children's gender-science stereotypes: A meta-analysis of 5 decades of U.S. draw-a-scientist studies. *Child Development*, *89*(6), 1943–1955.
Mitchell, S., Potash, E., Barocas, S., D'Amour, A., & Lum, K. (2021). Algorithmic fairness: Choices, assumptions, and definitions. *Annual Review of Statistics and Its Application*, *8*(1), 141–163.
Morton, T. A., Hornsey, M. J., & Postmes, T. (2009). Shifting ground: The variable use of essentialism in contexts of inclusion and exclusion. *British Journal of Social Psychology*, *48*(1), 35–59.
Morton, T. A., & Postmes, T. (2009). When differences become essential: Minority essentialism in response to majority treatment. *Personality and Social Psychology Bulletin*, *35*(5), 656–668.
Morton, T. A., Postmes, T., Haslam, S. A., & Hornsey, M. J. (2009). Theorizing gender in the face of social change: Is there anything essential about essentialism? *Journal of Personality and Social Psychology*, *96*(3), 653–664.
Moss, A. (2018). After the bot scare: Understanding what's been happening with data collection on MTurk and how to stop it. *Cloud Research*.
Moynihan, D. P. (1965). *The Negro family: The case for national action*. U.S. Department of Labor: Office of Policy Planning and Research.
Mukerji, M. A. (1951). *A study of stereotypes and attitude change* [Master's Thesis]. Lucknow University.
Murphy, N. (2019, December). Reddit's 2019 Year in Review. *Upvoted*. https://redditblog.com/2019/12/04/reddits-2019-year-in-review/
Neel, R., & Lassetter, B. (2015). Growing fixed with age: Lay theories of malleability are target age-specific. *Personality and Social Psychology Bulletin*, *41*(11), 1505–1522.

Nelkin, D., & Lindee, M. S. (1995). *The DNA Mystique*. The University of Michigan Press.
Newman, G. E., & Knobe, J. (2019). The essence of essentialism. *Mind & Language*, *34*(5), 585–605.
Nicolas, G., Bai, X., & Fiske, S. T. (2020). Comprehensive stereotype content dictionaries using a semi-automated method. *European Journal of Social Psychology*.
Niemann, Y. F., Jennings, L., Rozelle, R. M., Baxter, J. C., & Sullivan, E. (1994). Use of free responses and cluster analysis to determine stereotypes of eight groups. *Personality and Social Psychology Bulletin*, *20*(4), 379–390.
Nisbett, R. E., Zukier, H., & Lemley, R. E. (1981). The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology*, *13*(2), 248–277.
No, S., Hong, Y., Liao, H.-Y., Lee, K., Wood, D., & Chao, M. M. (2008). Lay theory of race affects and moderates Asian Americans' responses toward American culture. *Journal of Personality and Social Psychology*, *95*(4), 991–1004.
Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., Smith, C. T., Olson, K. R., Chugh, D., Greenwald, A. G., & Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, *18*(1), 36–88.
Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., Bar-Anan, Y., Bergh, R., Cai, H., Gonsalkorale, K., Kesebir, S., Maliszewski, N., Park, J., Schnabel, K., Shiomura, K., Tulbure, B. T., Wiers, R. W., Akrami, N., Ekehammar, B., … Greenwald, A. G. (2009). National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, *106*(26), 10593–10597.
Nzinga, K., Rapp, D. N., Leatherwood, C., Easterday, M., Rogers, L. O., Gallagher, N., & Medin, D. L. (2018). Should social scientists be distanced from or engaged with the people they study? *Proceedings of the National Academy of Sciences*, *115*(45), 11435–11441.
O'Connor, K. J., & D'Angelo, C. M. (2013). Pre-service teachers' perceptions of middle school students. *Educational Studies*, *39*(5), 548–551.
Ooms, J. (2014). The jsonlite package: A practical and consistent matching between JSON data and R objects. *ArXiv:1403.2805*. https://arxiv.org/abs/1403.2805
Oulmokhtar, D., Krauth-Gruber, S., Alexopoulos, T., & Drozda-Senkowska, E. (2011). Police officers' stereotype content and its evolution over two decades: From "neither nice nor able" to "still not nice but able." *Revue Internationale de Psychologie Sociale*, *24*(3), 87–100.
Oxford Languages. (2021). *Oxford Languages and Google—English | Oxford Languages*. https://languages.oup.com/google-dictionary-en/
Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, *23*(3), 184–188.
Parish, T. S., Bryant, W. T., & Shirazi, A. (1976). The Personal Attribute Inventory. *Perceptual and Motor Skills*, *42*, 715–720.
Parish, T. S., & Taylor, J. C. (1978). The Personal Attribute Inventory for children: A report on its validity and reliability as a self-concept scale. *Educational and Psychological Measurement*, *38*(2), 565–569.

Paxton, K. R. (1983). *A comparison of the effectiveness of three short-term instructional programs on the attitudes of fifth and sixth grade students toward the handicapped* [Doctoral Dissertation]. University of Alabama.

Payne, B. K., Vuletich, H. A., & Brown-Iannuzzi, J. L. (2019). Historical roots of implicit bias in slavery. *Proceedings of the National Academy of Sciences*, *116*, 11693–11698.

Pehrson, S., Brown, R., & Zagefka, H. (2009). When does national identification lead to the rejection of immigrants? Cross-sectional and longitudinal evidence for the role of essentialist in-group definitions. *British Journal of Social Psychology*, *48*(1), 61–76.

Perlmutter, H. V. (1954). Relations between the self-image, the image of the foreigner, and the desire to live abroad. *The Journal of Psychology*, *38*(1), 131–137.

Perry, S. P., Skinner, A. L., & Abaied, J. L. (2019). Bias awareness predicts color conscious racial socialization methods among white parents. *Journal of Social Issues*, *75*(4), 1035–1056.

Petsko, C., & Bodenhausen, G. V. (2019). Racial stereotyping of gay men: Can a minority sexual orientation erase race? *Journal of Experimental Social Psychology*, *83*, 37–54.

Pettigrew, T. F. (2018). The emergence of contextual social psychology. *Personality and Social Psychology Bulletin*, *44*(7), 963–971.

Philogène, G. (2001). Stereotype fissure: Katz and Braly revisited. *Social Science Information*, *40*(3), 411–432.

Plaut, V. C., Thomas, K. M., Hurd, K., & Romano, C. A. (2018). Do color blindness and multiculturalism remedy or foster discrimination and racism? *Current Directions in Psychological Science*, *27*(3), 200–206.

Prentice, D. A., & Miller, D. T. (2007). Psychological essentialism of human categories. *Current Directions in Psychological Science*, *16*(4), 202–206.

Prothro, E. T., & Melikian, L. H. (1954). Studies in stereotypes: III. Arab students in the Near East. *The Journal of Social Psychology*, *40*(2), 237–243.

Prothro, E. T., & Melikian, L. H. (1955). Studies in stereotypes: V. Familiarity and the kernel of truth hypothesis. *The Journal of Social Psychology*, *41*(1), 3–10.

Pruden, S. M., & Levine, S. C. (2017). Parents' spatial language mediates a sex difference in preschoolers' spatial-language use. *Psychological Science*, *28*(11), 1583–1596.

Purdie-Vaughns, V., & Eibach, R. P. (2008). Intersectional invisibility: The distinctive advantages and disadvantages of multiple subordinate-group identities. *Sex Roles*, *59*(5–6), 377–391.

Putnam, H. (1975). The meaning of "Meaning." *Philosophical Papers*, *2*, 131–193.

R Core Team. (2018). *R: A language and environment for statistical computing*.

Rad, M. S., & Ginges, J. (2018). Folk theories of nationality and anti-immigrant attitudes. *Nature Human Behaviour*, *2*(5), 343–347.

Rad, M. S., Martingano, A. J., & Ginges, J. (2018). Toward a psychology of *Homo sapiens*: Making psychological science more representative of the human population. *Proceedings of the National Academy of Sciences*, *115*(45), 11401–11405.

Rad, M. S., Shackleford, C., Lee, K. A., Jassin, K., & Ginges, J. (2019). Folk theories of gender and anti-transgender attitudes: Gender differences and policy preferences. *PLOS ONE*, *14*(12).

Rangel, U., & Keller, J. (2011). Essentialism goes social: Belief in social determinism as a component of psychological essentialism. *Journal of Personality and Social Psychology*, *100*(6), 1056–1078.
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). SAGE Publications.
Revelle, W. (2019). *psych: Procedures for psychological, psychometric, and personality research*. Northwestern University.
Rhodes, M., & Gelman, S. A. (2009). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology*, *59*(3), 244–274.
Rogers, L. O. (2020). "I'm kind of a feminist": Using master narratives to analyze gender identity in middle childhood. *Child Development*, *91*(1), 179–196.
Rogers, L. O., & Way, N. (2016). "I have goals to prove all those people wrong and not fit into any one of those boxes": Paths of resistance to stereotypes among Black adolescent males. *Journal of Adolescent Research*, *31*(3), 263–298.
Rogers, L. O., & Way, N. (2018). Reimagining social and emotional development: Accommodation and resistance to dominant ideologies in the identities and friendships of boys of color. *Human Development*, *61*(6), 311–331.
Rothbart, M., & Taylor, M. (1992). Category labels and social reality: Do we view social categories as natural kinds. In G. R. Semin & K. Fiedler (Eds.), *Language, interaction, and social cognition* (pp. 11–36). Sage.
Rudolf, J., Wettstein, G., & Pi, R. (1961). La formazion de estereotipos y su relaction con los medios de expresión. *Revista Mexicana de Sociologia*, *23*, 443–452.
Rüsch, N., Todd, A. R., Bodenhausen, G. V., & Corrigan, P. W. (2010). Biogenetic models of psychopathology, implicit guilt, and mental illness stigma. *Psychiatry Research*, *179*(3), 328–332.
Ryazanov, A. A., & Christenfeld, N. J. S. (2018). The strategic value of essentialism. *Social and Personality Psychology Compass*, *12*(1).
Saenger, G., & Flowerman, S. (1954). Stereotypes and prejudicial attitudes. *Human Relations*, *7*(2), 217–238.
Sattelberg, W. (2020). *The demographics of Reddit: Who uses the site?* Tech Junkie. https://social.techjunkie.com/demographics-reddit/
Schmitt, M. T., Spears, R., & Branscombe, N. R. (2003). Constructing a minority group identity out of shared rejection: The case of international students. *European Journal of Social Psychology*, *33*(1), 1–12.
Schneider, M. C., & Bos, A. L. (2011). An exploration of the content of stereotypes of Black politicians. *Political Psychology*, *32*(2), 205–233.
Schneider, M. C., & Bos, A. L. (2014). Measuring stereotypes of female politicians. *Political Psychology*, *35*(2), 245–266.
Schoenfeld, N. (1942). *An Experimental Study of Some Problems Relating to Stereotypes*. Archives of Psychology.
Schor, N., & Weed, E. (1994). *The essential difference*. Indiana University Press.
Seago, D. W. (1947). Stereotypes: Before Pearl Harbor and after. *The Journal of Psychology*, *23*(1), 55–63.

Segev, M., Bergman, Y. S., & Diesendruck, G. (2012). Is religion essential? Beliefs about religious categories. *Journal of Cognition and Culture*, *12*(3–4), 323–337.
Sellers, R. M., Rowley, S. A. J., Chavous, T. M., Shelton, J. N., & Smith, M. A. (1997). Multidimensional inventory of Black identity: A preliminary investigation of reliability and construct validity. *Journal of Personality and Social Psychology*, *73*(4), 805–815.
Shaban, H. (2019). Twitter reveals its daily active user numbers for the first time. *The Washington Post*. https://www.washingtonpost.com/technology/2019/02/07/twitter-reveals-its-daily-active-user-numbers-first-time/
Sharp, L. M., & Frankel, J. (1983). Respondent burden: A test of some common assumptions. *Public Opinion Quarterly*, *47*(1), 36.
Sheldon, J. P., Pfeffer, C. A., Jayaratne, T. E., Feldbaum, M., & Petty, E. M. (2007). Beliefs about the etiology of homosexuality and about the ramifications of discovering its possible genetic origin. *Journal of Homosexuality*, *52*(3–4), 111–150.
Shrout, P. E., & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology*, *69*(1), 487–510.
Sides, J., & Gross, K. (2013). Stereotypes of Muslims and support for the War on Terror. *The Journal of Politics*, *75*(3), 583–598.
Silge, J., & Robinson, D. (2016). tidytext: Text mining and analysis using tidy data principles in R. *The Open Journal*, *1*(3).
Simon, B. (1992). The perception of ingroup and outgroup homogeneity: Reintroducing the intergroup context. *European Review of Social Psychology*, *3*(1), 1–30.
Simon, B., & Klandermans, B. (2001). Politicized collective identity. *American Psychologist*, *56*(4), 319–331.
Sinha, A. K. P., & Upadhyaya, O. P. (1960). Stereotypes of male and female university students in India toward different ethnic groups. *The Journal of Social Psychology*, *51*(1), 93–102.
Siperstein, G. N. (1980). *Instruments for Measuring Children's Attitudes Toward the Handicapped*. Center for Human Services, University of Massachusetts/Boston.
Six, B., & Eckes, T. (1991). A closer look at the complex structure of gender stereotypes. *Sex Roles*, *24*(1–2), 57–71.
Skewes, L., Fine, C., & Haslam, N. (2018). Beyond Mars and Venus: The role of gender essentialism in support for gender inequality and backlash. *PLOS ONE*, *13*(7).
Small, M. L., Harding, D. J., & Lamont, M. (2010). Reconsidering culture and poverty. *The ANNALS of the American Academy of Political and Social Science*, *629*(1), 6–27.
Smith, J., & Russell, G. (1984). Why do males and females differ? Children's beliefs about sex differences. *Sex Roles*, *11*(11–12), 1111–1120.
Smyth, K., Feeney, A., Eidson, R. C., & Coley, J. D. (2017). Development of essentialist thinking about religion categories in Northern Ireland (and the United States). *Developmental Psychology*, *53*(3), 475–496.
Sousa, P., Atran, S., & Medin, D. (2002). Essentialism and folkbiology: Evidence from Brazil. *Journal of Cognition and Culture*, *2*(3), 195–223.
Spivak, G. C. (1988). Can the subaltern speak? In C. Nelson & L. Grossberg (Eds.), *Marxism and the Interpretation of Culture*. University of Illinois Press.

Stangor, C., & Lange, J. E. (1994). Mental representations of social groups: Advances in understanding stereotypes and stereotyping. In *Advances in Experimental Social Psychology* (Vol. 26, pp. 357–416). Elsevier.

Staton, L. J., Panda, M., Chen, I., Genao, I., Kurz, J., Pasanen, M., Mechaber, A. J., Menon, M., O'Rorke, J., Wood, J., Rosenberg, E., Faeslis, C., Carey, T., Calleson, D., & Cykert, S. (2007). When race matters: Disagreement in pain perception between patients and their physicians in primary care. *Journal of the National Medical Association*, *99*(5), 7.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*.

Steele, C. M. (2010). *Whistling Vivaldi*. W.W. Norton & Company.

Stopar, A. (2015). Encounters with national stereotypes in foreign language teaching: Adjectives describing Americans. *ELOPE: English Language Overseas Perspectives and Enquiries*, *12*(1), 105–118.

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, *33*(1), 1–39.

Tajfel, H., & Forgas, J. P. (2000). Social categorization: Cognitions, values, and groups. In C. Stagnor (Ed.), *Stereotypes and Prejudice: Essential Readings*.

Thorndike, R. L. (1977). Content and evaluation in ethnic stereotypes. *The Journal of Psychology*, *96*(1), 131–140.

Toosi, N. R., & Ambady, N. (2011). Ratings of essentialism for eight religious identities. *International Journal for the Psychology of Religion*, *21*(1), 17–29.

Tsatsoulis-Bonnekessen, B. (1993). Exploring ethnic contact with stories and stereotypes: The Thematic Apperception Technique and the Adjective Checklist as supporting interview tools. *CAM*, *5*(2), 1–3.

Udolf, R. (1973). Liberal and conservative political stereotypes in terms of belief response hierarchies. *Psychological Reports*, *32*(1), 275–284.

Unkel, J., & Kümpel, A. S. (2020). (A)synchronous communication about TV series on social media: A multi-method investigation of Reddit discussions. *Media and Communication*, *8*(3), 180–190.

Vaes, J., Latrofa, M., Suitner, C., & Arcuri, L. (2019). They are all armed and dangerous! Biased language use in crime news with ingroup and outgroup perpetrators. *Journal of Media Psychology*, *31*(1), 12–23.

Vale, M. T., Bisconti, T. L., & Sublett, J. F. (2020). Benevolent ageism: Attitudes of overaccommodative behavior toward older women. *The Journal of Social Psychology*, *160*(5), 548–558.

van Ryn, M., Burgess, D. J., Dovidio, J. F., Phelan, S. M., Saha, S., Malat, J., Griffin, J. M., Fu, S. S., & Perry, S. (2011). The impact of racism on clinician cognition, behavior, and clinical decision making. *Du Bois Review: Social Science Research on Race*, *8*(1), 199–218.

Verkuyten, M. (2003). Discourses about ethnic group (de-)essentialism: Oppressive and progressive aspects. *British Journal of Social Psychology*, *42*(3), 371–391.

Verkuyten, M., & Brug, P. (2004). Multiculturalism and group status: The role of ethnic identification, group essentialism and protestant ethic. *European Journal of Social Psychology*, *34*(6), 647–661.

Wages, J., Perry, S., Skinner-Dorkenoo, A., & Bodenhausen, G. V. (2020). *Reckless gambles and responsible ventures: Racialized prototypes of risk-taking* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/2nphk

Walker, I., & Read, J. (2002). The differential effectiveness of psychosocial and biogenetic causal explanations in reducing negative attitudes toward "mental illness." *Psychiatry: Interpersonal and Biological Processes*, *65*(4), 313–325.

Waxman, S., Medin, D., & Ross, N. (2007). Folkbiological reasoning from a cross-cultural developmental perspective: Early essentialist notions are shaped by cultural beliefs. *Developmental Psychology*, *43*(2), 294–308.

Way, N., Cressen, J., Bodian, S., Preston, J., Nelson, J., & Hughes, D. (2014). "It might be nice to be a girl... Then you wouldn't have to be emotionless": Boys' resistance to norms of masculinity during adolescence. *Psychology of Men & Masculinity*, *15*(3), 241–252.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, *92*(4), 548–573.

Weitz, R., & Gordon, L. (1993). Images of Black women among Anglo college students. *Sex Roles*, *28*(1–2), 19–34.

Who uses YouTube, WhatsApp and Reddit. (2019). Pew Research Center. https://www.pewresearch.org/internet/chart/who-uses-youtube-whatsapp-and-reddit/

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V. S., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43).

Wigginton, M. L., Wiersma, L. D., Sherman, C. P., & Rubin, D. (2009). Weight stigmatization in children. *Californian Journal of Health Promotion*, *7*(1), 9.

Willemsen, T. M. (2002). Gender typing of the successful manager—A stereotype reconsidered. *Sex Roles*.

Williams, J. E., & Bennett, S. M. (1975). The definition of sex stereotypes via the adjective check list. *Sex Roles*, *1*(4).

Williams, J. E., & Best, D. L. (1982). *Measuring sex stereotypes: A thirty-nation study*. Sage Publications.

Williams, M. J., & Eberhardt, J. L. (2008). Biological conceptions of race and the motivation to cross racial boundaries. *Journal of Personality and Social Psychology*, *94*(6), 1033–1047.

Wilton, L. S., Apfelbaum, E. P., & Good, J. J. (2019). Valuing differences and reinforcing them: Multiculturalism increases race essentialism. *Social Psychological and Personality Science*, *10*(5), 681–689.

Wojcik, S., & Hughes, A. (2019). *U.S. adult Twitter users are younger and more likely to be Democrats than the general public. Most users rarely tweet, but the most prolific 10% create 80% of tweets from adult U.S. users*. Pew Research Center.

Woods, C. L., & Williams, D. E. (1976). Traits attributed to stuttering and normally fluent males. *Journal of Speech and Hearing Research*, *19*(2), 267–278.

Xu, L., & Zhang, W.-J. (2001). Comparison of different methods for variable selection. *Analytica Chimica Acta*, *446*(1–2), 475–481.

Yalcinkaya, N. S., Estrada-Villalta, S., & Adams, G. (2017). The (biological or cultural) essence of essentialism: Implications for policy support among dominant and subordinated groups. *Frontiers in Psychology*, *8*(900).

Ye, Y., & Gawronski, B. (2018). Validating the semantic misattribution procedure as an implicit measure of gender stereotyping: Gender stereotyping. *European Journal of Social Psychology*, *48*(3), 348–364.

Yzerbyt, V., Corneille, O., & Estrada, C. (2001). The interplay of subjective essentialism and entitativity in the formation of stereotypes. *Personality and Social Psychology Review*, *5*(2), 141–155.

Yzerbyt, V., Estrada, C., Corneille, O., Seron, E., & Demoulin, S. (2004). Subjective essentialism in action. In V. Yzerbyt, C. M. Judd, & O. Corneille (Eds.), *The psychology of group perception: Perceived variability, entitativity, and essentialism* (p. 18). Psychology Press.

Yzerbyt, V., Judd, C. M., & Corneille, O. (Eds.). (2004). *The psychology of group perception: Perceived variability, entitativity, and essentialism*. Psychology Press.

Yzerbyt, V., Rocher, S., & Schadron, G. (1997). Stereotypes as explanations A subjective essentialistic view of group perception. In R. Spears, P. J. Oakes, N. Ellemers, & S. A. Haslam (Eds.), *The Social Psychology of Stereotyping and Group Life*. Blackwell Publishers Ltd.

Zagefka, H., Nigbur, D., Gonzalez, R., & Tip, L. (2013). Why does ingroup essentialism increase prejudice against minority members? *International Journal of Psychology*, *48*(1), 60–68.

Zanna, M. P., & Hamilton, D. L. (1977). Further evidence for meaning change in impression formation. *Journal of Experimental Social Psychology*, *13*(3), 224–238.

Zou, L. X., & Cheryan, S. (2017). Two axes of subordination: A new model of racial position. *Journal of Personality and Social Psychology*, *112*(5), 696–717.

Appendix A: Locating Papers Using the Adjective Checklist Measure

The initial pool of papers for potential inclusion was identified by searching "adjective checklist stereotype" in Google Scholar. Each search result was read, up until ten in a row did not meet the inclusion criteria. The pool of papers for potential inclusion was expanded by looking through the citations of papers which passed all criteria for inclusion. This was done iteratively until no further papers were identified for potential inclusion.

To be included in this narrative review, the manuscript needed to have participants fill out an adjective checklist measure about a social group or group member. The measure needed to use a binary choice response format; measures were excluded if they instead used bipolar scales (e.g., Fowlie & Cooper, 1978; Gordon et al., 1989; Woods & Williams, 1976), trait ratings (e.g., Bishop et al., 2014; Ewens, 1969; Fryer & Cohen, 1988; Harper, 2007; Hopkins et al., 1997; Madon, 1997; Matheson & Strickland, 1986; Perlmutter, 1954; Philogène, 2001; Thorndike, 1977), or categorization of traits as associated with one group relative to another (Bastide & Van Den Berghe, 1957; Bendig & Hountras, 1958; Blake & Dennis, 1943; Der-Karabetian & Smith, 1977; Six & Eckes, 1991; Tsatsoulis-Bonnekessen, 1993; Willemsen, 2002; J. E. Williams & Bennett, 1975; J. E. Williams & Best, 1982). The measure needed to measure attitudes towards a social group; measures were excluded if they only assessed impressions of individual people already known to the participant (e.g., Bourne, 1977; MacIntyre, 1981), including the self (e.g., Carlson & Levy, 1968; Gough, 1960; Koestner et al., 1990; Parish & Taylor, 1978) or the participant's current mood (e.g., Matthews et al., 1990). The measure had to be available in English; measures were excluded if they were only available in other languages (Haratani et al., 1960; Rudolf et al., 1961). The paper also had to provide enough information to understand how the measure was used; in two cases, documents were excluded for lacking substantive

information about process (Cantril & Strunk, 1951; Freeman, 1969). Finally, seven works cited within this literature were inaccessible and could not be evaluated for inclusion (e.g., unpublished manuscripts; Carneggie, 1990; Edelstein, 1972; Koeske, 1971; Kusunoti, 1936; Lewis et al., 1972; Mukerji, 1951; Paxton, 1983).

A set of 49 papers identified during the search process used an adjective checklist measure, but only to the end of determining affective attitudes towards a particular group. For example, Wigginton, Wiersma, Sherman, and Rubin (2009) asked participants to fill out the Siperstein (1980) adjective checklist in response to images of average-weight and obese children. Participant responses were scored by subtracting the number of negative adjectives chosen from the number of positive adjectives chosen and adding twenty; this score was used as the dependent variable from this measure. Because these papers do not investigate stereotype content beyond this affective output, they are excluded from the general narrative review. However, they are retained in the data set of papers which have used adjective checklist measures. Some measures (e.g., Gough & Heilbrun, 1983) are used differently in different papers; only papers that met the criteria were retained.

Appendix B: All Search Terms Used for Online Adjective Sourcing

Table 28. Online search terms for each social group.

Social Domain	Social Group	Online Search Terms
Body Type	Fat people	Fat Folk, Fat Folks, Fat Folx, Fat People, Fat Person, Fat Ppl, FatFolk, FatFolks, FatFolx, FatPeople, FatPerson, FatPpl, Obese Folk, Obese Folks, Obese Folx, Obese People, Obese Person, Obese Ppl, ObeseFolk, ObeseFolks, ObeseFolx, ObesePeople, ObesePerson, ObesePpl
	Muscular people	Muscular Folk, Muscular Folks, Muscular Folx, Muscular People, Muscular Person, Muscular Ppl, MuscularFolk, MuscularFolks, MuscularFolx, MuscularPeople, MuscularPerson, MuscularPpl, Strong Folk, Strong Folks, Strong Folx, Strong People, Strong Person, Strong Ppl, StrongFolk, StrongFolks, StrongFolx, StrongPeople, StrongPerson, StrongPpl
	Short people	Short Folk, Short Folks, Short Folx, Short People, Short Person, Short Ppl, ShortFolk, ShortFolks, ShortFolx, ShortPeople, ShortPerson, ShortPpl
	Skinny people	Skinny Folk, Skinny Folks, Skinny Folx, Skinny People, Skinny Person, Skinny Ppl, SkinnyFolk, SkinnyFolks, SkinnyFolx, SkinnyPeople, SkinnyPerson, SkinnyPpl, Thin Folk, Thin Folks, Thin Folx, Thin People, Thin Person, Thin Ppl, ThinFolk, ThinFolks, ThinFolx, ThinPeople, ThinPerson, ThinPpl
	Tall people	Tall Folk, Tall Folks, Tall Folx, Tall People, Tall Person, Tall Ppl, TallFolk, TallFolks, TallFolx, TallPeople, TallPerson, TallPpl
Diet	Dieters	Dieter, Dieters, Dieting Folk, Dieting Folks, Dieting Folx, Dieting People, Dieting Person, Dieting Ppl, DietingFolk, DietingFolks, DietingFolx, DietingPeople, DietingPerson, DietingPpl
	Meat-eaters	Carnivorous Folk, Carnivorous Folks, Carnivorous Folx, Carnivorous People, Carnivorous Person, Carnivorous Ppl, CarnivorousFolk, CarnivorousFolks, CarnivorousFolx, CarnivorousPeople, CarnivorousPerson, CarnivorousPpl, Meat Eating Folk, Meat Eating Folks, Meat Eating Folx, Meat Eating People, Meat Eating Person, Meat Eating Ppl, Meat EatingFolk, Meat EatingFolks, Meat EatingFolx, Meat EatingPeople, Meat EatingPerson, Meat EatingPpl, Meat-Eating Folk, Meat-Eating Folks, Meat-Eating Folx, Meat-Eating People, Meat-Eating Person, Meat-Eating Ppl, Meat-EatingFolk, Meat-EatingFolks, Meat-EatingFolx, Meat-EatingPeople, Meat-EatingPerson, Meat-EatingPpl, MeatEating Folk, MeatEating Folks, MeatEating Folx, MeatEating People, MeatEating

Social Domain	Social Group	Online Search Terms
		Person, MeatEating Ppl, MeatEatingFolk, MeatEatingFolks, MeatEatingFolx, MeatEatingPeople, MeatEatingPerson, MeatEatingPpl
	Omnivores	Omnivorous Folk, Omnivorous Folks, Omnivorous Folx, Omnivorous People, Omnivorous Person, Omnivorous Ppl, OmnivorousFolk, OmnivorousFolks, OmnivorousFolx, OmnivorousPeople, OmnivorousPerson, OmnivorousPpl
	Pescatarians	Pescatarian, PescatarianFolk, PescatarianFolks, PescatarianFolx, PescatarianPeople, PescatarianPerson, PescatarianPpl, Pescatarians
	Vegans	Vegan, VeganFolk, VeganFolks, VeganFolx, VeganPeople, VeganPerson, VeganPpl, Vegans
	Vegetarians	Vegetarian, VegetarianFolk, VegetarianFolks, VegetarianFolx, VegetarianPeople, VegetarianPerson, VegetarianPpl, Vegetarians
Employment Class	Blue collar workers	Blue Collar Folk, Blue Collar Folks, Blue Collar Folx, Blue Collar People, Blue Collar Person, Blue Collar Ppl, Blue Collar Worker, Blue Collar Workers, Blue CollarFolk, Blue CollarFolks, Blue CollarFolx, Blue CollarPeople, Blue CollarPerson, Blue CollarPpl, Blue-Collar Folk, Blue-Collar Folks, Blue-Collar Folx, Blue-Collar People, Blue-Collar Person, Blue-Collar Ppl, Blue-Collar Worker, Blue-Collar Workers, Blue-CollarFolk, Blue-CollarFolks, Blue-CollarFolx, Blue-CollarPeople, Blue-CollarPerson, Blue-CollarPpl, BlueCollar Folk, BlueCollar Folks, BlueCollar Folx, BlueCollar People, BlueCollar Person, BlueCollar Ppl, BlueCollarFolk, BlueCollarFolks, BlueCollarFolx, BlueCollarPeople, BlueCollarPerson, BlueCollarPpl, BlueCollarWorker, BlueCollarWorkers
	Unemployed people	Jobless, JoblessFolk, JoblessFolks, JoblessFolx, JoblessPeople, JoblessPerson, JoblessPpl, Unemployed, UnemployedFolk, UnemployedFolks, UnemployedFolx, UnemployedPeople, UnemployedPerson, UnemployedPpl
	White collar workers	White Collar Folk, White Collar Folks, White Collar Folx, White Collar People, White Collar Person, White Collar Ppl, White Collar Worker, White Collar Workers, White CollarFolk, White CollarFolks, White CollarFolx, White CollarPeople, White CollarPerson, White CollarPpl, White-Collar Folk, White-Collar Folks, White-Collar Folx, White-Collar People, White-Collar Person, White-Collar Ppl, White-Collar Worker, White-Collar Workers, White-CollarFolk, White-CollarFolks, White-CollarFolx, White-CollarPeople, White-CollarPerson, White-CollarPpl, WhiteCollar Folk, WhiteCollar Folks, WhiteCollar Folx, WhiteCollar People, WhiteCollar Person,

Social Domain	Social Group	Online Search Terms
		WhiteCollar Ppl, WhiteCollarFolk, WhiteCollarFolks, WhiteCollarFolx, WhiteCollarPeople, WhiteCollarPerson, WhiteCollarPpl, WhiteCollarWorker, WhiteCollarWorkers
Eye Color	Blue-eyed people	Blue Eyed Folk, Blue Eyed Folks, Blue Eyed Folx, Blue Eyed People, Blue Eyed Person, Blue Eyed Ppl, Blue EyedFolk, Blue EyedFolks, Blue EyedFolx, Blue EyedPeople, Blue EyedPerson, Blue EyedPpl, Blue-Eyed Folk, Blue-Eyed Folks, Blue-Eyed Folx, Blue-Eyed People, Blue-Eyed Person, Blue-Eyed Ppl, Blue-EyedFolk, Blue-EyedFolks, Blue-EyedFolx, Blue-EyedPeople, Blue-EyedPerson, Blue-EyedPpl, BlueEyed Folk, BlueEyed Folks, BlueEyed Folx, BlueEyed People, BlueEyed Person, BlueEyed Ppl, BlueEyedFolk, BlueEyedFolks, BlueEyedFolx, BlueEyedPeople, BlueEyedPerson, BlueEyedPpl
	Brown-eyed people	Brown Eyed Folk, Brown Eyed Folks, Brown Eyed Folx, Brown Eyed People, Brown Eyed Person, Brown Eyed Ppl, Brown EyedFolk, Brown EyedFolks, Brown EyedFolx, Brown EyedPeople, Brown EyedPerson, Brown EyedPpl, Brown-Eyed Folk, Brown-Eyed Folks, Brown-Eyed Folx, Brown-Eyed People, Brown-Eyed Person, Brown-Eyed Ppl, Brown-EyedFolk, Brown-EyedFolks, Brown-EyedFolx, Brown-EyedPeople, Brown-EyedPerson, Brown-EyedPpl, BrownEyed Folk, BrownEyed Folks, BrownEyed Folx, BrownEyed People, BrownEyed Person, BrownEyed Ppl, BrownEyedFolk, BrownEyedFolks, BrownEyedFolx, BrownEyedPeople, BrownEyedPerson, BrownEyedPpl
	Green-eyed people	Green Eyed Folk, Green Eyed Folks, Green Eyed Folx, Green Eyed People, Green Eyed Person, Green Eyed Ppl, Green EyedFolk, Green EyedFolks, Green EyedFolx, Green EyedPeople, Green EyedPerson, Green EyedPpl, Green-Eyed Folk, Green-Eyed Folks, Green-Eyed Folx, Green-Eyed People, Green-Eyed Person, Green-Eyed Ppl, Green-EyedFolk, Green-EyedFolks, Green-EyedFolx, Green-EyedPeople, Green-EyedPerson, Green-EyedPpl, GreenEyed Folk, GreenEyed Folks, GreenEyed Folx, GreenEyed People, GreenEyed Person, GreenEyed Ppl, GreenEyedFolk, GreenEyedFolks, GreenEyedFolx, GreenEyedPeople, GreenEyedPerson, GreenEyedPpl
Gender	Boys	Boy, Boys, Lad, Lads
	Gender non-binary people	Enbie, EnbieFolk, EnbieFolks, EnbieFolx, EnbiePeople, EnbiePerson, EnbiePpl, Enbies, Enby, EnbyFolk, EnbyFolks, EnbyFolx, EnbyPeople, EnbyPerson, EnbyPpl, Enbys, Non-Binary, Non-BinaryFolk, Non-BinaryFolks, Non-BinaryFolx, Non-BinaryPeople, Non-BinaryPerson, Non-BinaryPpl, Nonbinary, NonbinaryFolk, NonbinaryFolks, NonbinaryFolx, NonbinaryPeople, NonbinaryPerson, NonbinaryPpl

Social Domain	Social Group	Online Search Terms
	Gender non-conforming people	Gender Non-conforming, Gender Non-conformingFolk, Gender Non-conformingFolks, Gender Non-conformingFolx, Gender Non-conformingPeople, Gender Non-conformingPerson, Gender Non-conformingPpl, Gender Nonconforming, Gender NonconformingFolk, Gender NonconformingFolks, Gender NonconformingFolx, Gender NonconformingPeople, Gender NonconformingPerson, Gender NonconformingPpl, GenderNon-conforming, GenderNon-conformingFolk, GenderNon-conformingFolks, GenderNon-conformingFolx, GenderNon-conformingPeople, GenderNon-conformingPerson, GenderNon-conformingPpl, GenderNonconforming, GenderNonconformingFolk, GenderNonconformingFolks, GenderNonconformingFolx, GenderNonconformingPeople, GenderNonconformingPerson, GenderNonconformingPpl
	Genderqueer people	Gender queer, Gender queerFolk, Gender queerFolks, Gender queerFolx, Gender queerPeople, Gender queerPerson, Gender queerPpl, Gender queers, Gender-queer, Gender-queerFolk, Gender-queerFolks, Gender-queerFolx, Gender-queerPeople, Gender-queerPerson, Gender-queerPpl, Genderqueer, GenderqueerFolk, GenderqueerFolks, GenderqueerFolx, GenderqueerPeople, GenderqueerPerson, GenderqueerPpl
	Girls	Girl, Girls, Lass, Lasses
	Men	Dude, Dudes, Gentleman, Gentlemen, Guy, Guys, Male, Males, Man, Men
	Transgender people	Tranny, TrannyFolk, TrannyFolks, TrannyFolx, TrannyPeople, TrannyPerson, TrannyPpl, Trannys, Trans, TransFolk, TransFolks, TransFolx, Transgender, TransgenderFolk, TransgenderFolks, TransgenderFolx, TransgenderPeople, TransgenderPerson, TransgenderPpl, Transgenders, TransPeople, TransPerson, TransPpl, Transsexual, TranssexualFolk, TranssexualFolks, TranssexualFolx, TranssexualPeople, TranssexualPerson, TranssexualPpl, Transsexuals
	Transmen	Transman, Transmasc, TransmascFolk, TransmascFolks, TransmascFolx, TransmascPeople, TransmascPerson, TransmascPpl, Transmascs, Transmen
	Transwomen	Transfemme, TransfemmeFolk, TransfemmeFolks, TransfemmeFolx, TransfemmePeople, TransfemmePerson, TransfemmePpl, Transfemmes, Transwoman, Transwomen
	Women	Female, Females, Gal, Gals, Ladies, Lady, Woman, Women
Handedness	Ambidextrous people	Ambidextrous, AmbidextrousFolk, AmbidextrousFolks, AmbidextrousFolx, AmbidextrousPeople, AmbidextrousPerson, AmbidextrousPpl
	Left-handed people	Left Handed, Left HandedFolk, Left HandedFolks, Left HandedFolx, Left HandedPeople, Left HandedPerson, Left HandedPpl, Left-Handed, Left-HandedFolk, Left-HandedFolks, Left-

Social Domain	Social Group	Online Search Terms
		HandedFolx, Left-HandedPeople, Left-HandedPerson, Left-HandedPpl, LeftHanded, LeftHandedFolk, LeftHandedFolks, LeftHandedFolx, LeftHandedPeople, LeftHandedPerson, LeftHandedPpl
	Right-handed people	Right Handed, Right HandedFolk, Right HandedFolks, Right HandedFolx, Right HandedPeople, Right HandedPerson, Right HandedPpl, Right-Handed, Right-HandedFolk, Right-HandedFolks, Right-HandedFolx, Right-HandedPeople, Right-HandedPerson, Right-HandedPpl, RightHanded, RightHandedFolk, RightHandedFolks, RightHandedFolx, RightHandedPeople, RightHandedPerson, RightHandedPpl
Hobbies	Gardeners	Gardener, Gardeners
	Movie buffs	Movie Buff, Movie Buffs
	Sports fans	Sports Fan, Sports Fans
	Woodworkers	Carpenter, Carpenters, Woodworker, Woodworkers
Kinship	Aunts	Aunt, Auntie, Aunties, Aunts, Aunty, Auntys
	Brothers	Brother, Brothers
	Children	Child, Children, Kid, Kids
	Cousins	Cousin, Cousins
	Daughters	Daughter, Daughters
	Fathers	Dad, Daddy, Daddys, Dads, Father, Fathers, Papa, Papas, Poppa, Poppas
	Grandchildren	Grandchild, Grandchildren, Grandkid, Grandkids
	Granddaughters	Granddaughter, Granddaughters
	Grandfathers	Granddad, Granddaddy, Granddads, Grandfather, Grandfathers, Grandpa, Grandpappy
	Grandmothers	Gram, Gramps, Grandma, Grandmama, Grandmamas, Grandmas, Grandmom, Grandmoms, Grandmother, Grandmothers, Grannies, Granny, NaNanas
	Grandparents	Grandparent, Grandparents, Grandpas
	Grandsons	Grandson, Grandsons
	Mothers	Mama, Mamas, Mamma, Mammas, Mom, Mommy, Mommys, Moms, Mother, Mothers
	Parents	Parent, Parents
	Siblings	Sibling, Siblings
	Sisters	Sister, Sisters
	Sons	Son, Sons

Social Domain	Social Group	Online Search Terms
	Uncles	Uncle, Uncles
Occupation	Administrative assistants	Admin Assistant, Admin Assistants, Admin Asst, Admin Assts, AdminAssistant, AdminAssistants, AdminAsst, AdminAssts, Administrative Assistant, Administrative Assistants, AdministrativeAssistant, AdministrativeAssistants
	Bankers	Banker, Bankers
	Construction workers	Builder, Builders, Construction Worker, Construction Workers, ConstructionWorker, ConstructionWorkers, Contractor, Contractors
	Secretaries	Secretaries, Secretary
	Students	CollegeStudent, CollegeStudents, GradStudent, GradStudents, GraduateStudent, GraduateStudents, HighSchoolStudent, HighSchoolStudents, Student, StudentFolk, StudentFolks, StudentFolx, StudentPeople, StudentPerson, StudentPpl, Students, Undergrad, UndergradFolk, UndergradFolks, UndergradFolx, UndergradPeople, UndergradPerson, UndergradPpl, Undergrads, UndergradStudent, UndergradStudents, Undergraduate, UndergraduateFolk, UndergraduateFolks, UndergraduateFolx, UndergraduatePeople, UndergraduatePerson, UndergraduatePpl, Undergraduates, UndergraduateStudent, UndergraduateStudents
	Tech entrepreneurs	Tech Entrepreneur, Tech Entrepreneurs, TechEntrepreneur, TechEntrepreneurs, Technology Entrepreneur, Technology Entrepreneurs, TechnologyEntrepreneur, TechnologyEntrepreneurs
Political Ideology	Conservatives	Conservative Folk, Conservative Folks, Conservative Folx, Conservative People, Conservative Person, Conservative Ppl, ConservativeFolk, ConservativeFolks, ConservativeFolx, ConservativePeople, ConservativePerson, ConservativePpl, Conservatives
	Democrats	Democrats, Dems
	Liberals	Liberal Folk, Liberal Folks, Liberal Folx, Liberal People, Liberal Person, Liberal Ppl, LiberalFolk, LiberalFolks, LiberalFolx, LiberalPeople, LiberalPerson, LiberalPpl, Liberals, Libs
	Libertarians	Libertarian Folk, Libertarian Folks, Libertarian Folx, Libertarian People, Libertarian Person, Libertarian Ppl, LibertarianFolk, LibertarianFolks, LibertarianFolx, LibertarianPeople, LibertarianPerson, LibertarianPpl, Libertarians
	Progressives	Progressive Folk, Progressive Folks, Progressive Folx, Progressive People, Progressive Person, Progressive Ppl, ProgressiveFolk, ProgressiveFolks, ProgressiveFolx, ProgressivePeople, ProgressivePerson, ProgressivePpl, Progressives, Progs

Social Domain	Social Group	Online Search Terms
	Republicans	GOPer, GOPers, Republican Folk, Republican Folks, Republican Folx, Republican People, Republican Person, Republican Ppl, RepublicanFolk, RepublicanFolks, RepublicanFolx, RepublicanPeople, RepublicanPerson, RepublicanPpl, Republicans
Race-Ethnicity	Asian people	Asian American, Asian AmericanFolk, Asian AmericanFolks, Asian AmericanFolx, Asian AmericanPeople, Asian AmericanPerson, Asian AmericanPpl, Asian Americans, Asian Folk, Asian Folks, Asian Folx, Asian People, Asian Person, Asian Ppl, Asian-American, Asian-AmericanFolk, Asian-AmericanFolks, Asian-AmericanFolx, Asian-AmericanPeople, Asian-AmericanPerson, Asian-AmericanPpl, Asian-Americans, AsianAmerican, AsianAmericanFolk, AsianAmericanFolks, AsianAmericanFolx, AsianAmericanPeople, AsianAmericanPerson, AsianAmericanPpl, AsianAmericans, AsianFolk, AsianFolks, AsianFolx, AsianPeople, AsianPerson, AsianPpl, Asians
	Black people	African American, African AmericanFolk, African AmericanFolks, African AmericanFolx, African AmericanPeople, African AmericanPerson, African AmericanPpl, African Americans, African-American, African-AmericanFolk, African-AmericanFolks, African-AmericanFolx, African-AmericanPeople, African-AmericanPerson, African-AmericanPpl, African-Americans, AfricanAmerican, AfricanAmericanFolk, AfricanAmericanFolks, AfricanAmericanFolx, AfricanAmericanPeople, AfricanAmericanPerson, AfricanAmericanPpl, AfricanAmericans, Black American, Black AmericanFolk, Black AmericanFolks, Black AmericanFolx, Black AmericanPeople, Black AmericanPerson, Black AmericanPpl, Black Americans, Black Folk, Black Folks, Black Folx, Black People, Black Person, Black Ppl, Black-American, Black-AmericanFolk, Black-AmericanFolks, Black-AmericanFolx, Black-AmericanPeople, Black-AmericanPerson, Black-AmericanPpl, Black-Americans, BlackAmerican, BlackAmericanFolk, BlackAmericanFolks, BlackAmericanFolx, BlackAmericanPeople, BlackAmericanPerson, BlackAmericanPpl, BlackAmericans, BlackFolk, BlackFolks, BlackFolx, BlackPeople, BlackPerson, BlackPpl, Blacks, Nigga, Niggas
	Hispanic people	Hispanic Folk, Hispanic Folks, Hispanic Folx, Hispanic People, Hispanic Person, Hispanic Ppl, HispanicFolk, HispanicFolks, HispanicFolx, HispanicPeople, HispanicPerson, HispanicPpl, Hispanics, Latin American, Latin AmericanFolk, Latin AmericanFolks, Latin AmericanFolx, Latin AmericanPeople, Latin AmericanPerson, Latin AmericanPpl, Latin Americans, Latin-American, Latin-AmericanFolk, Latin-AmericanFolks, Latin-AmericanFolx, Latin-AmericanPeople, Latin-AmericanPerson, Latin-AmericanPpl, Latin-Americans, Latina Folk, Latina Folks, Latina Folx,

Social Domain	Social Group	Online Search Terms
		Latina People, Latina Person, Latina Ppl, LatinaFolk, LatinaFolks, LatinaFolx, LatinAmerican, LatinAmericanFolk, LatinAmericanFolks, LatinAmericanFolx, LatinAmericanPeople, LatinAmericanPerson, LatinAmericanPpl, LatinAmericans, LatinaPeople, LatinaPerson, LatinaPpl, Latinas, Latine Folk, Latine Folks, Latine Folx, Latine People, Latine Person, Latine Ppl, LatineFolk, LatineFolks, LatineFolx, LatinePeople, LatinePerson, LatinePpl, Latines, Latino Folk, Latino Folks, Latino Folx, Latino People, Latino Person, Latino Ppl, LatinoFolk, LatinoFolks, LatinoFolx, LatinoPeople, LatinoPerson, LatinoPpl, Latinos, Latinx, LatinxFolk, LatinxFolks, LatinxFolx, LatinxPeople, LatinxPerson, LatinxPpl
	Native American people	American Indian, American IndianFolk, American IndianFolks, American IndianFolx, American IndianPeople, American IndianPerson, American IndianPpl, American Indians, American-Indian, American-IndianFolk, American-IndianFolks, American-IndianFolx, American-IndianPeople, American-IndianPerson, American-IndianPpl, American-Indians, AmericanIndian, AmericanIndianFolk, AmericanIndianFolks, AmericanIndianFolx, AmericanIndianPeople, AmericanIndianPerson, AmericanIndianPpl, AmericanIndians, Indigenous American, Indigenous AmericanFolk, Indigenous AmericanFolks, Indigenous AmericanFolx, Indigenous AmericanPeople, Indigenous AmericanPerson, Indigenous AmericanPpl, Indigenous Americans, Indigenous Folk, Indigenous Folks, Indigenous Folx, Indigenous People, Indigenous Person, Indigenous Ppl, Indigenous-American, Indigenous-AmericanFolk, Indigenous-AmericanFolks, Indigenous-AmericanFolx, Indigenous-AmericanPeople, Indigenous-AmericanPerson, Indigenous-AmericanPpl, Indigenous-Americans, IndigenousAmerican, IndigenousAmericanFolk, IndigenousAmericanFolks, IndigenousAmericanFolx, IndigenousAmericanPeople, IndigenousAmericanPerson, IndigenousAmericanPpl, IndigenousAmericans, IndigenousFolk, IndigenousFolks, IndigenousFolx, IndigenousPeople, IndigenousPerson, IndigenousPpl, Native American Folk, Native American Folks, Native American Folx, Native American People, Native American Person, Native American Ppl, Native AmericanFolk, Native AmericanFolks, Native AmericanFolx, Native AmericanPeople, Native AmericanPerson, Native AmericanPpl, Native Americans, Native Folk, Native Folks, Native Folx, Native People, Native Person, Native Ppl, Native-American Folk, Native-American Folks, Native-American Folx, Native-American People, Native-American Person, Native-American Ppl, Native-AmericanFolk, Native-AmericanFolks, Native-AmericanFolx, Native-AmericanPeople, Native-AmericanPerson, Native-AmericanPpl, Native-Americans, NativeAmerican Folk,

Social Domain	Social Group	Online Search Terms
		NativeAmerican Folks, NativeAmerican Folx, NativeAmerican People, NativeAmerican Person, NativeAmerican Ppl, NativeAmericanFolk, NativeAmericanFolks, NativeAmericanFolx, NativeAmericanPeople, NativeAmericanPerson, NativeAmericanPpl, NativeAmericans, NativeFolk, NativeFolks, NativeFolx, NativePeople, NativePerson, NativePpl
	White people	Caucasian Folk, Caucasian Folks, Caucasian Folx, Caucasian People, Caucasian Person, Caucasian Ppl, CaucasianFolk, CaucasianFolks, CaucasianFolx, CaucasianPeople, CaucasianPerson, CaucasianPpl, Caucasians, European American, European AmericanFolk, European AmericanFolks, European AmericanFolx, European AmericanPeople, European AmericanPerson, European AmericanPpl, European Americans, European-American, European-AmericanFolk, European-AmericanFolks, European-AmericanFolx, European-AmericanPeople, European-AmericanPerson, European-AmericanPpl, European-Americans, EuropeanAmerican, EuropeanAmericanFolk, EuropeanAmericanFolks, EuropeanAmericanFolx, EuropeanAmericanPeople, EuropeanAmericanPerson, EuropeanAmericanPpl, EuropeanAmericans, White American, White AmericanFolk, White AmericanFolks, White AmericanFolx, White AmericanPeople, White AmericanPerson, White AmericanPpl, White Americans, White Folk, White Folks, White Folx, White People, White Person, White Ppl, White-American, White-AmericanFolk, White-AmericanFolks, White-AmericanFolx, White-AmericanPeople, White-AmericanPerson, White-AmericanPpl, White-Americans, WhiteAmerican, WhiteAmericanFolk, WhiteAmericanFolks, WhiteAmericanFolx, WhiteAmericanPeople, WhiteAmericanPerson, WhiteAmericanPpl, WhiteAmericans, WhiteFolk, WhiteFolks, WhiteFolx, WhitePeople, WhitePerson, WhitePpl, Whites, Wypipo, Wyppl, Wyppo, Yt Folk, Yt Folks, Yt Folx, Yt People, Yt Person, Yt Ppl, YtFolk, YtFolks, YtFolx, YtPeople, YtPerson, YtPpl
Religion	Agnostic people	Agnostic, AgnosticFolk, AgnosticFolks, AgnosticFolx, AgnosticPeople, AgnosticPerson, AgnosticPpl, Agnostics
	Atheists	Atheist, AtheistFolk, AtheistFolks, AtheistFolx, AtheistPeople, AtheistPerson, AtheistPpl, Atheists
	Buddhists	Buddhist, BuddhistFolk, BuddhistFolks, BuddhistFolx, BuddhistPeople, BuddhistPerson, BuddhistPpl, Buddhists
	Catholics	Catholic, CatholicFolk, CatholicFolks, CatholicFolx, CatholicPeople, CatholicPerson, CatholicPpl, Catholics

Social Domain	Social Group	Online Search Terms
	Christians	Christian, ChristianFolk, ChristianFolks, ChristianFolx, ChristianPeople, ChristianPerson, ChristianPpl, Christians
	Hindus	Hindu, HinduFolk, HinduFolks, HinduFolx, HinduPeople, HinduPerson, HinduPpl, Hindus
	Jewish people	Jew, Jewish Folk, Jewish Folks, Jewish Folx, Jewish People, Jewish Person, Jewish Ppl, JewishFolk, JewishFolks, JewishFolx, JewishPeople, JewishPerson, JewishPpl, Jews
	Muslims	Muslim, MuslimFolk, MuslimFolks, MuslimFolx, MuslimPeople, MuslimPerson, MuslimPpl, Muslims
	Protestants	Protestant, ProtestantFolk, ProtestantFolks, ProtestantFolx, ProtestantPeople, ProtestantPerson, ProtestantPpl, Protestants
	Quakers	Quaker, QuakerFolk, QuakerFolks, QuakerFolx, QuakerPeople, QuakerPerson, QuakerPpl, Quakers
School Tropes	Jocks	Jock, Jocks, Student Athlete, Student Athletes, Student-Athlete, Student-Athletes, StudentAthlete, StudentAthletes
	Loners	Loner, Loners
	Nerds	Dork, Dorks, Geek, Geeks, Nerd, Nerds
	Stoners	Pothead, Potheads, Stoner, Stoners
Sexuality	Bisexual people	Bi Folk, Bi Folks, Bi Folx, Bi People, Bi Person, Bi Ppl, BiFolk, BiFolks, BiFolx, BiPeople, BiPerson, BiPpl, Bisexual, BisexualFolk, BisexualFolks, BisexualFolx, BisexualPeople, BisexualPerson, BisexualPpl, Bisexuals
	Gay people	Gay Folk, Gay Folks, Gay Folx, Gay People, Gay Person, Gay Ppl, GayFolk, GayFolks, GayFolx, GayPeople, GayPerson, GayPpl, Gays
	Heterosexual people	Hetero, HeteroFolk, HeteroFolks, HeteroFolx, HeteroPeople, HeteroPerson, HeteroPpl, Heteros, Heterosexual, HeterosexualFolk, HeterosexualFolks, HeterosexualFolx, HeterosexualPeople, HeterosexualPerson, HeterosexualPpl, Heterosexuals, Str8 Folk, Str8 Folks, Str8 Folx, Str8 People, Str8 Person, Str8 Ppl, Str8Folk, Str8Folks, Str8Folx, Str8People, Str8Person, Str8Ppl, Straight Folk, Straight Folks, Straight Folx, Straight People, Straight Person, Straight Ppl, StraightFolk, StraightFolks, StraightFolx, StraightPeople, StraightPerson, StraightPpl
	Homosexual people	Homo, HomoFolk, HomoFolks, HomoFolx, HomoPeople, HomoPerson, HomoPpl, Homos, Homosexual, HomosexualFolk, HomosexualFolks, HomosexualFolx, HomosexualPeople, HomosexualPerson, HomosexualPpl, Homosexuals

Social Domain	Social Group	Online Search Terms
	Pansexual people	Pan Folk, Pan Folks, Pan Folx, Pan People, Pan Person, Pan Ppl, PanFolk, PanFolks, PanFolx, PanPeople, PanPerson, PanPpl, Pansexual, PansexualFolk, PansexualFolks, PansexualFolx, PansexualPeople, PansexualPerson, PansexualPpl, Pansexuals
	Queer people	Queer Folk, Queer Folks, Queer Folx, Queer People, Queer Person, Queer Ppl, QueerFolk, QueerFolks, QueerFolx, QueerPeople, QueerPerson, QueerPpl, Queers
Sexuality-Gender	Bisexual men	Bi Man, Bi Men, BiMan, BiMen, Bisexual Man, Bisexual Men, BisexualMan, BisexualMen
	Bisexual women	Bi Woman, Bi Women, Bisexual Woman, Bisexual Women, BisexualWoman, BisexualWomen, BiWoman, BiWomen
	Gay men	Gay Man, Gay Men
	Lesbian women	Lesbian, Lesbian Woman, Lesbian Women, LesbianFolk, LesbianFolks, LesbianFolx, LesbianPeople, LesbianPerson, LesbianPpl, Lesbians, Lesbo, LesboFolk, LesboFolks, LesboFolx, LesboPeople, LesboPerson, LesboPpl, Lesbos
US Region	East Coasters	East Coast Folk, East Coast Folks, East Coast Folx, East Coast People, East Coast Person, East Coast Ppl, East Coaster, East Coasters, East CoastFolk, East CoastFolks, East CoastFolx, East CoastPeople, East CoastPerson, East CoastPpl, East-Coast Folk, East-Coast Folks, East-Coast Folx, East-Coast People, East-Coast Person, East-Coast Ppl, East-Coaster, East-Coasters, East-CoastFolk, East-CoastFolks, East-CoastFolx, East-CoastPeople, East-CoastPerson, East-CoastPpl, EastCoast Folk, EastCoast Folks, EastCoast Folx, EastCoast People, EastCoast Person, EastCoast Ppl, EastCoaster, EastCoasters, EastCoastFolk, EastCoastFolks, EastCoastFolx, EastCoastPeople, EastCoastPerson, EastCoastPpl
	Midwesterners	Midwest Folk, Midwest Folks, Midwest Folx, Midwest People, Midwest Person, Midwest Ppl, Midwestern Folk, Midwestern Folks, Midwestern Folx, Midwestern People, Midwestern Person, Midwestern Ppl, Midwesterner, Midwesterners, MidwesternFolk, MidwesternFolks, MidwesternFolx, MidwesternPeople, MidwesternPerson, MidwesternPpl, MidwestFolk, MidwestFolks, MidwestFolx, MidwestPeople, MidwestPerson, MidwestPpl
	Southerners	Southern Folk, Southern Folks, Southern Folx, Southern People, Southern Person, Southern Ppl, Southerner, Southerners, SouthernFolk, SouthernFolks, SouthernFolx, SouthernPeople, SouthernPerson, SouthernPpl
	West Coasters	West Coast Folk, West Coast Folks, West Coast Folx, West Coast People, West Coast Person, West Coast Ppl, West Coaster, West Coasters, West CoastFolk, West CoastFolks, West CoastFolx, West

Social Domain	Social Group	Online Search Terms
		CoastPeople, West CoastPerson, West CoastPpl, West-Coast Folk, West-Coast Folks, West-Coast Folx, West-Coast People, West-Coast Person, West-Coast Ppl, West-Coaster, West-Coasters, West-CoastFolk, West-CoastFolks, West-CoastFolx, West-CoastPeople, West-CoastPerson, West-CoastPpl, WestCoast Folk, WestCoast Folks, WestCoast Folx, WestCoast People, WestCoast Person, WestCoast Ppl, WestCoaster, WestCoasters, WestCoastFolk, WestCoastFolks, WestCoastFolx, WestCoastPeople, WestCoastPerson, WestCoastPpl
Wealth Class	Middle class people	Middle Class, Middle ClassFolk, Middle ClassFolks, Middle ClassFolx, Middle ClassPeople, Middle ClassPerson, Middle ClassPpl, MiddleClass, MiddleClassFolk, MiddleClassFolks, MiddleClassFolx, MiddleClassPeople, MiddleClassPerson, MiddleClassPpl
	Poor people	Broke Folk, Broke Folks, Broke Folx, Broke People, Broke Person, Broke Ppl, BrokeFolk, BrokeFolks, BrokeFolx, BrokePeople, BrokePerson, BrokePpl, Poor Folk, Poor Folks, Poor Folx, Poor People, Poor Person, Poor Ppl, PoorFolk, PoorFolks, PoorFolx, PoorPeople, PoorPerson, PoorPpl
	Rich people	Affluent Folk, Affluent Folks, Affluent Folx, Affluent People, Affluent Person, Affluent Ppl, AffluentFolk, AffluentFolks, AffluentFolx, AffluentPeople, AffluentPerson, AffluentPpl, Rich Folk, Rich Folks, Rich Folx, Rich People, Rich Person, Rich Ppl, RichFolk, RichFolks, RichFolx, RichPeople, RichPerson, RichPpl, Wealthy Folk, Wealthy Folks, Wealthy Folx, Wealthy People, Wealthy Person, Wealthy Ppl, WealthyFolk, WealthyFolks, WealthyFolx, WealthyPeople, WealthyPerson, WealthyPpl
	Working class people	Working Class, Working ClassFolk, Working ClassFolks, Working ClassFolx, Working ClassPeople, Working ClassPerson, Working ClassPpl, Working-Class, Working-ClassFolk, Working-ClassFolks, Working-ClassFolx, Working-ClassPeople, Working-ClassPerson, Working-ClassPpl, WorkingClass, WorkingClassFolk, WorkingClassFolks, WorkingClassFolx, WorkingClassPeople, WorkingClassPerson, WorkingClassPpl

Appendix C: Data Processing Pipeline for Social Media Data

Gathering adjectives from the social media sources was time and resource-intensive. A more detailed description of the pipeline used to gather them is defined separately for each source in Table 29. The table describes the processing steps, including both the analytic decisions that are made (e.g., defining the sampling frame) and the technical processes, and has information about the files produced at each stage. The data from each source goes through similar stages with technical differences because of different storage of the base data. I have organized the table with overall steps, where each step has one processing goal and may have sub-steps.

Table 29. Social media data processing steps in detail

Step of Process	Processing Goal	Reddit Data	Twitter Data
Step 1	Defining search terms	All references to the target groups, including a variety of synonyms and alternate references to the groups (see Appendix B: All Search Terms Used for Online Adjective Sourcing).	
Step 2	Defining population of online posts	All reddit data from 2019, as stored by the Pushshift dataset (Baumgartner et al., 2020). This includes both submissions and comments, which are stored separately.	All twitter data from 2019, as accessed by Twitter's academic API (*Advancing Academic Research with Twitter Data*, 2021).
Step 3	Defining sampling frame of online posts	Separate each month of 2019 data into files that are approximately 6.8 GB (5500000 rows of comments, or 2060000 rows of submissions). The number of rows per month does not divide perfectly into those, so the last file per month is smaller. From each month, randomly select two of the comment files and one submission file.	From all of 2019, randomly select one day per month. Then, from that one day, select a six-hour window that is broken into two three-hour windows. For each month, one of the three-hour windows is randomly assigned to set 1-1, and the other three-hour window is randomly assigned to set 1-2.
Step 3.5	Change files to	Partway through Step 4 on the 6.8 GB files, the computing resource stopped processing files of that	--

Step of Process	Processing Goal	Reddit Data	Twitter Data
	accommodate computer system	size. The only solution was to further subset the randomly-selected files into files that were about 500MB (381945 lines of a comment file, 163493 lines of a submission file). Each file is also converted into json format using the jq program (Dolan, 2020).	
Step 4	Gather all posts from sampling frame that include any of the search terms	From each file, use the grepl() function to see whether any of the search terms appear in the text as standalone words. This stops, for example, a post with the word *human* being tagged as containing *man*, or *granddaughter* being tagged as containing *daughter*. Retain only posts that reference at least one search term. If the scrape was being done on one of the 6.8 GB files (before the computer system changed), it is saved in subsets of the overall file. There are two sets of output files, because some synonyms for Mothers and Fathers were added after the main scraping was done. One set has most of the search terms, and the second set has the additional synonyms. Output files are Rdata files which correspond to the screened file. Each Rdata file contains the base information about the reddit post (comment or submission) and has a column for each search term.	Using the twitter API, request tweets from each three-hour window that include each of the terms. This is done separately for Set 1-1 and Set 1-2. Within the API call, request only tweets in English and omit retweets. The twitter API uses the search term as unique word, so does not (for example) return tweets which contain *human* as containing *man*. Issue 1: This will return tweets that contain the term in an included link (e.g., a tweet that included a link to an article stored at *https://www.nytimes.com/2021/05/26/opinion/democrats-republicans-wokeness-cancel-culture.html* would count as containing *Democrats* and *Republicans*, even if it otherwise didn't have the search term). This is surprising because the text of the tweet contains an altered version of the link, not the base link. To solve this, tweets are double-checked for whether they contain the search term. This retains tweets which reference a search term as a hashtag like *#Democrats*, or where a search term is

Step of Process	Processing Goal	Reddit Data	Twitter Data
		Each search term column has a 1 or 0 for whether or not the search term appears in the post text.	included in a twitter handle such as *@HouseDemocrats*. Issue 2: Each call to the API pulls up to 500 tweets. However, some search terms appear more than 500 times in a three-hour window (e.g., *White People*). The first 500 terms returned are not a random selection, so the function paginates through sets of 500 to get all the tweets referencing the term in the window. Output files are labelled by search term and set (1-1 or 1-2), and include all returned tweets for that search term in that window.
		Use a file that defines the target group of each search term.	
Step 5	Reorganize posts by target group	Because of computational memory limits, this is split into two parts (with two separate R scripts). Part 1: The same process is done on each of the files produced in Step 4. Step 5.1.1: The data is stacked, so there is one row per post-search term combination. Only rows which show the search term did appear in the post are retained. Step 5.1.2: Each search term is associated with a target group. The subset of rows associated with each target group is saved separately, producing many small CSV files.	The same process is done on each of the files produced in Step 4. Step 5.1: If tweets with exactly duplicate text were pulled for the same search term in the same time-window (i.e., Set 1-1 or Set 1-2), only one is retained. Duplicated tweets are aggregated across all files. Step 5.2: If country information is provided, all tweets produced outside the US are excluded. Distribution of Tweet location is aggregated across all files.

Step of Process	Processing Goal	Reddit Data	Twitter Data
		Part 1 Output Files: CSVs labelled with target group name. Numbered iteratively until all sources files have been accounted for. This means some target groups have a greater number of CSVs (e.g., *White People*) than other target groups (e.g., *Blue-Eyed People*) because they appeared in more files produced by Step 4. Part 2: Take all of the CSV files that are for a given target group and aggregate them together. Part 2 Output Files: Rdata file of all reddit submissions and comments for all search terms for each target group. This aggregates within target group for all search terms, but retains duplicate posts if the post references multiple search terms for the target group.	Step 5.3: The search term is referenced to the target group, and saved as a file for that target group. This aggregates within target group across all search terms, but still retains duplicate tweets if, for instance, the same tweet includes multiple search terms for the same target group (e.g., a tweet that contains both *working-class* and *workingclass*) Output files: Rdata file that has all duplicate tweets pulled for any search term. Rdata file that has distribution of locations for all search terms. Rdata file of all unique tweets pulled for all search terms for each target group
Step 6	Decide on specific content of interest out of a post	Any sentence which contains a reference to a target group, excluding the rest of the post. This can mean multiple sentences from the same post (e.g., the first and tenth sentence of a twenty-sentence post). A sentence is defined as text separated by a period (.), question mark (*?*), or exclamation point (*!*), which may or may not be followed by a space. The space is not a hard requirement because online text often attempts to prioritize brevity.	
Step 7	Retain only sentences that reference the target group	The same process is done for each Target Group file from Step 5. This is done separately for the Reddit and Twitter files, but follows the exact same process. Step 7.1: Convert all text to lowercase. Step 7.2: Separate each post into distinct sentences.	

Step of Process	Processing Goal	Reddit Data	Twitter Data
		Step 7.3: Check whether each sentence of a given post contains the search term for which the post was tagged. Omit any sentences which do not contain that specific search term. Step 7.4: In some cases, a sentence from a post includes multiple search terms for the same target group (e.g., a sentence that contains *working-class* and *workingclass*). Remove such duplicates. Step 7.5: Save Rdata file for all sentences which reference target group. Output files: One Rdata file per target group per social media source which retains only the sentences of the posts that specifically reference the target group.	
Step 8	Extract all individual strings that co-occur with each target group	Step 8 loops over all the files produced in Step 7 to create two overall output file, and one per-group output file. The same process occurs separately for Twitter and Reddit. Step 8.1: Get rid of all non-ASCII characters, numbers, and odd punctuation (e.g., *n* indicates a new line in the text). Step 8.2: Separate sentence into strings based on characters contained between spaces. This retains hyphenated words (e.g., "anti-racism") Step 8.3: Count how many times the string co-occurs with the target group. Save this as a csv and an Rdata file. This is what is used to identify the top adjectives per group per source. Step 8.4: Create output CSV showing how many unique posts referenced each target group.	

Appendix D: Online Screening

Participants from Amazon Mechanical Turk were required to answer all questions in the screening from to access the survey (see below; questions 1 - 7 were presented in a random order). To pass the screener, participants had to: indicate they were able to consent to participation, enter an age 18 or older, correctly answer the word-scramble question, say they were living in a US state or territory and enter a valid two-letter state or territory code. Data quality on Mturk is an ongoing concern (e.g., Ahler et al., 2019; Moss, 2018), and the screening process is intended to filter out bots (question 2, 3, 4, and 9), people who are not fluent in English (all questions, especially question 4), and people who do not live in the US (question 3, question 9). The word scramble question and the requested format of the state code varied across studies to avoid the accepted answers becoming shared knowledge among potential participants. Participants who passed the screener proceeded to the consent process. For Chapter III, the screener was also used to recruit a sample with only LGB+ participants (question 7).

These questions are to screen you for your eligibility to participate in the study. If you are not eligible, you will be screened out of the study after you answer these questions. Screener answers will be retained to ensure participants do not change their answers in order to qualify. Do not go back and redo the screener with different answers; your work would be rejected and you would be restricted from further participation.

(1) How old are you? (enter number of years) _____
(2) Are you able to consent to participating in research?
    □ Yes
    □ No
(3) Do you live in a US state or territory?
    □ Yes
    □ No
(4) *Example:* This question will require you to unscramble some words. How mnay suaqres are lfet if you have tewnty and gvie aawy twevle?
    [Choices 1 – 50]
(5) What is your gender?
    □ Man
    □ Woman
    □ Not Listed: _____

(6) Do you identify as transgender, or within the trans* umbrella?
- □ Yes
- □ No

(7) How do you identify?
- □ Heterosexual or Straight
- □ Gay or Lesbian
- □ Bisexual
- □ Uncertain or Questioning
- □ Not Listed: _____

(8) What is your mturk WorkerID? _____

(9) (if yes to question 3): In which US state or territory do you live? Enter the two-letter state code as you would in your address. Do not include periods or spaces. _____

Appendix E: Adjective Checklist Measure Instructions

These instructions were based on those in the existing literature, amended to reduce social desirability by making the task low-stakes, licensing gut reactions, and assuring participants there are no right or wrong answers.

In this study, we're gathering exploratory data about **everyday social perceptions**. We're interested in your **immediate gut reactions** to the questions. There are no right or wrong answers.

You will see the name of a social group [information about a person] and be asked to pick which traits, out of a list, you associate with that group [person]. You will do this ### times.

For each group, we are interested in studying the "typical" characteristics of the group. Not all group members are alike; we're interested in the description of the group in general. If you feel a trait describes the group in general, then click on that adjective. If you feel a trait does not describe the group in general, do not click on it.

[For each person, we're interested in your description of the person in general. If you feel a trait describes the person in general, then click on that adjective. If you feel a trait does not describe the person in general, do not click on it.]

If you can think of any characteristics that are missing from the list, please add them at the bottom of the list of traits.

Appendix F: Attention Check and Seriousness Check

The same attention (1) and seriousness (2) checks were used across studies.

(1) To show that you have read the instructions, please ignore the question and choose "I don't know" as your answer. How would you best describe your feelings right now?

□ Delighted

- Pleased
- Neutral
- Unhappy
- Miserable
- I don't know
- None of the above

(2) Now that you're done, we'd like to ask you whether you took this study seriously as you completed it. **There is no possible penalty to you, so please be honest**. We use this information in our analysis of the data.

- I took this study seriously
- I did not take this study seriously

Appendix G: Valence of the 213 Adjectives

Table 30. The mean and standard deviation of valence ratings for each normed adjective.

Adjective	Mean Valence	SD Valence	N of Raters	Adjective	Mean Valence	SD Valence	N of Raters
absent	3.04	1.49	54	just	5.22	1.33	49
academic	5.45	1.27	44	late	3.32	1.57	57
accepted	5.47	1.36	49	lazy	2.77	1.66	53
activist	4.80	1.51	55	left	3.78	1.21	49
adult	5.00	0.99	48	leftist	3.69	1.64	55
aggressive	3.20	1.77	44	legal	5.47	1.19	49
alive	5.58	1.48	45	level	5.15	1.13	48
alone	3.28	1.57	53	limited	3.92	1.41	51
amateur	3.96	1.32	53	local	4.47	1.22	45
ambitious	5.53	1.35	53	lost	2.83	1.44	54
annoying	2.74	1.69	57	loving	6.23	0.85	57
anonymous	3.91	1.21	43	low	3.07	1.30	54
anxious	3.50	1.39	56	loyal	5.74	1.15	47
arrogant	2.57	1.64	47	lucky	5.75	1.16	44
athletic	5.46	1.16	41	luxurious	5.15	1.37	48
available	5.77	0.97	43	mad	2.59	1.61	49
aware	5.54	1.25	56	magic	5.32	1.18	47
bad	2.04	1.41	57	married	5.04	1.68	49
biological	4.73	1.06	51	masculine	5.05	1.36	42
boring	2.89	1.39	47	mean	3.21	2.11	48
brave	5.77	1.30	53	medical	4.71	1.38	51
busy	4.37	1.29	54	mental	3.55	1.75	51

Adjective	Mean Valence	SD Valence	N of Raters	Adjective	Mean Valence	SD Valence	N of Raters
capitalist	4.65	1.29	57	messy	3.36	1.63	56
certain	5.14	1.07	50	misunderstood	3.26	1.65	47
cheap	3.15	1.46	46	moderate	4.53	1.02	47
close-minded	3.21	1.89	52	moral	5.42	1.47	45
comfortable	5.87	0.92	53	motivated	5.83	1.18	59
common	4.75	1.32	55	national	4.93	1.36	58
confident	5.82	1.02	50	nationalist	4.80	1.99	56
confused	2.94	1.36	47	natural	5.55	1.19	58
confusing	2.79	1.40	56	nice	6.16	1.10	44
converted	4.16	1.40	44	obedient	4.71	1.32	49
cool	5.73	1.02	44	official	4.91	1.21	55
corrupt	2.37	1.78	57	old	3.89	1.49	56
crazy	3.35	1.95	52	opposed	3.38	1.62	64
creepy	3.02	2.05	50	oppressed	3.04	1.47	48
cultural	5.26	1.20	57	pagan	3.61	1.54	46
curious	5.28	1.33	60	perfect	6.00	1.23	46
current	5.24	1.09	51	picky	4.05	1.31	56
cute	5.77	1.12	44	political	4.10	1.28	51
dark	3.45	1.82	55	powerful	5.55	1.08	53
dear	5.63	1.10	38	pregnant	5.11	1.39	37
dedicated	5.61	1.32	49	pretty	5.73	1.40	49
deep	5.11	1.03	47	private	4.50	1.11	44
defending	4.38	1.23	53	privileged	4.24	1.55	49
delicate	4.64	1.28	47	protective	5.59	1.21	44
delicious	5.73	1.21	44	proud	5.74	1.21	53
democratic	5.24	1.45	50	published	5.13	0.98	48
dependent	4.33	1.54	57	quiet	4.81	1.38	47
desperate	2.90	1.75	48	racist	2.74	1.94	54
determined	5.71	1.04	51	radical	3.98	1.86	47
different	4.32	1.38	50	raw	4.10	1.14	51
difficult	3.07	1.61	46	ready	5.75	1.14	52
dirty	2.92	1.65	51	real	5.78	1.20	50
disabled	3.28	1.71	54	relaxed	5.73	1.08	51
disadvantaged	3.03	1.79	37	religious	4.80	1.66	49
distant	3.75	1.58	40	respectful	6.08	1.23	48
diverse	4.74	1.45	62	responsible	5.67	1.30	42

Adjective	Mean Valence	SD Valence	N of Raters	Adjective	Mean Valence	SD Valence	N of Raters
divorced	3.38	1.37	45	right	5.39	1.24	46
drunk	2.80	1.56	50	romantic	5.64	1.05	50
early	5.08	0.97	50	rowdy	2.84	1.65	57
easy	5.06	1.45	48	rude	2.25	1.60	57
economic	4.91	1.29	53	ruling	4.25	1.50	44
emotional	4.73	1.37	51	rural	4.17	1.48	54
employed	5.94	1.03	53	sad	2.49	1.28	47
entitled	3.81	1.92	42	safe	5.54	1.22	56
equal	5.28	1.39	46	scared	3.33	1.79	51
ethnic	4.53	1.30	43	secular	4.72	1.09	54
excited	5.44	1.39	57	selfish	2.88	1.90	48
expressive	5.46	0.95	50	serious	4.22	1.40	46
fake	2.02	1.33	47	sexist	3.12	1.76	42
famous	5.61	1.14	46	sexual	4.58	1.24	40
fashionable	5.46	1.08	56	sexy	5.37	1.25	52
favorite	5.70	1.21	57	sick	2.78	1.55	45
fellow	4.79	1.32	56	skilled	6.00	1.32	49
feminine	4.51	1.42	49	smart	5.91	1.20	57
feminist	4.38	1.59	58	sorry	3.68	1.47	60
fine	5.67	0.95	45	specific	4.80	1.05	54
fluid	4.15	1.22	48	spicy	4.69	1.33	49
forever	5.47	1.15	59	stressed	3.34	1.72	44
free	5.63	1.09	52	strict	4.06	1.44	47
friendly	6.12	1.18	51	struggling	3.50	1.82	48
front	4.58	1.12	55	stubborn	3.14	1.51	49
funny	5.76	1.10	46	stupid	2.56	1.72	54
gentle	5.84	1.14	44	successful	6.02	1.19	47
giving	5.65	1.07	43	suitable	5.45	0.97	51
glutenfree	4.33	1.33	48	terrorist	2.08	1.77	48
graduate	5.59	1.36	51	thoughtful	5.80	1.01	41
great	5.92	1.17	49	tired	3.42	1.69	53
growing	5.40	1.38	47	tough	4.37	1.40	57
handsome	5.93	1.07	55	traditional	4.94	1.36	50
happy	6.31	1.07	51	tribal	4.21	1.46	48
hardworking	5.96	1.25	52	trustworthy	6.33	1.05	42
healthy	6.16	1.16	51	ugly	2.74	1.71	50

Adjective	Mean Valence	SD Valence	N of Raters	Adjective	Mean Valence	SD Valence	N of Raters
helpful	5.92	1.21	50	uncomfortable	3.09	1.65	54
holy	5.04	1.29	55	unfortunate	2.55	1.50	40
homeless	2.92	1.53	52	unique	5.38	1.14	50
honest	6.13	1.03	54	united	5.40	1.39	47
hurt	3.10	1.53	50	visiting	5.02	1.06	45
illegal	2.67	1.78	43	wanting	4.35	1.31	48
important	5.55	1.28	53	warm	5.43	1.25	54
inclusive	5.40	1.21	53	weak	2.87	1.49	53
independent	5.20	1.32	54	weird	3.40	1.63	53
innocent	5.16	1.48	55	whole	5.38	1.40	52
insulting	2.47	1.76	47	wise	5.29	1.68	48
interesting	5.91	1.15	43	young	5.31	1.21	49
international	4.98	1.19	50				

Appendix H: Descriptive Analyses of Adjective Choice for Chapter II

The main focus of Chapter II was to look at when participants believe individuals do (or do not) retain traits associated with their initial group membership. Those analyses are discussed and displayed in the chapter. Here, I provide additional descriptive results for Study 1 and Study 2, addressing the coverage of the list for each of the conditions as well as the five most-chosen traits per condition.

Study 1

Table 31. Study 1 - proportion of participants offering new adjectives to list.

Social Dimension	Condition	Proportion Adding Adjectives	Added Adjectives
Eating Habits	Vegan to Vegetarian	3.92%	animal loving, indecisive
	Vegetarian to Vegan	8.51%	empathetic, fanatical, kind, liberal
Gender	Man to Woman	0.00%	
	Woman to Man	0.00%	

Social Dimension	Condition	Proportion Adding Adjectives	Added Adjectives
Political Ideology	Democrat to Republican	13.33%	confused, concerned, intelligent, noncommitted, rude, stupid, untrustworthy, wishy-washy
	Republican to Democrat	10.64%	adaptive, coward, free-thinking, open-minded, phony, progressive, untrustworthy
Race	Black to White	0.00%	
	White to Black	10.53%	impossible, racial, strong
Religion	Jewish to Muslim	8.11%	converted, indecisive, naïve
	Muslim to Jewish	14.71%	open-minded, ambivalent, religious, strong, tolerant, trusting, unsure
Wealth Class	Poor to Rich	5.88%	able, capitalist, closed, content, logical
	Rich to Poor	8.16%	irresponsible, bohemian, unhappy, unlucky

Table 32. Study 1 - top adjectives selected for each condition.

Social Dimension	Condition	Adjective	Proportion Chosen
Eating Habits	Vegan to Vegetarian	healthy	56.86%
		thoughtful	39.22%
		natural	35.29%
		dedicated	31.37%
		emotional	25.49%
		gentle	25.49%
		respectful	25.49%
	Vegetarian to Vegan	healthy	48.94%
		thoughtful	42.55%
		dedicated	40.43%
		responsible	31.91%
		natural	29.79%
Gender	Man to Woman	feminine	50.00%
		expressive	46.15%
		unique	42.31%
		comfortable	38.46%
		adult	34.62%
		happy	34.62%
	Woman to Man	brave	45.45%
		interesting	36.36%
		unique	36.36%
		expressive	31.82%
		masculine	31.82%

Social Dimension	Condition	Adjective	Proportion Chosen
Political Ideology	Democrat to Republican	political	60.00%
		nationalist	35.56%
		traditional	35.56%
		responsible	26.67%
		adult	24.44%
		aggressive	24.44%
	Republican to Democrat	democratic	51.06%
		political	51.06%
		curious	36.17%
		thoughtful	34.04%
		adult	27.66%
		responsible	27.66%
Race	Black to White	interesting	27.27%
		motivated	27.27%
		selfish	27.27%
		adult	18.18%
		employed	18.18%
		expressive	18.18%
		hardworking	18.18%
		masculine	18.18%
		natural	18.18%
		stressed	18.18%
		uncomfortable	18.18%
		unique	18.18%
	White to Black	expressive	47.37%
		misunderstood	42.11%
		proud	42.11%
		cultural	36.84%
		unique	36.84%
Religion	Jewish to Muslim	religious	72.97%
		interesting	43.24%
		unique	40.54%
		cultural	32.43%
		holy	32.43%
	Muslim to Jewish	religious	58.82%
		curious	35.29%
		thoughtful	35.29%
		cultural	32.35%
		adult	26.47%
		interesting	26.47%
		proud	26.47%
Wealth Class	Poor to Rich	successful	74.51%

Social Dimension	Condition	Adjective	Proportion Chosen
		hardworking	56.86%
		motivated	52.94%
		skilled	49.02%
		powerful	41.18%
		responsible	41.18%
	Rich to Poor	struggling	67.35%
		unfortunate	61.22%
		stressed	42.86%
		uncomfortable	40.82%
		alone	34.69%
		scared	34.69%

Study 2

Table 33. Study 2 - proportion of participants offering new adjectives to list.

Social Dimension	Condition	Age Group	Proportion Adding Adjectives	Added Adjectives
Political Ideology	Democrat to Republican	Child	13.21%	republican, conservative, easily influenced, judgemental, misguided, sarcastic, scared, secretive, stubborn, unfriendly
		Teenager	10.00%	authoritarian, defiant, independent, self important, sexist, uneducated
		Young adult	11.67%	conservative, ignorant, naïve, non gender, republican, self-absorbed, stupid, unremarkable
		Adult	8.47%	conservative, open minded, stupid, wise, wishy washy
		Middle-aged person	7.89%	fortunate, ignorant, intelligent, moral, selfish
	Republican to Democrat	Child	6.78%	independent, determined, liberal, mercurial, unique
		Teenager	11.11%	caring, liberal, challenging, compassionate, independent, influenced by media, open, open-minded, rebellious, self-righteous, woke
		Young adult	8.06%	brainwashed, democrat, open-minded, smart, snowflake, untrustworthy, welcoming

Social Dimension	Condition	Age Group	Proportion Adding Adjectives	Added Adjectives
		Adult	9.52%	caring, empathetic, flexible, intellectual, intelligent, liberal, elite, middle-aged, normal, open-minded, sophisticated, wise
		Middle-aged person	22.45%	educated, creepy, depressed, desperate, drunk, empathetic, independent, open, open minded, peculiar, retired, unhealthy, wandering, wishy washy
Religion	Jewish to Muslim	Child	1.96%	unique
		Teenager	4.08%	intelligent, muslim
		Young adult	5.00%	gullible, independent
		Adult	6.00%	defensive, egotistical, inconsiderate, open, quarrelsome, steadfast
		Middle-aged person	6.25%	confused, educated, family-oriented
	Muslim to Jewish	Child	13.64%	kind, confused, discriminatory, poignant, questioning, resentful, untraditional
		Teenager	2.08%	fortunate, thorough
		Young adult	11.32%	ambivalent, brave, helpful, introspective, open, scrupulous, spiritual, undecided
		Adult	13.33%	flexible, adaptable, determined, dissatisfied, indecisive, lost, open minded, progressive, searching, seeker
		Middle-aged person	7.69%	confused, curious, deep, fluid, old, uncertain

Table 34. Study 2 - top adjectives selected for each condition.

Condition	Age Group	Adjective	Proportion Chosen
Democrat to Republican	Child	traditional	56.60%
		political	54.72%
		proud	32.08%
		responsible	30.19%
		adult	28.30%
		hardworking	28.30%

Condition	Age Group	Adjective	Proportion Chosen
	Teenager	political	64.00%
		traditional	46.00%
		privileged	34.00%
		employed	28.00%
		motivated	26.00%
		nationalist	26.00%
	Young adult	political	61.67%
		adult	53.33%
		traditional	45.00%
		privileged	30.00%
		nationalist	25.00%
		responsible	25.00%
	Adult	political	79.66%
		traditional	55.93%
		adult	38.98%
		responsible	33.90%
		employed	30.51%
	Middle-aged person	political	65.79%
		traditional	36.84%
		adult	35.53%
		privileged	32.89%
		masculine	28.95%
		successful	28.95%
Republican to Democrat	Child	democratic	66.10%
		political	55.93%
		thoughtful	35.59%
		friendly	25.42%
		responsible	25.42%
	Teenager	democratic	71.43%
		political	61.90%
		thoughtful	49.21%
		smart	46.03%
		curious	42.86%
	Young adult	political	72.58%
		democratic	70.97%
		adult	43.55%
		thoughtful	40.32%
		smart	30.65%
	Adult	democratic	74.60%
		political	69.84%
		adult	50.79%
		thoughtful	47.62%

Condition	Age Group	Adjective	Proportion Chosen
		responsible	39.68%
	Middle-aged person	democratic	63.27%
		political	57.14%
		adult	55.10%
		thoughtful	44.90%
		interesting	32.65%
		responsible	32.65%
Jewish to Muslim	Child	religious	74.51%
		cultural	35.29%
		holy	33.33%
		curious	31.37%
		misunderstood	27.45%
	Teenager	religious	83.67%
		dedicated	36.73%
		thoughtful	36.73%
		adult	34.69%
		cultural	34.69%
		interesting	34.69%
	Young adult	religious	70.00%
		cultural	50.00%
		adult	35.00%
		holy	30.00%
		curious	27.50%
		interesting	27.50%
		traditional	27.50%
	Adult	religious	78.00%
		cultural	56.00%
		curious	38.00%
		holy	36.00%
		traditional	36.00%
	Middle-aged person	religious	85.42%
		adult	47.92%
		cultural	47.92%
		holy	37.50%
		interesting	35.42%
		thoughtful	35.42%
Muslim to Jewish	Child	religious	81.82%
		thoughtful	52.27%
		interesting	43.18%
		brave	40.91%
		unique	40.91%
	Teenager	religious	79.17%

Condition	Age Group	Adjective	Proportion Chosen
		cultural	58.33%
		holy	41.67%
		curious	35.42%
		thoughtful	35.42%
	Young adult	religious	84.91%
		cultural	49.06%
		adult	41.51%
		interesting	33.96%
		dedicated	32.08%
		holy	32.08%
	Adult	religious	80.00%
		curious	55.56%
		interesting	42.22%
		adult	40.00%
		brave	40.00%
		cultural	40.00%
	Middle-aged person	religious	71.79%
		cultural	48.72%
		interesting	35.90%
		unique	35.90%
		adult	33.33%
		curious	33.33%