Bayesian Reasoning and Demographic Misperceptions

Brian Guay*[1], Tyler Marghetis[2], Cara Wong[3], and David Landy[4]

[1]Ph.D. Candidate, Dept. of Political Science, Duke University
[2]Assistant Professor, Dept. of Cognitive and Information Sciences, University of California, Merced & Omidyar Fellow, Santa Fe Institute
[3]Associate Professor, Dept. of Political Science, University of Illinois, Urbana-Champaign
[4]Senior Data Scientist, Netflix

October 1, 2020

Abstract

Misperceptions about the size of demographic groups are among the most cited instances of political misinformation, yet little is understood about their origins. Existing theories emphasize the role of perceived threat and social contact, yet these accounts are limited by inconsistent empirical support. In this paper we argue that demographic misperceptions are one instance of a larger pattern of Bayesian proportion rescaling, whereby individuals systematically bias their estimates of proportions toward a prior belief, regardless of what the proportions represent. We find strong support for our theory across over 35,000 estimates and 42 estimated groups from existing studies and original data. We then evaluate our theory alongside current explanations using a rich dataset containing both national and local estimates of multiple racial groups and measures of perceived threat and social contact. Our findings have implications for how researchers interpret misperceptions about politically-relevant quantities.

Word count: 9,947

*Correspondence should be directed to Brian Guay (brian.guay@duke.edu)

A central question in political science concerns the extent to which citizens in democratic societies are informed about politics. Decades of research suggests that the public is characterized by relatively low levels of political knowledge (Converse, 1964; Carpini and Keeter, 1996) and a growing body of work documents the many misperceptions citizens hold across a wide range of policy domains (Flynn et al., 2017). Among the most cited instances of citizen ignorance are demographic misperceptions—inaccurate beliefs about the size of groups in the population. For instance, Americans dramatically overestimate the share of the population that is African American, Latino, Muslim, Jewish, or gay (e.g., Alba et al., 2005; Wong, 2007; Martinez et al., 2008) and people around the world overestimate the size of the immigrant population (Sides and Citrin, 2007; Herda, 2010; Hopkins et al., 2019).

These misperceptions are consequential not only because numeric facts "prevent debates from becoming disconnected from the material conditions they attempt to address" (Carpini and Keeter, 1996, pg. 11), but also because they are associated with both attitudes toward the groups and support for policies that affect them (Kuklinski et al., 2000; Sides and Citrin, 2007). When perceptions of group size serve as cognitive shortcuts in political decision-making, *mis*perceptions can lead to biased attitudes and behavior. For instance, past work shows that people who overestimate the size of the immigrant population are more likely to support restrictive immigration policies (Sides and Citrin, 2007). More recent work suggests that misperceptions about the size of stereotypical partisan groups (e.g., the share of Democrats who are gay) contribute to rising levels of political polarization in the U.S. (Ahler and Sood, 2018).

A critical question thus concerns the origins of demographic misperceptions. One theory, rooted in Realistic Conflict Theory (Key, 1966; Blalock, 1967; Blumer, 1958) posits that people overestimate the size of minority groups that they perceive as threatening (Allport, 1954; Nadeau et al., 1993; Semyonov et al., 2004; Dixon, 2006). Another posits that contact with members of a minority group—either in-person or indirectly through media exposure—influences perceptions of that group's size, with greater levels of exposure driving larger estimates of a group's size (Nadeau et al., 1993; Sigelman and Niemi, 2001; Herda, 2010). However, empirical support for these theories is limited, with even the most comprehensive models accounting for little variation in people's systematic overestimation of the size of minority groups (Alba et al., 2005; Herda, 2010; Nadeau et al., 1993). More importantly, while these theories provide a theoretical account for why members of the majority overestimate the size of minority groups, they do not explain why members of minority groups make nearly identical errors (Wong, 2007; Duffy, 2018). Nor do these theories explain why people make similar errors when estimating the non-racial groups, such as the share of the population that is unemployed, in poverty, or donates to charity (Lawrence and Sides, 2014; Theiss-Morse, 2003)

In this paper we provide an alternative explanation for demographic misperceptions, rooted in the psychology of individual decision-making under uncertainty. We propose that demographic misperceptions are less about attitudes towards the specific group being estimated and more about the systematic cognitive errors people make when estimating the size of proportions. When people are asked to estimate the proportion of the population that belongs to a certain group, they engage in a form of Bayesian reasoning: They incorporate both information about the size of that group *and* prior beliefs about the size of groups more generally. This process, which we refer to as *Bayesian rescaling* results in overestimating of the size of smaller groups and underestimating of the size of larger ones, the same pattern that political scientists have observed in demographic misperceptions over the last two decades. Unlike other theories of demographic misperceptions, Bayesian rescaling explains a wider range of demographic misperceptions—not only why members of the majority overestimate the size of minority groups, but also why members of minority groups overestimate their own prevalence, and why members of both minority and majority groups underestimate the size of majority groups. While Bayesian rescaling has been widely documented in other instances of proportion estimation, such as economic decision-making (e.g., Tversky and Kahneman, 1992), estimates of general numerical magnitudes (e.g., Barth and Paladino, 2011; Landy et al., 2018; Cohen and Blanc-Goldhammer, 2011), and estimates of proportion of shapes and sounds with specific characteristics (e.g., Erlick, 1964; Varey et al., 1990; Nakajima, 1987), it has been largely overlooked by research on political misperceptions until now.

We formalize a model of Bayesian rescaling and test it using a large collection of estimates of the size of demographic groups from the American National Election Study, European Social Survey, General Social Survey, and six published studies. Since past studies on misperceptions focus almost exclusively on the size of relatively small racial and ethnic groups, we also conducted two original surveys in which we asked respondents in the Cooperative Congressional Election Study and an online survey conducted on Lucid to estimate the prevalence of a wider variety of groups, including groups for which misperception is difficult to explain using theories of perceived threat and social contact. Together, these data constitute the largest and most diverse collection of demographic estimates analyzed to date, containing 42 unique demographic groups from over 35,000 respondents in multiple countries over several decades. We then evaluate Bayesian rescaling alongside existing theories of demographic misperception using the 2000 General Social Survey, which contains estimates of the proportion of the population that is Black, Hispanic, Asian American, and White, as well as measures of how much contact respondents have with each group and how threatening they perceive each group to be.

We find strong evidence that demographic misperceptions are largely the result of Bayesian rescaling. Across multiple data sets, our model of Bayesian rescaling closely predicts the general pattern of misestimation documented by political scientists for decades. Moreover, and as predicted by our model, people made nearly identical errors when estimating the proportion of the population that is Black, Hispanic, and White as they did when estimating the proportion of the population that owns a dishwasher, holds a valid passport, and has indoor plumbing. Indeed, the pattern of errors in our data bears a striking resemblance to those that have been observed in estimates of non-demographic quantities across multiple domains (e.g., economic decision-making; Tversky and Kahneman (1992)). Moreover, we show that accounting for Bayesian rescaling consistently and substantially increases the amount of variance explained in estimates of racial in-groups and out-groups at the national and local level, while perceived threat and contact account for little variance. Indeed, we find little evidence that perceived threat and contact are associated with these misperceptions at all, lending further support to our hypothesis that demographic misperceptions are largely the result of domain-general process rather than characteristics of the specific groups being estimated. Taken together, these findings have implications not only for our understanding of where demographic misperceptions originate, but also for how they should be interpreted.

Theories of Demographic Misperception

A growing body of research documents the misperceptions people hold about the size of politically-relevant demographic groups. Across Europe and the U.S., people dramatically overestimate the size of the immigrant population (Hopkins et al., 2019; Sides and Citrin, 2007; Gorodzeisky and Semyonov, 2018). Americans overestimate the size of racial and ethnic minority groups—such as the proportion of the population that is Black, Hispanic, Asian, and Jewish— and underestimate the size of majority groups, such as Whites and Christians (Nadeau et al., 1993; Alba et al., 2005; Lawrence and Sides, 2014; Theiss-Morse, 2003; Sigelman and Niemi, 2001; Wong, 2007; Gallup and Newport, 1990). Similarly, people overestimate the share of the population that is college-educated, unemployed, lives under the poverty line, and receives welfare, as well as the share of welfare recipients who are Black, uneducated, and rely on welfare for more than 8 years (Lawrence and Sides, 2014; Kuklinski et al., 2000). Such misperceptions are frequently interpreted as political ignorance or innumeracy both by academics and the media, which often reports survey findings with headlines like "Today's Key Fact: You are Probably Wrong About Almost Everything" (The Guardian, 2014), "Americans Drastically Overestimate How Many Unauthorized Immigrants Are in The Country, And They Don't Want to Know the Truth" (Slate, 2012), "Here's how little Americans really know about immigration" (The Washington Post, 2016).[1]

The wide range of groups that are misperceived in society raises normative concerns about the ability of citizens to form political attitudes that are tethered to reality. Even when Americans are ideologically unconstrained, they often base their policy preferences on the groups that are affected by policies (Converse, 1964). Sides (2013, pg. 2) explains that "group-centric reasoning allows citizens to make political decisions without much detailed information or more sophisticated abstract reasoning," similar in rationale to Dawson's (1995) Black utility heuristic. If voters think in terms of racial and ethnic groups as they cast their ballots, misperceptions about groups can bias what might otherwise be useful cognitive shortcuts in political decision-making. Research examining the relationship between misperceptions and attitudes lends credence to these concerns. For instance, people who overestimate the size of the immigrant population are more opposed to immigration and hold more negative views of immigrants (Sides and Citrin, 2007; Herda, 2010). Similarly, Gilens (1999) finds that overestimating the percentage of poor people who are black leads to greater opposition to welfare programs. Likewise, Ahler and Sood (2018) find that misperceptions about the composition of political parties in the U.S., such as the proportions of Democrats who are gay and Republicans who are wealthy, fuel negative partisan affect and allegiance to one's own party.

To date, two theories explaining the origins of demographic misperceptions have emerged. The first posits that individuals overestimate the size of groups that they perceive as threatening. This explanation is rooted in one of the core tenets of Realistic Group Conflict Theory—that members of the majority group perceive minority groups as more threatening as the size of the minority group increases (Bobo, 1999; Key, 1966). As minority groups grow in size, majority group members fear competition over scarce economic and political

[1]https://slate.com/business/2012/01/americans-drastically-overestimate-how-many-unauthorized-immigrants-are-in-the-country-and-they-don-t-want-to-know-the-truth.html
https://www.theguardian.com/news/datablog/2014/oct/29/todays-key-fact-you-are-probably-wrong-about-almost-everything
https://www.washingtonpost.com/news/wonk/wp/2016/09/01/heres-how-little-americans-really-know-about-immigration/

resources, which leads to greater prejudice and discrimination against the minority group members (Blalock, 1967; Bonilla-Silva, 2001; Dixon, 2006; Sides and Citrin, 2007). Multiple studies have documented higher levels of perceived threat and greater prevalence of anti-minority attitudes in regions with higher concentrations of racial and ethnic minorities (Fossett and Kiecolt, 1989; Quillian, 1995). This relationship has been leveraged to explain variation in the *perceived* size of minority groups. Allport (1954) alludes to this when describing South Africans' perceptions of the size of the Jewish population as 20% (vs. 1%), suggesting that "quite likely fear of a Jewish 'menace' underlay the inflated estimate" (pg. 166). More recent studies have similarly suggested that demographic misperceptions are influenced by perceptions of threat, arguing that Americans overestimate the size of Black, Hispanic, and Jewish populations when these groups are seen as threatening (Nadeau et al., 1993; Alba et al., 2005). Gallagher (2003) concludes that "the media, residential segregation, racial stereotypes, and perception of group threat each contribute to Whites' underestimation of the size of the White population and the inflation of group size among racial minorities" (pg. 381).

A second theory posits that perceptions of group size are influenced by an individual's exposure to members of that group (e.g., Lee et al., 2019). People construct beliefs about the world based on experiences and observations made in the course of daily life, including those with whom they interact (Howard et al., 2003). Accordingly, these experiences and observations should influence perceptions of the size of demographic groups. Nadeau et al. (1993), for example, find greater overestimation of minority groups by individuals who report more frequent interactions with them. Similarly, Sigelman and Niemi (2001) find that "for both African Americans and Whites, individuals who interacted more with African Americans were more likely to overestimate the size of the Black population" (pg. 93). Some have also suggested that less intimate forms of exposure to groups, such as through the media, can similarly increase overestimation, though empirical support is limited (Herda, 2010).

While theories of perceived threat and social contact offer an intuitive explanation of some demographic misperceptions, they are constrained by the narrow subset of observations they explain. Past work has almost exclusively sought to explain Whites' estimates of racial and ethnic minority groups (e.g., Nadeau et al., 1993; Alba et al., 2005; Herda, 2010; Sigelman and Niemi, 2001). It is unclear how these theories account for the misperceptions held by people belonging to minority groups, or the misperceptions members of majority groups people hold about their own group. For instance, theories of social contact predict that members of majority groups should overestimate the size of their own group, since people tend to socialize with people who are similar to themselves (Lee et al., 2019). Theories of perceived threat predict that minorities, too, should overestimate the size of majority populations that they perceive as threatening, while underestimating the size of minority populations they perceive as non-threatening. However, the evidence demonstrates the opposite—members of both minority and majority groups similarly overestimate the size of majority groups and underestimate the size of minority groups (e.g., Wong, 2007; Duffy, 2018).

Furthermore, there is limited empirical support for both existing theories. For example, prior studies show that models accounting for perceived threat and contact explain a relatively small proportion of variance in the misperceptions White respondents hold about the size of minority groups (Nadeau et al., 1993; Herda, 2010; Alba et al., 2005). There is also inconsistent empirical support for theories of social contact. For instance, Herda (2010) measures exposure to immigrants five ways and finds that only two of them are associated with overestimating the immigrant population, while one is associated with *underestimating* the size of the immigrant population. Moreover, the errors people make when estimating the size of racial groups are almost identical to those made when estimating quantities that cannot be explained by perceived threat and contact. For instance people make similar errors when estimating racial and non-racial demographic groups, such as the share of the population that donates to charity (Theiss-Morse, 2003) or receives welfare (Kuklinski et al., 2000). The same is true for estimates of entirely non-demographic quantities, such as the share of the federal budget spent on foreign aid (Gilens, 2001) and the inflation rate (Conover et al., 1986).

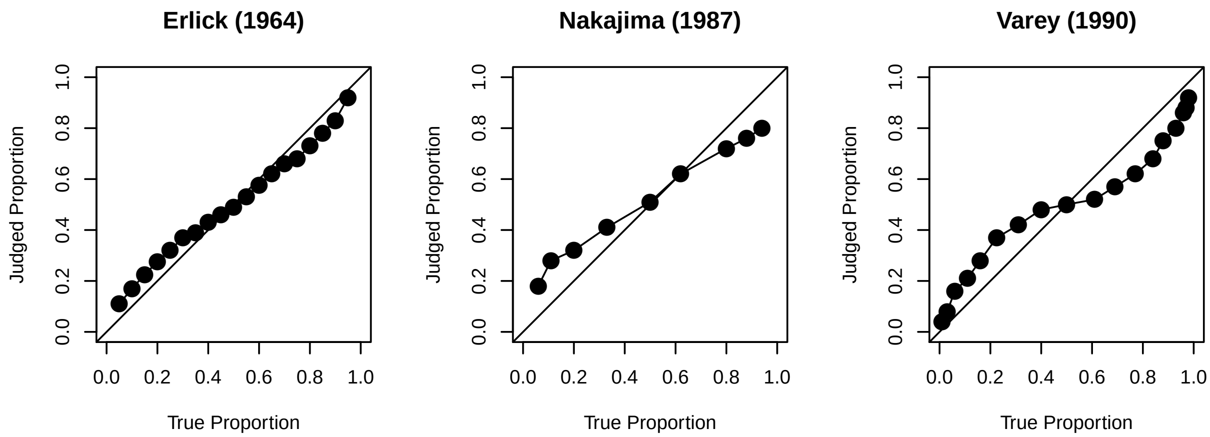
In the next section, we propose a more general explanation of demographic misperceptions, one that explains the errors people make when estimating the size of demographic populations regardless of the group being estimated or the person making the estimate. Whereas the focus of prior work on the origins of demographic misperceptions have been rooted in people's perceptions of fear or contact with a particular group being estimated, we focus instead on the general cognitive errors individuals make when estimating the size of proportions.

The Psychology of Proportion Estimation

A common assumption holds that people's responses to surveys do not correspond perfectly with their underlying beliefs and attitudes. Survey respondents "must sample from a set of available considerations in order to construct an answer to the question" (Flynn et al., 2017, pg. 138), and this construction process can often introduce error. In the specific case of survey questions that require respondents to report beliefs about specific quantities, such as the size of demographic groups, Kuklinski et al. (2000) note that "we do not expect [individuals] to infer details such as specific amounts and percentages in the ordinary course of events. Instead, they will construct and store more general factual beliefs [...]. When they have the occasion—for example, answering a survey—they will translate these general notions into more specific ones" (p. 795).

The central claim of this paper is that the translation from these "general notions" about the size of demographic groups to responses on surveys is characterized by the same types of systematic error that occur when people estimate proportions more generally. Several decades of research on how people estimate and interact with quantitative information has found that translations from "general notions" to explicit estimates of proportions are systematically skewed—individuals overestimate the size of small proportions and under-

Figure 1: Examples of Proportion Estimation Error from Prior Studies



Mean proportion estimates from prior studies. From left to right, estimates of the proportion of letters in a sentence that are 'A', time intervals containing a specific sound, and dots that are a certain color. For a comprehensive overview, see Hollands and Dyre (2000).

estimate large ones (Stevens, 1957; Gonzalez and Wu, 1999). Moreover, these estimates consistently follow an inverted S-shaped pattern, with the most dramatic over-under estimation occurring near the ends of the proportion space, close to .20 and .80.

Moreover, the systematic overestimation of small proportions and underestimation of large proportions appears to be domain-general, or unrelated to the specific quantity that the estimated proportion represents. Researchers examining quantitative judgments have observed the same pattern of over-under estimation across a wide variety of domains. People consistently overestimate small proportions and underestimate large ones when estimating the proportion of 'A's in a random sequence of letters (Erlick, 1964), the number of dots on a page that are a specific color (Varey et al., 1990), and the proportion of time intervals containing a specific sound (Nakajima, 1987). Figure 1 illustrates the pattern of over-under estimation from these early studies on proportion estimation. Similar forms of misestimation error characterize economic decision-making (Tversky and Kahneman, 1992), estimates of general numerical magnitudes (Barth and Paladino, 2011; Cohen and Blanc-Goldhammer, 2011), and interpretations of bar graphs and pie charts (Spence, 1990).

Bayesian Rescaling

Why do people overestimate the size of small proportions and underestimate the size of large proportions across such a diverse set of domains? Psychologists have produced different theories over the decades; here we present one that captures features shared by several domain-general theories. Our model specifies that the specific pattern of systematic overestimation of small proportions and underestimation of large ones follows from two generic properties of human reasoning about numeric quantities: 1) rescaling new information toward a prior belief and 2) processing proportions as log-odds. In what follows, we briefly review each of these properties of quantitative reasoning, provide illustrative examples, and formalize these processes in a model of generalized proportion estimation error, which we term *Bayesian rescaling.*

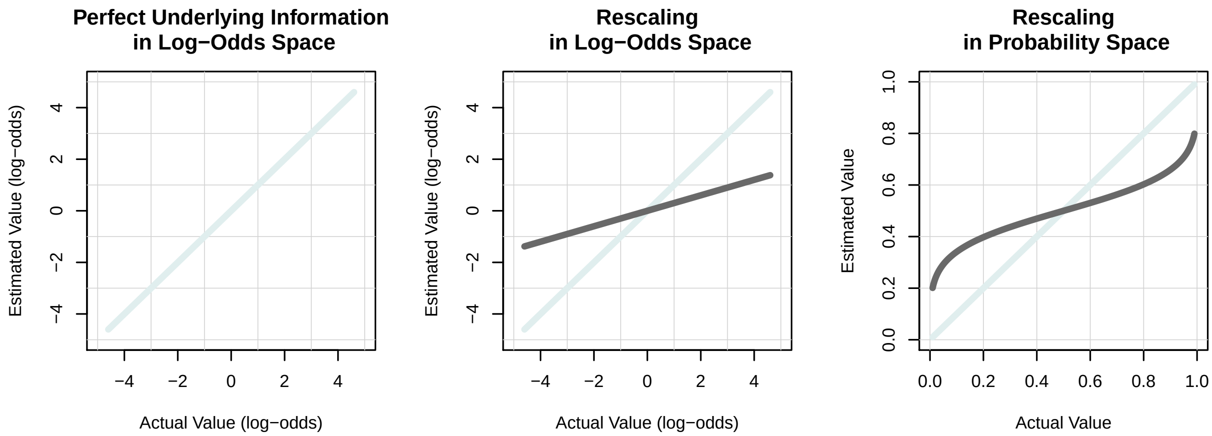
The first property of quantitative reasoning that produces generalized proportion estimation error is that when estimating a proportion, individuals rely not only on information specific to that proportion (e.g., the number of immigrants in a country), but also *prior information about the size of proportions more generally.* Survey researchers have long implicitly made the assumption that respondents incorporate prior information about the range of possible values into their estimates. Indeed, if people did not incorporate *any* prior information about proportions, they might completely ignore the fact that proportions are bounded by 0 and 1 and estimate that 120% of the population is foreign-born. However, the Bayesian approach goes beyond this by assuming that people sometimes take into account not just the boundaries, but the distribution of typical proportions more generally.

When individuals are uncertain about the true size of a specific proportion, such as the proportion of the population that is foreign-born, a rational strategy is to not only rely on information implicitly gathered from one's exposure to immigrants in daily life, but also knowledge of proportions more generally. Indeed, if an individual has *no information* about the size of the immigrant population, and so regards each proportion as equally likely (the uniform prior), the estimate that minimizes response error lies directly in the middle of the proportion space, .50. A consequence of this reliance on prior information about the size of proportions more generally is that as individuals are increasingly uncertain about the information they are estimating, they will increasingly move, or *hedge*, their estimates toward the center of the distribution of their prior. While this behavior has been referred to by a wide variety of names (e.g., regularization, evidence-pooling, rescaling), we refer to it as *Bayesian rescaling*. 'Bayesian' because individuals are incorporating their prior beliefs into their explicit estimate of a proportion, and 'rescaling' because they are doing so by shifting or rescaling their estimates toward those priors. We illustrate this process of Bayesian rescaling in the first and second panels of Figure 2.

To demonstrate, consider the case of an airline passenger who has a layover in a foreign country and immediately upon landing is asked by a pollster to estimate the proportion of the population that is foreign-born in that country. If this individual knows nothing about the size of the immigrant population and has no experiences in that country with which to inform an estimate, she will likely rely heavily on her prior beliefs about the size of the immigrant populations in other countries she has visited, or perhaps on the range of all possible responses more generally (i.e., 0-100%). Indeed, with no information at all, the best guess is the center of one's prior distribution. Conversely, if this person has perfect underlying information about the specific quantity at hand—for instance, if they have visited the country for decades and read extensively about its immigrant population—she would likely not need to rely on this prior information at all. In reality, most people will fall somewhere in-between these two extremes, hedging their estimates of proportions smaller than the center of their prior upwards, and hedging their estimates of proportions larger than that downward.

Importantly, the prior belief is not always .50 (Schille-Hudson and Landy, 2020). For instance, when estimating the size of a group one knows to be a minority, the range of possible estimates is constrained above by .50, because a minority group cannot, by definition, account for more than 50% of the population. With no information about a group other than that it is a minority, a reasonable prior will be closer to .25. Likewise, because the size of majority

Figure 2: Bayesian Rescaling



The process of domain-general proportion estimation error, or Bayesian rescaling. The three panels correspond with Equations 1, 2, and 3 below. The first panel illustrates the individual's accurate perception of the size of proportions, expressed as log-odds. In the second panel, perceptions are rescaled toward a prior of 50%–that is, small proportions are rescaled upwards and large proportions are rescaled downwards. The third panel illustrates these rescaled perceptions translated in proportion space.

groups is naturally greater than .5, plausible priors will be constrained to values between .5 and 1.

The process of domain-general proportion estimation error, or Bayesian rescaling, illustrated for a hypothetical individual with perfect underlying information about the proportions being estimated. The first panel illustrates the individual's accurate perception of the size of proportions in log-odds space. In the second panel, perceptions of small proportions are rescaled upwards and perceptions of large proportions are rescaled downwards, resulting in a linear but inaccurate perceptions. The third panel illustrates these same misperceptions, but in proportion space, which is how respondents are asked to estimate the size of demographic groups on surveys.

The second property of quantitative reasoning that produces systematic errors in proportion estimation is that *proportions are mentally processed as log-odds*. Any model of quantitative estimation must specify the format of individuals' internal representations. There are many natural ways to represent proportional information, for instance as percentages, proportions, fractions, odds, or log-odds. Although log-odds are not as familiar to non-statisticians, there are many reasons to favor them as a baseline model of human representation of implicit numerical values. First, they naturally align with the logarithmic representation of other magnitudes, such as weight, loudness, numerosity, and many others (e.g., Gonzalez and Wu, 1999). Second, log-odds are unbounded, making Bayesian inference in terms of normal distributions feasible. Third, they produce S-shaped curves extremely similar to those empirically found in a large range of cases. For these reasons, in line with recent work (Landy et al., 2018; Marghetis et al., 2018; Gonzalez and Wu, 1999), our model of proportion estimation assumes that individuals' mental representations are encoded as log-odds. To be clear, we are not claiming that people are *aware* of the format of their internal representations of proportions. Rather, we are suggesting that people implicitly store these values this way, and that log-odds characterize the influence of Bayesian inference on that process.

Next, we formalize both components of proportional reasoning: that mental representations of proportions are in the form of log-odds, and that people engage in rescaling under uncertainty (see also Landy et al. (2018)). First, mental representations of proportions (r_p) are processed as the proportion (p) in log-odds:

$$r_p = \log\left(\frac{p}{1-p}\right) \tag{1}$$

When survey respondents are asked to estimate the size of a group, their estimates are a linear combination of the specific information they have about the size of a group and their prior belief about the sizes of groups in general, with those sizes represented as log odds:

$$\Psi'(r_p) = \gamma r_p + (1-\gamma)\log(\delta_{odds}) \tag{2}$$

Equation 2 formalizes the model of Bayesian rescaling introduced above. The first term represents how people's estimates rely on their own information about the size of the group in log odds space (r_p), with the degree to which they rely on this information is represented by the weighting parameter (γ). The second term represents how people's estimates rely on prior information about the size of groups in general ($\log(\delta_{odds})$), weighted by $(1-\gamma)$. These two terms are combined to produce a person's estimate of the group's size, represented in log odds ($\Psi'(r_p)$).

Of course, respondents express their knowledge on surveys as proportions, not log-odds. Equation 3 therefore translates Equation 2 into the proportion space:

$$\Psi(p) = \frac{\delta_{odds}^{(1-\gamma)} p^{\gamma}}{\delta_{odds}^{(1-\gamma)} p^{\gamma} + (1-p)^{\gamma}} \tag{3}$$

In what follows, we test our model of Bayesian rescaling on the largest and most diverse collection of demographic misperceptions to date. We then evaluate our theory alongside existing theories of perceived threat and social contact using a rich dataset containing both estimates of the national and local prevalence of racial minority groups in the U.S.

Data & Methods

Mapping Demographic Misperceptions

We begin by applying our model of Bayesian rescaling to a dataset containing demographic estimates from three large government-funded surveys, six published studies, and two original surveys. We use these data to examine the overarching pattern of estimation errors that people make when evaluating the size of demographic groups. While past work on demographic misperceptions considers estimates of specific demographic groups in isolation, examining a wide range of estimates enables us to test for the existence of a broader pattern of systemic over-under estimation found in other domains of proportion estimation. This large collection of estimates also enables a comparison between estimation errors made by respondents to those predicted by the Bayesian rescaling model presented in Equation 4.

First, we obtain estimates included on large high-quality public surveys frequently used in studies examining demographic misperceptions: the 1991 American National Election Study Pilot (ANES), 2000 General Social Survey (GSS), and 2002 European Social Survey (ESS). Together, these data contain 40,576 individual estimates of 10 demographic groups from 33,508 respondents in 21 countries during a period of 11 years. We also include estimates from six existing studies that use original survey data to measure demographic misperceptions (Ahler and Sood, 2018; Hopkins et al., 2019; Lawrence and Sides, 2014; Theiss-Morse, 2003; Gallup and Newport, 1990).[2]

These data are limited in two respects. First, of the 22 unique groups asked about, only 3 have a true size of more than 50%. This makes it difficult to observe a broader pattern of over-under estimation if it exists and may help to explain why past work on demographic misperceptions has not addressed the systematic over-under pattern of misestimation observed in other domains. Second, these surveys primarily contain estimates of racial and ethnic groups. While theories of perceived threat and contact were developed to explain the widespread overestimation of racial groups, demographic misperceptions are clearly not limited to only these groups (e.g., Ahler and Sood, 2018; Lawrence and Sides, 2014; Kuklinski et al., 2000). Similar patterns of error in estimates of racial and non-racial groups might suggest an underlying cause other than perceived threat or social contact.

We therefore conducted two original surveys to obtain estimates of a more diverse range of demographic groups. First, we asked 1,000 respondents on the 2016 Cooperative Congressional Election Study (CCES) to estimate the size of 10 demographic groups, including adults in the U.S. who are White (.77), Republican (.44), Democrat (.48), and own a home (.63). Additionally, we asked respondents from an online non-probability sample of 1,220 U.S. adults to estimate the size of 19 non-racial groups that cannot be easily explained by

[2]We included estimates from studies that reported group mean or median estimates and true values, or had publicly available replication data with which these values could be calculated, and had a sample size of more than 200 respondents.

existing theories of demographic misperception, such as the proportion of U.S. adults who are younger than 95, clinically obese, earn less than $30,000 annually, and who possess common objects, such as a cell phone, microwave, stove, washing machine, clothes dryer, dishwasher, car, driver's license, and passport.[3]

Comparison to Existing Theories

Next, we compare Bayesian rescaling with existing accounts of perceived threat and contact by modeling the errors people make when estimating the size of racial groups. To do so, we use the 2000 General Social Survey (GSS)[4], which includes estimates of the share of the U.S. population that is Black, Hispanic, Asian, and White, as well as individual-level measures of perceived threat and contact for these groups. We restrict our analysis to the 1,398 respondents who were randomly selected to receive the *Multi-Ethnic United States* module, which contains measures of the perceived size of racial and ethnic groups in the U.S. and attitudes towards these groups.

While past work primarily considers national level demographics, the GSS separately asked respondents to estimate the prevalence of each of these groups at both the national and local level. Local estimates are important for two reasons. First, while each demographic group has only one true size at the national level (e.g., 12% of the U.S. population is Black), group sizes vary widely at the local level in the U.S. For instance, the local population in our sample ranged from less than 1% to 57%). This variation enables us to determine how rescaling, threat, and contact vary within estimates of a single racial group. Second, modeling local estimates enables a more conservative test of Bayesian rescaling alongside theories of perceived threat and contact. Since the latter posit that misperceptions about the size of groups are largely driven by everyday interactions with individuals through personal observation, we might expect these factors to be even *more* influential in estimates of the

[3]We recruited an online non-probability sample using Lucid, a a survey sampling firm that connects researchers to a large pool of online research participants (see Coppock and McClellan (2019) for an overview).

[4]The 2000 GSS was conducted in-person from February to May 2000 on a probability sample of 2,817 U.S. adults.

local community than in the nation as a whole.

Another unique characteristics of the GSS data is that respondents estimated both the size of groups to which they *do and do not belong* (i.e., in-groups and out-groups, respectively). Due to the focus of prior studies on estimates of out-groups, we lack an understanding of why individuals make similar errors when estimating the size of in-groups as they do when estimating the size of out-groups. These data allow us to compare theories of Bayesian rescaling, perceived threat, and social contact for estimates of in-groups and out-groups separately.

The GSS includes two items measuring respondents' contact with members of out-groups. Respondents were asked "do you know any [Whites/Blacks/Hispanics/Asians]," and, if they indicated that they did, were asked "are any of these [Whites/Blacks/Hispanics/Asians] people you feel close to?" We constructed an index using these two items: respondents who reported not knowing anyone from a group were assigned a value of 0 (46% of the sample), respondents who reported knowing but not feeling close to anyone from a group were assigned a value of .5 (29%), and respondents who reported knowing and feeling close to someone from a group were assigned a value of 1 (25%).

Perceived threat has been operationalized in a number of ways by past work, often by asking respondents directly about whether they believe there is a zero-sum inter-group competition for political, economic, or cultural influence. However, competition is not a necessary condition for threat to manifest in prejudice and discrimination (Wilcox and Roof, 1978). As Blumer (1958) explains, a perceived challenge to the status quo (via out-group population concentration) can lead dominant groups to seek to maintain their social distance from other groups (and even increase the salience of racial boundaries) and their relatively privileged position. To measure perceived threat we construct an index of eight items measuring attitudes toward each of the four racial groups (Cronbach's α = .76). Respondents were asked to what extent they perceived members of each group as violent (vs. peaceful), unintelligent (vs. intelligent), lazy (vs. hardworking), and committed to strong families and the equal treatment of all members of society (vs. not committed). Additionally, respondents were asked how comfortable they would be marrying and living in a neighborhood where half of their neighbors were a member of each group and to rate how important the contributions each group makes are to the country as a whole.

While these items enable us to measure perceived threat identically for each of the four estimated groups and captures the negative group affect, prejudice, and discrimination Blalock (1967) theorized are intertwined with perceptions of threat, they do not directly capture the competition dimension of perceived threat. Since it is possible that this dimension of threat is the principal driver of misestimation error, we constructed a second measure of perceived threat that closely matches the extant literature on the relationship between demographic misperceptions and perceived threat, but is available for only two of the groups being estimated. We follow Alba et al.'s (2005) operationalization of perceived threat using survey items asking White respondents about African Americans and Hispanics specifically, including questions measuring racial resentment, threat posed by Hispanic immigrants.[5]

Results

We begin by considering the large collection of racial and non-racial demographic estimates from existing work and original studies. In Figure 3, mean estimates are plotted against the true values of the proportions being estimated, along with predictions from the model specified in Equation 4. A pattern of over-under estimation is immediately apparent when considering demographic misperceptions in the aggregate. This pattern is even more recognizable after accounting for the wider range of population sizes in our original data (panel 2). On average, respondents underestimate the size of majority groups and overestimate the size of minority groups. In fact, all of the 68 minority groups (i.e., less than 50% of the

[5]Details on Alba et al.'s (2005) measures are included in the Appendix (page 8). Since two of the Black perceived threat items were featured on a portion of the survey using a split-ballot design, and therefore only asked of a random 50% sample of respondents, this portion of the analysis is limited to 503 of the 1,088 White respondents in the GSS when using this measure of perceived threat. For Hispanic estimates, we are able to use 769 of the 1,088 White respondents.

population) are overestimated, while 20 of the 21 majority groups (i.e., less than 50% of the population) are underestimated.

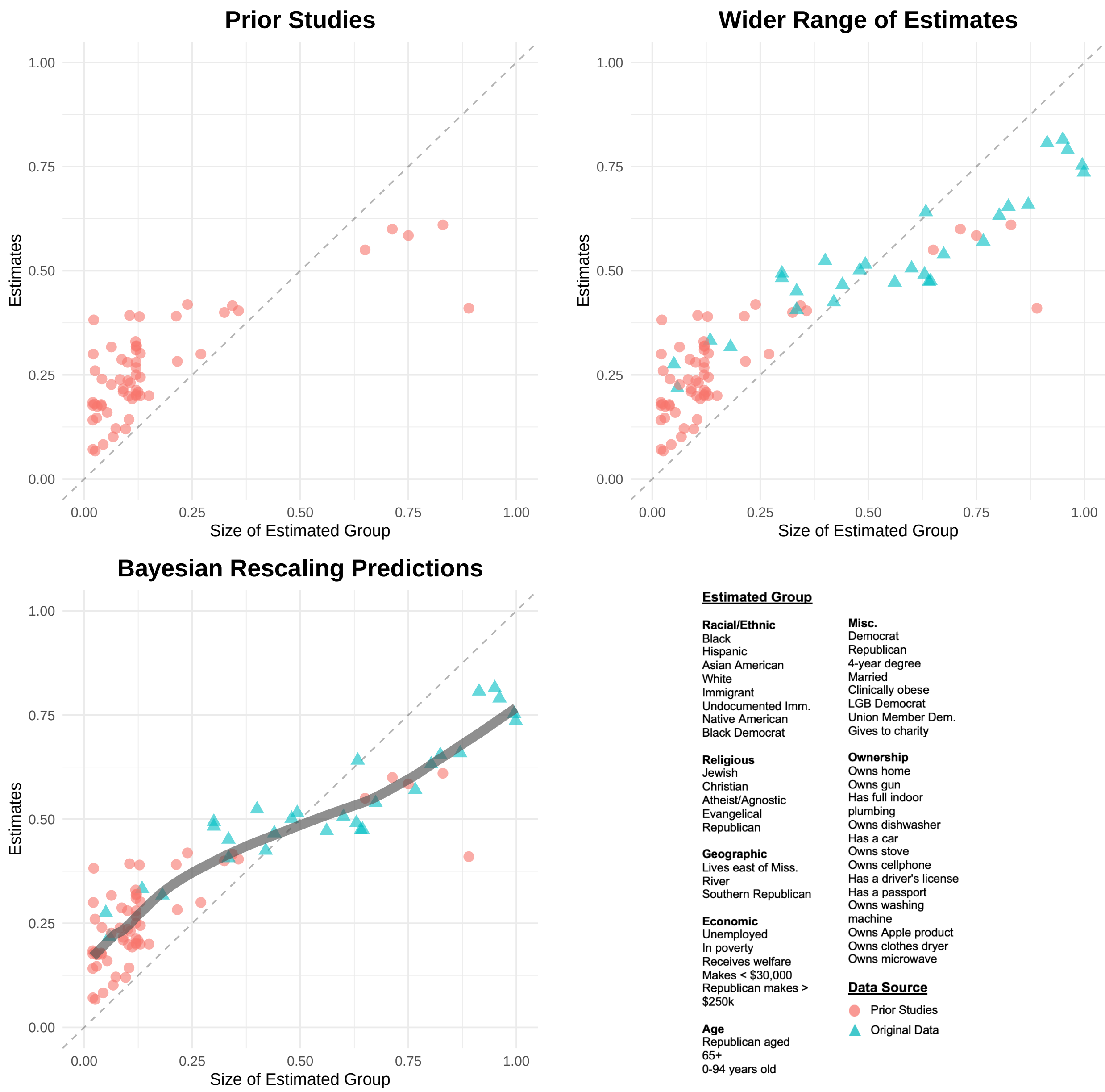
Moreover, this pattern of over- and under-estimation is systematic, following the familiar inverted S-shaped curve characteristic of proportion estimation outside the domain of demographic groups (recall Figure 2). In the second panel of Figure 3 we observe that estimates of racial and ethnic groups follow similar patterns to those of the proportion of the U.S. population that, for instance, holds a college degree, has a driver's license or passport, lives east of the Mississippi River and owns an Apple product, dishwasher, and car (see Tables 1-3 in the Appendix for all estimates and true values from Figure 3). The close similarity in the pattern of errors observed in estimates of racial and non-racial groups is suggestive of a domain-general process, rather than processes that are specific to the perception of racial groups.

In the third panel of Figure 3, we illustrate how well our two-parameter model of Bayesian rescaling captures this pattern of over- and under-estimation in aggregate estimates. Predictions from Equation 4 are represented by the grey line[6]. Our model assumes that people have perfect underlying information about the groups in question but engage in Bayesian rescaling when translating this information into proportion estimates on a survey. Note that our two-parameter model predicts the systematic errors that people make across a range of group sizes, including both racial and non-racial groups.

It is also evident from Figure 3 why Bayesian rescaling has so far been overlooked as a potential explanation for demographic misperceptions. Prior work documenting and explaining demographic misperceptions focuses almost exclusively on estimates of relatively small proportions, represented in the first panel of Figure 3 as points in the shape of a circle. It is therefore unsurprising that the conclusions drawn from this work has emphasized the overestimation of minority groups. Analyzing estimates from these studies and our two original surveys (represented as points in the shape of a triangle, second panel of Figure 3),

[6]See Appendix page 14 for model fitting details.

Figure 3: Estimates of Population Sizes



Estimates of the size of 42 groups (vertical axis) plotted against true values (horizontal axis). The first panel presents mean estimates from existing studies and surveys: the 1991 ANES, 2000 GSS, 2002 ESS, and 6 published studies. The second panel includes additional estimates from original surveys asking about a wider range of demographic groups. In the third panel includes predictions from the Bayesian rescaling model specified in Equation 4. Tables 1-3 in the Appendix report mean estimates and true values for the data in Figure 3.

illustrates the same pattern of over-under estimation characteristic of proportion estimation more generally.

Comparison to Existing Theories

We now turn our attention to understanding the extent to which perceived threat, contact, and Bayesian rescaling are associated with demographic misperceptions. To do so, we partition the 2000 GSS data into four mutually exclusive subsets depending on the type of estimate made by respondents: estimates of the size of local out-groups, local in-groups, national out-groups, and national in-groups, where in-groups and out-groups refer to groups to which respondents do and do not belong, respectively.

With each subset of the data we estimate four models. First, we estimate a *Baseline model* that predicts respondents' estimates with a set of demographic characteristics that prior research suggests may be associated with misestimation error: age, gender, educational attainment, income, marital status, political ideology (e.g., Alba et al., 2005; Herda, 2010). Next, we estimate two theoretically-driven models. The *Threat & Contact model* adds measures of perceived threat and contact to the Baseline model, while the *Rescaling model* adds Bayesian rescaling to the Baseline model. Finally, we estimate a *Full model* containing respondents' demographic characteristics, perceived threat, contact, and Bayesian rescaling. Given the lack of theory suggesting a relationship between perceived threat or contact with estimates of one's in-group, and because the GSS does not measure perceptions of threat for in-groups, we limit our analysis of in-group estimates to the Baseline and Rescaling models.

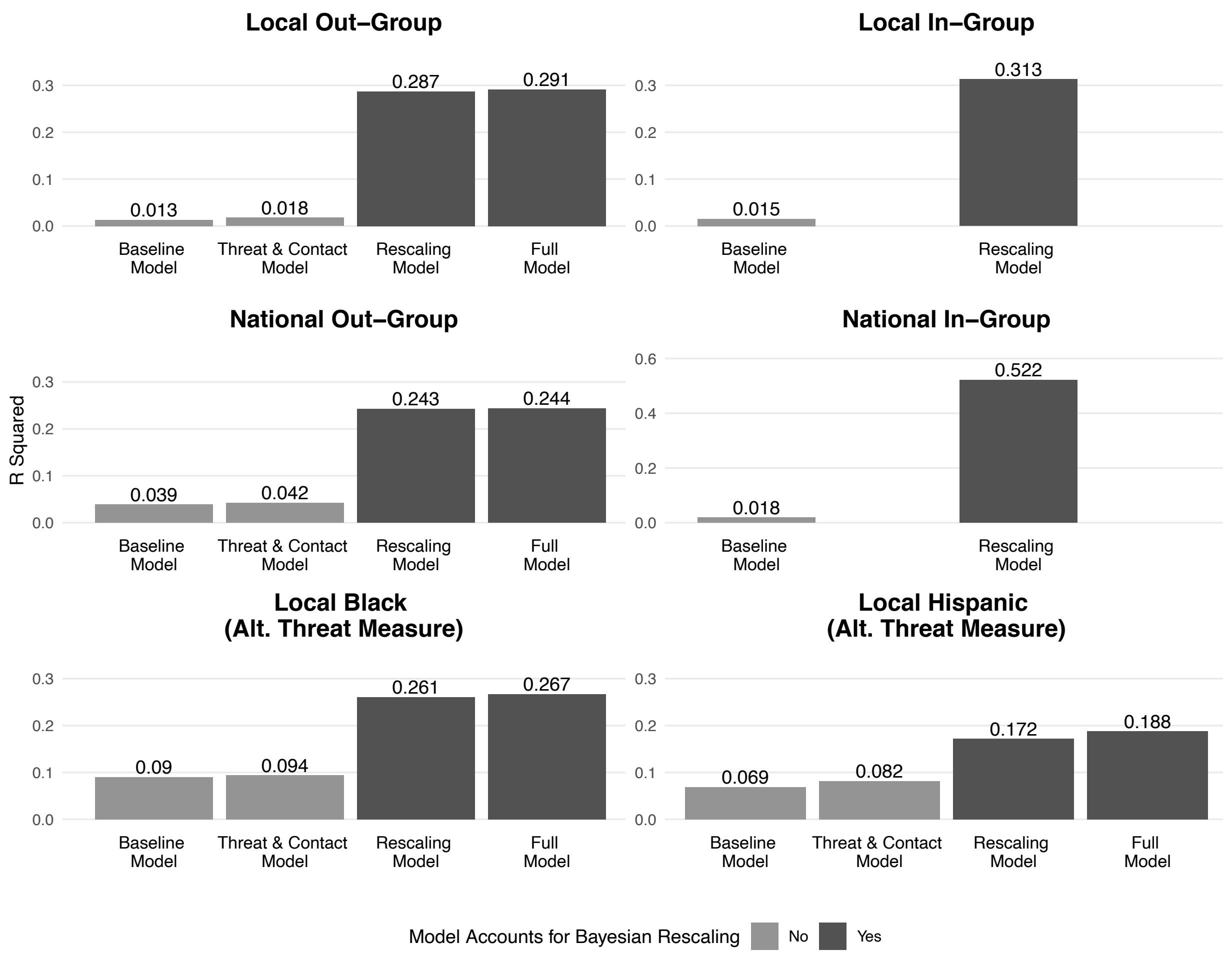
We follow prior work on demographic misperceptions by modeling overestimation *error*, rather than raw estimations. Error is defined as the difference between the actual size of a group and the estimated size of a group. In order to incorporate Bayesian rescaling into our models via Equation 4, we follow a computationally equivalent approach, modeling respondents' estimates and including the true value being estimated on the right hand side of the equation.[7] All models were fit using maximum likelihood estimation (using the *R* function *optim*).[8] To provide a more conservative test of Bayesian rescaling, we assume that everybody engages in Bayesian rescaling in the same way–that is, we estimate only two rescaling parameters (i.e., the prior, δ, and weight, γ) in the *Rescaling* and *Full* models, rather than estimating individual rescaling parameters for each respondent. Estimating individual rescaling parameters would risk model overfitting, given that each respondent only estimated the size of four groups. Rather than allowing each individual to rescale by a different amount and toward a different prior, our conservative approach assumes that all individuals are engaging in Bayesian rescaling in the exact same way, with the same prior, δ, that they weight by the same amount, γ.

Across all subsets of the data, models accounting for Bayesian rescaling explain substantially more variation in estimation errors than those that do not. Figure 4 reports the proportion of variation accounted for (R^2) in each model. For instance, models that account for Bayesian rescaling explain approximately 30% of the variance in respondents' estimates of local out-groups and in-groups (top row of Figure 4), compared to less than 2% in models that do not. Similarly, models that account for Bayesian rescaling explain approximately 24% of variation in estimates of national out-groups and 52% of variation in estimates of national in-groups, compared to to less than 2% in models that do not (middle row of Figure 4). When we restrict our analysis to only White respondents' estimates of local Black and Hispanic population and use Alba et al.'s (2005) operationalization of perceived threat,

[7] When modeling estimates of the size of *national* groups, we assume that individuals are rescaling the groups' true *national* prevalence. Prior work has sometimes accounted simultaneously for both local and national prevalence, since individuals may rely on local group size to estimate national demographics (e.g., Wong, 2007). However, a group's local prevalence is likely strongly correlated with individuals' contact with and perceived threat of that group. As a conservative test of our account, therefore, we incorporate only the true national size into our models of Bayesian rescaling.

[8] To make more direct comparisons between models within each data group (e.g., local in-group, local out-group, national in-group, national out-group), we use only observations without missing values for each of the variables contained in the Full model. To ensure that we were isolating a stable maximum, we reran our models with multiple starting parameters. Confidence intervals were calculated using a one thousand sample bootstrap, in which we randomly re-sampled individuals from the data set. We used a fixed normal error term in probability space. While errors in the probability space are not normal, this decision results in the maximum likelihood minimizing squared error, which is simpler to calculate. This simplification did not appear to substantially affect our results.

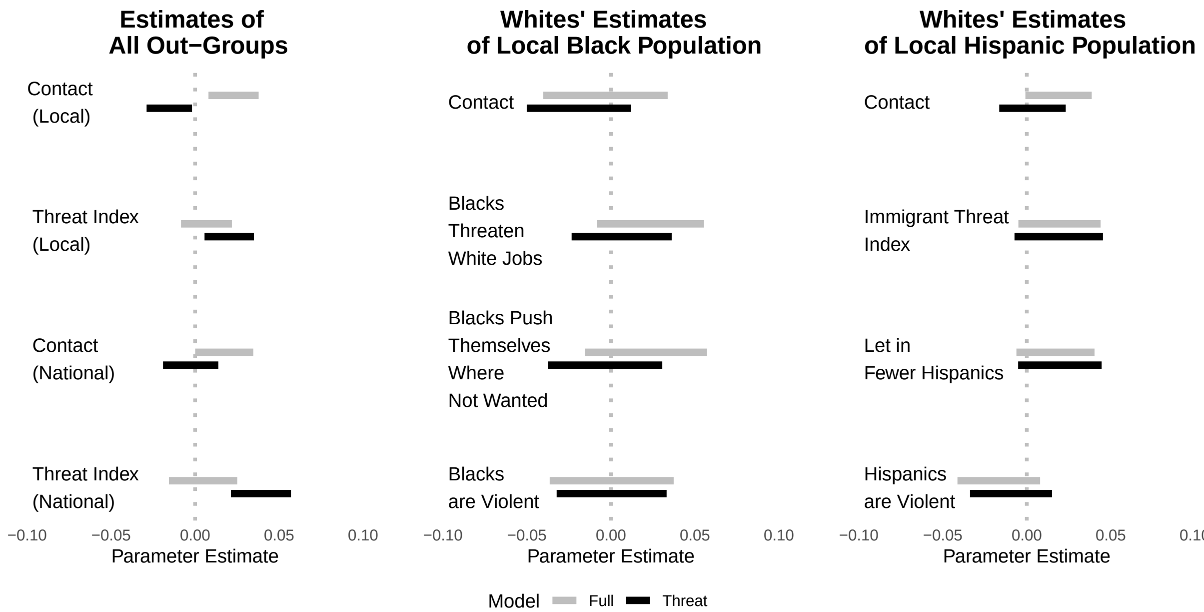
Figure 4: Model Fit Across Different Estimate Categories



Model fit statistics (R^2) for all four models, applied to six different types of estimates. The Baseline model includes only respondents' demographic characteristics (e.g., gender, education); the Threat & Contact model adds one parameter for perceived threat and one parameter for contact to the Baseline model; the Rescaling model adds two rescaling parameters to the Baseline model; and the Full model includes all parameters (i.e., for threat, contact, and Bayesian rescaling, plus demographic characteristics).

accounting for Bayesian rescaling improves model fit from 9% to 26% and 7% to 17%, respectively (bottom row of Figure 4).

Figure 5: Perceived Threat and Contact Parameter Estimates



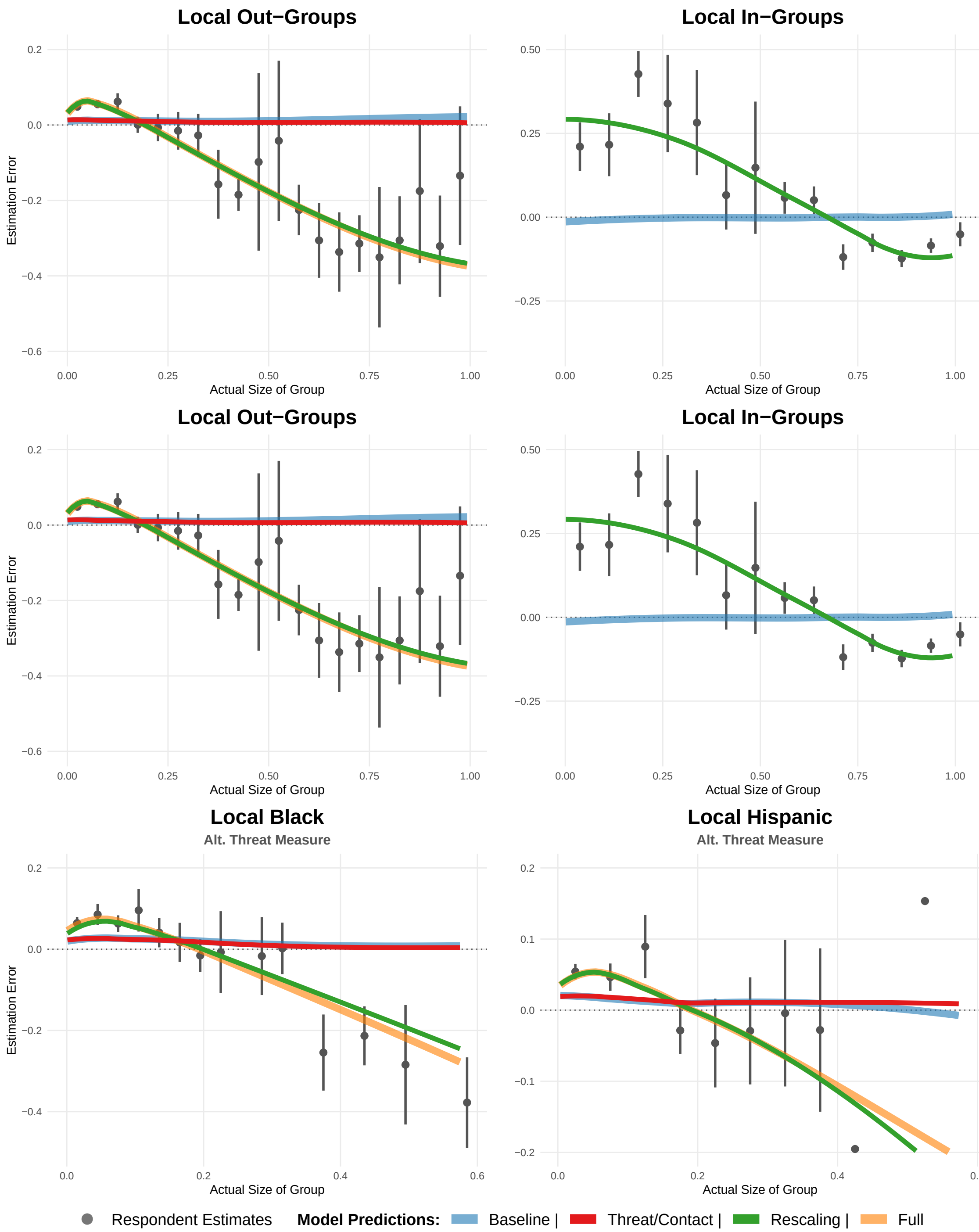
Bootstrapped 95% confidence intervals, calculated with 1,000 bootstrapped simulations, of parameter estimates of perceived threat and contact from the Threat (does not include Bayesian rescaling) and Full (includes Bayesian rescaling) models. Parameter estimates represent the change in respondents' estimates associated with a change of one standard deviation in the measure of threat or contact.

In contrast, perceived threat and contact explain very little variation in respondents' estimates. Across all subsets of the data, adding measures of perceived threat and contact to the Baseline model accounts for almost no additional variation in respondents' estimates. Similarly, almost no additional variation is explained by adding these measures to the Rescaling model (i.e., to create the Full model). Indeed, the parameter estimates reported in Figure 5 reveal small and inconsistent associations between demographic estimates and perceived threat and contact (see Tables 5-7 in the Appendix for full model results). For estimates of all local and national out-groups (Fig. 5, first panel), the parameter estimates for contact are positive when accounting for Bayesian rescaling and slightly negative when not; in both cases, they are negligible. For the threat index in particular, the parameter estimates in the Threat & Contact model were numerically small but statistically significantly greater than zero for both local and national out-group estimates, but were small and indistinguishable from zero in the Full model that accounts for Bayesian rescaling. Even in the model that found the largest association between perceived threat and demographic estimates—the Threat & Contact model applied to estimates of national out-groups—an increase of one standard deviation in threat was only associated with an increase of .03 increase in respondents' proportion estimates. We observe similarly small associations when using Alba et al.'s (2005) operationalization of perceived threat (Figure 5, second and third panels). Neither perceived threat nor contact have a reliable association with Whites' estimates of the local Black (second panel) or Hispanic (third panel) populations; parameter estimates of those associations are all numerically small and none differ statistically from zero.

By contrast, Bayesian rescaling accounts for much of the errors in demographic estimation. In Figure 6, we compare the actual errors made by respondents (points) to those predicted by each model (lines) for groups of various sizes (horizontal axis). For estimates of both in-groups and out-groups, at both the local and national level, we observe the familiar pattern of systematic over-estimation for small populations (i.e., positive errors) and under-estimation for large populations (i.e., negative errors). As indicated by the parameter estimates for the prior (δ) reported in Tables 5-7 in the Appendix, the central value toward which individuals are adjusting their estimates depends on the particular set of populations they are estimating. For estimates of out-groups, which mostly included estimates of minority groups, this central value is relatively small (δ = .19 for local out-groups, .35 for national out-groups). For estimates of in-groups, which included many estimates by White respondents of the size of their own majority group, the prior was larger (δ = .66 for local in-groups, .51 for national in-groups). This suggests that respondents aligned their prior to the particular groups they were estimating: a small prior for small groups, a larger prior for larger groups.

In sum, across every subset of the data, models that account for Bayesian rescaling fit the

Figure 6: Model Predictions



Respondents' mean binned estimates are presented as gray points with vertical 95% confidence intervals, while predictions from each model are presented as colored lines. Full model results, including the Bayesian Information Criterion (BIC) for each model, are reported in Tables 5-7 of the Appendix.

pattern of errors made by respondents substantially better than those that do not. Whereas the Baseline and Threat & Contact models predict almost no error in estimates of groups of any size, the Rescaling and Full models closely predict the errors made by respondents. This is particularly evident for larger groups, where predictions from the Rescaling model deviate substantially from those from the Baseline and from the Threat & Contact predictions. However, even for smaller out-groups with true values of less than .20, predictions from the Rescaling model fit the data substantially better. In contrast, we observe no improvement in model fit when adding measures of perceived threat and contact to the Baseline and Rescaling models.

Discussion

This paper examines the origins of demographic misperceptions by considering the psychology of how people perceive and estimate numeric information more broadly. We present a model of Bayesian rescaling, a process in which individuals adjust their proportion estimates (represented mentally in log-odds) toward a prior belief about the size of groups in general. We find strong support for Bayesian rescaling in a dataset containing a larger and more diverse range of demographic estimates than past work. Moreover, we show that misperceptions about both racial and non-racial groups follow the same S-shaped pattern documented in other domains of proportion estimation. We also find strong support for our Bayesian rescaling model in estimates of the size of racial groups on the GSS, both for estimates of in-groups and out-groups at the local and national level. In contrast, we find almost no empirical support for theories of perceived threat and social contact.

Our findings have implications for how scholars interpret demographic misperceptions reported on surveys. Previous interpretations attribute demographic misperceptions to underlying ignorance or misinformation about the size of certain groups, driven by differential social contact with minority groups or perceptions of these groups as threatening (Allport, 1954; Nadeau et al., 1993; Semyonov et al., 2004; Dixon, 2006). However, we demonstrate that these misperceptions are the product of a reasonable approach to estimating quantities under uncertainty. Indeed, we show that the types of demographic misperceptions observed for decades closely mirror precisely what we would expect to see when people have accurate underlying information, but under uncertainty adjust their estimates toward a reasonable prior belief.

The findings presented here also have implications for how misperceptions about *non-demographic* quantities are interpreted. Political scientists are often interested in people's perceptions of quantities relating to the economy, such as the proportion of government spending dedicated to welfare, the unemployment rate, and inflation (Conover et al., 1986; Holbrook and Garand, 1996; Kuklinski et al., 2000). For instance, past studies have documented errors in the public's perception of the human and financial cost of armed conflict (Berinsky, 2007) and the proportion of the federal budget spent on foreign aid (Gilens, 2001; Scotto et al., 2017). Given the findings presented here, it is very likely that Bayesian rescaling also drives at least some of the error in those estimates.

These findings also raise questions for the growing body of research that attempts to change attitudes by correcting misperceptions. These studies show that correct information often succeeds in correcting misperceptions, but fails to change downstream attitudes (e.g., Kuklinski et al., 2000; Lawrence and Sides, 2014; Hopkins et al., 2019). For instance, providing correct information about the size of the immigrant population leads to substantial belief updating, but almost no change in attitudes toward immigration policy (Hopkins et al., 2019). Given the results presented here, one likely explanation is that attitudes are rooted in internal perceptions of group size, not the rescaled proportions provided on surveys. Indeed, one of the key implications of Bayesian rescaling is that people will make systematic estimation errors even when these internal perceptions are perfectly accurate.

Future research is also needed to identify other factors that *can* account for error in both demographic and non-demographic groups beyond that predicted by Bayesian rescaling. Our model offers one way of capturing and understanding the systematic errors that appear in estimates of all demographic groups, but it does not preclude other correlates of demographic misperceptions. When seeking to explain misperceptions about the size of a particular group, candidate explanations should first account for the error that appears systematically across estimates of all groups before invoking factors specific to a particular group. Indeed, the results presented here may help to explain the associations observed in previous studies that document the correlates of demographic misperceptions, which are small in size relative to the large estimation errors they seek to explain (e.g., Gilens, 1999; Sides and Citrin, 2007; Ahler and Sood, 2018). By first accounting for variation due to the way people estimate the size of groups in general, future work can better account for the variation related to estimates of specific groups.

.

References

Ahler, D. J. and G. Sood (2018). The parties in our heads: Misperceptions about party composition and their consequences. *The Journal of Politics 80*(3), 964–981.

Alba, R., R. G. Rumbaut, and K. Marotz (2005). A Distorted Nation: Perceptions of Racial/Ethnic Group Sizes and Attitudes Toward Immigrants and Other Minorities. *Social Forces 84*(2), 901–919.

Allport, G. (1954). The Nature of Prejudice. Reading, MA: Addison-Wesley.

Barth, H. C. and A. M. Paladino (2011). The Development of Numerical Estimation: Evidence Against a Representational Shift. *Developmental Science 14*(1), 125–135.

Berinsky, A. J. (2007). Assuming the Costs of War: Events, Elites, and American Public Support for Military Conflict. *The Journal of Politics 69*(4), 975–997.

Blalock, H. M. (1967). *Toward a Theory of Minority-group Relations*, Volume 325. New York: Wiley.

Blumer, H. (1958). Race Prejudice as a Sense of Group Position. *Pacific Sociological Review 1*(1), 3–7.

Bobo, L. D. (1999). Prejudice as Group Position: Microfoundations of a Sociological Approach to Racism and Race Relations. *Journal of Social Issues 55*(3), 445–472.

Bonilla-Silva, E. (2001). *White Supremacy and Racism in the Post-Civil Rights Era.* Lynne Rienner Publishers.

Carpini, M. X. D. and S. Keeter (1996). *What Americans Know about Politics and Why it Matters.* Yale University Press.

Cohen, D. J. and D. Blanc-Goldhammer (2011). Numerical Bias in Bounded and Unbounded Number Line Tasks. *Psychonomic Bulletin & Review 18*(2), 331–338.

Conover, P. J., S. Feldman, and K. Knight (1986). Judging Inflation and Unemployment: The Origins of Retrospective Evaluations. *The Journal of Politics 48*(3), 565–588.

Converse, P. (1964). The Nature of Belief Systems in Mass Publics. In *Ideology and Discontent*, pp. 206–61. New York: Free Press.

Coppock, A. and O. A. McClellan (2019). Validating the Demographic, Political, Psychological, and Experimental Results Obtained From a New Source of Online Survey Respondents. *Research & Politics 6*(1).

Dawson, M. C. (1995). *Behind the Mule: Race and Class in African-American Politics*. Princeton University Press.

Dixon, J. C. (2006). The Ties That Bind and Those That Don't: Toward Reconciling Group Threat and Contact Theories of Prejudice. *Social Forces 84*(4), 2179–2204.

Duffy, B. (2018). *The Perils of Perception: Why We're Wrong About Nearly Everything*. Atlantic Books.

Erlick, D. E. (1964). Absolute Judgments of Discrete Quantities Randomly Distributed Over Time. *Journal of Experimental Psychology 67*(5), 475.

Flynn, D., B. Nyhan, and J. Reifler (2017). The Nature and Origins of Misperceptions: Understanding False and Unsupported Beliefs About Politics. *Political Psychology 38*, 127–150.

Fossett, M. A. and K. J. Kiecolt (1989). The Relative Size of Minority Populations and White Racial Attitudes. *Social Science Quarterly 70*(4), 820.

Gallagher, C. A. (2003, September). Miscounting Race: Explaining Whites' Misperceptions of Racial Group Size. *Sociological Perspectives 46*(3), 381–396.

Gallup, G. and F. Newport (1990). Americans Strongly in Tune with The Purpose of Earth Day 1990. *Gallup Poll Monthly 295*(3), 5–21.

Gilens, M. (1999). *Why Americans Hate Welfare: Race, Media, and the Politics of Antipoverty Policy.* University of Chicago Press.

Gilens, M. (2001). Political Ignorance and Collective Policy Preferences. *American Political Science Review 95*(2), 379–396.

Gonzalez, R. and G. Wu (1999). On the Shape of the Probability Weighting Function. *Cognitive Psychology 38*(1), 129–166.

Gorodzeisky, A. and M. Semyonov (2018). Perceptions and Misperceptions: Actual Size, Perceived Size and Opposition to Immigration in European Societies. *Journal of Ethnic and Migration Studies*, 1–19.

Herda, D. (2010). How Many Immigrants? Foreign-Born Population Innumeracy in Europe. *Public Opinion Quarterly 74*(4), 674–695.

Holbrook, T. and J. C. Garand (1996). Homo Economus? Economic Information and Economic Voting. *Political Research Quarterly 49*(2), 351–375.

Hollands, J. and B. P. Dyre (2000). Bias in Proportion Judgments: the Cyclical Power Model. *Psychological Review 107*(3), 500.

Hopkins, D. J., J. Sides, and J. Citrin (2019). The Muted Consequences of Correct Information About Immigration. *The Journal of Politics 81*(1), 315–320.

Howard, J. A., D. G. Renfrow, and J. Delamater (2003). Handbook of social psychology.

Key, V. O. (1966). *The Responsible Electorate.* Harvard University Press Cambridge, MA.

Kuklinski, J. H., P. J. Quirk, J. Jerit, D. Schwieder, and R. F. Rich (2000). Misinformation and the Currency of Democratic Citizenship. *The Journal of Politics 62*(3), 790–816.

Landy, D., B. Guay, and T. Marghetis (2018). Bias and Ignorance in Demographic Perception. *Psychonomic Bulletin & Review 25*(5), 1606–1618.

Lawrence, E. D. and J. Sides (2014). The consequences of political innumeracy. *Research & Politics 1*(2), 2053168014545414.

Lee, E., F. Karimi, C. Wagner, H.-H. Jo, M. Strohmaier, and M. Galesic (2019). Homophily and minority-group size explain perception biases in social networks. *Nature Human Behaviour 3*(10), 1078–1087.

Marghetis, T., B. M. Guay, A. Karlapudy, and D. Landy (2018). The Psychophysics of Society: Uncertain Estimates of Invisible Entities. In *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Martinez, M. D., K. D. Wald, and S. C. Craig (2008). Homophobic Innumeracy? Estimating the Size of the Gay and Lesbian Population. *Public Opinion Quarterly 72*(4), 753–767.

Nadeau, R., R. G. Niemi, and J. Levine (1993). Innumeracy About Minority Populations. *Public Opinion Quarterly 57*(3), 332–347.

Nakajima, Y. (1987). A model of empty duration perception. *Perception 16*(4), 485–520.

Quillian, L. (1995). Prejudice as a Response to Perceived Group Threat: Population Composition and Anti-immigrant and Racial Prejudice in Europe. *American Sociological Review*, 586–611.

Schille-Hudson, E. B. and D. Landy (2020). Scaling uncertainty in visual perception and estimation tasks.

Scotto, T. J., J. Reifler, and J. vanHeerde Hudson (2017). We Spend How Much? Misperceptions, Innumeracy, and Support for the Foreign Aid in the United States and Great Britain. *Journal of Experimental Political Science 4*(2), 119–128.

Semyonov, M., R. Raijman, A. Y. Tov, and P. Schmidt (2004, December). Population size, perceived threat, and exclusion: a multiple-indicators analysis of attitudes toward foreigners in Germany. *Social Science Research 33*(4), 681–701.

Sides, J. (2013). Group-Centrism in American Public Opinion. *Unpublished Manuscript, George Washington University*. http://home.gwu.edu/ jsides/groupcentrism.pdf.

Sides, J. and J. Citrin (2007, July). European Opinion About Immigration: The Role of Identities, Interests and Information. *British Journal of Political Science 37*(3), 477–504.

Sigelman, L. and R. G. Niemi (2001). Innumeracy about Minority Populations: African Americans and Whites Compared. *The Public Opinion Quarterly 65*(1), 86–94.

Spence, I. (1990). Visual Psychophysics of Simple Graphical Elements. *Journal of Experimental Psychology: Human Perception and Performance 16*(4), 683.

Stevens, S. S. (1957). On the Psychophysical Law. *Psychological Review 64*(3), 153.

Theiss-Morse, E. (2003). Characterizations and Consequences: How Americans Envision the American People. In *Annual Meeting of the Midwest Political Science Association, Chicago, IL.*

Tversky, A. and D. Kahneman (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty 5*(4), 297–323.

Varey, C. A., B. A. Mellers, and M. H. Birnbaum (1990). Judgments of Proportions. *Journal of Experimental Psychology: Human Perception and Performance 16*(3), 613.

Wilcox, J. and W. C. Roof (1978). Percent Black and Black-white Status Inequality: Southern Versus Non-southern Patterns. *Social Science Quarterly 59*(3), 421.

Wong, C. J. (2007). "Little" and "Big" Pictures in Our Heads: Race, Local Context, and Innumeracy About Racial Groups in the United States. *Public Opinion Quarterly 71*(3), 392–412.

Supplementary Appendix

Bayesian Reasoning and Demographic Misperceptions

Contents

1 Aggregated Estimates Methodology

1.1 1991 American National Election Study

The 1991 American National Election Study (ANES) Pilot featured a probability-based sampling design. Therefore, weights are not necessary except to account for small differences in the probability of a household member being selected, which we do not account for. Respondents were first asked: "In the country as a whole, what percent of the U.S. population today would you say is black?" before being asked "What percent would you say is Jewish?" and "What percent would you say is Hispanic?"

1.2 2000 General Social Survey

The 2000 General Social Survey (GSS) was conducted using probability sampling. NORC indicates that no weighting is necessary when analysing these data. Respondents were asked "Just your best guess-what percentage of the United States population is each group?" for the following groups: Black, White, Hispanic, Asian, Jewish, and Native American.

1.3 2002 European Social Survey

Respondents from 22 of the 24 countries surveyed in the European Social Survey (ESS) answered a question asking, "Out of every 100 people living in [country], how many do you think were born outside [country]?" Estimates of the proportion of citizens in a country that are foreign born are weighted to account for the unequal probability of selection within each of those countries. We followed the procedure described by Sides & Citrin (2007) to calculate the true size of the foreign-born population in each country.

1.4 2016 Cooperative Congressional Election Study

The 2016 Cooperative Congressional Election Study (CCES) was conducted on an online non-probability sample collected by YouGov. Respondents were asked about the size of the following groups: White, Black, Hispanic, Asian, Republican, Democrat, unemployed, gun owner, college graduate, homeowner.

1.5 2018 Lucid Survey

Due to the nature of the study that was the main purpose of the 2018 Lucid Survey, the survey was run on a convenience sample of 1,258 internet users in the U.S. collected by Lucid. Due to this sampling approach, we make no claims of generalizability and estimates are not weighted.

1.6 Estimates from Published Studies

For nearly all of the published studies, respondents' mean estimates and the sizes of the groups being estimated available directly in the text (see Table 1 below for a list of all studies, along with the respondents' mean estimates and the size of the groups being estimated).

In one case, Hopkins et al. (2018), we obtained these quantities by re-analyzing the data made available on Dataverse. Hopkins and colleagues conducted 7 experiments in 5 surveys to examine the effect of correct information on estimates of the size of the foreign-born population in the U.S. and immigration attitudes. The authors removed Hispanic respondents from all analyses. No weights are reported and thus our analysis features unweighted estimates. Since the studies all featured experiments where correct information was given to a subset of respondents, we obtain estimates of the foreign-born population in the U.S. from respondents in conditions where correct information was not provided prior to the estimation question. We exclude the second survey because the subset of respondents in the control condition, who were provided with correct information, is very small (N = 103). The data from these studies come from the 2006 Cooperative Congressional Election Study (Study 1), 2010 Knowledge Networks survey (Study 3), a 2017 Morning Consult survey (Study 4), and the 2010 Cooperative Congressional Election Study (Study 5).

2 Aggregated Estimates Data

2.1 Table Containing Quantities from Figure 3 in Main Text

Table 1: Data from Prior Surveys

Source	Group	Actual	Est. (Mean)	Estimate (SE)	Model Pred.
1990 ANES	Jewish	0.020	0.184	0.008	0.158
1990 ANES	Hispanic	0.090	0.216	0.008	0.262
1990 ANES	Black	0.120	0.318	0.008	0.287
2000 GSS	Native American	0.020	0.141	0.005	0.146
2000 GSS	Jewish	0.020	0.177	0.005	0.146
2000 GSS	Asian	0.040	0.176	0.005	0.188
2000 GSS	Black	0.120	0.310	0.005	0.277
2000 GSS	Hispanic	0.130	0.245	0.005	0.285
2000 GSS	White	0.750	0.585	0.004	0.590
2002 ESS	Immigrant	0.020	0.071	0.002	0.091
2002 ESS	Immigrant	0.025	0.067	0.002	0.103
2002 ESS	Immigrant	0.029	0.147	0.004	0.110
2002 ESS	Immigrant	0.039	0.179	0.009	0.129
2002 ESS	Immigrant	0.044	0.083	0.004	0.136
2002 ESS	Immigrant	0.053	0.160	0.005	0.151
2002 ESS	Immigrant	0.063	0.227	0.007	0.164
2002 ESS	Immigrant	0.067	0.101	0.003	0.170
2002 ESS	Immigrant	0.073	0.121	0.003	0.177
2002 ESS	Immigrant	0.083	0.239	0.005	0.188
2002 ESS	Immigrant	0.100	0.280	0.006	0.207
2002 ESS	Immigrant	0.101	0.236	0.004	0.208
2002 ESS	Immigrant	0.103	0.199	0.004	0.210
2002 ESS	Immigrant	0.104	0.143	0.004	0.211
2002 ESS	Immigrant	0.107	0.231	0.004	0.214
2002 ESS	Immigrant	0.111	0.193	0.003	0.219
2002 ESS	Immigrant	0.120	0.203	0.004	0.227
2002 ESS	Immigrant	0.125	0.209	0.004	0.232
2002 ESS	Immigrant	0.216	0.282	0.004	0.304
2002 ESS	Immigrant	0.325	0.400	0.007	0.376

Note: Survey abbreviations: American National Election Study (ANES), General Social Survey (GSS), European Social Survey (ESS). Estimates of the size of the immigrant population on the 2002 European Social Survey come from 22 different countries, each of which is recorded as a separate estimate in the table as each as a different true value.

Table 2: Data from Prior Studies

Source	Group	Actual	Est. (Mean)	Model Pred.
Ahler & Sood (2018)	Republican > $250k	0.022	0.382	0.293
Ahler & Sood (2018)	LGB Democrat	0.063	0.317	0.338
Ahler & Sood (2018)	Atheist/Agnostic Democrat	0.087	0.287	0.353
Ahler & Sood (2018)	Union Member Dem.	0.105	0.393	0.363
Ahler & Sood (2018)	Republican aged 65+	0.213	0.391	0.400
Ahler & Sood (2018)	Black Democrat	0.239	0.419	0.407
Ahler & Sood (2018)	Evangelical Republican	0.343	0.416	0.430
Ahler & Sood (2018)	Southern Republican	0.357	0.404	0.433
Citrin & Sides (2008)	Immigrant	0.120	0.280	0.279
Gallup & Newport (1990)	Jewish	0.024	0.180	0.162
Gallup & Newport (1990)	Hispanic	0.090	0.210	0.256
Gallup & Newport (1990)	Black	0.121	0.320	0.283
Hopkins et al. (2018)	Undocumented Imm.	0.030	0.174	0.155
Hopkins et al. (2018)	Immigrant	0.120	0.268	0.265
Hopkins et al. (2018)	Immigrant	0.120	0.250	0.265
Hopkins et al. (2018)	Immigrant	0.120	0.214	0.265
Hopkins et al. (2018)	Immigrant	0.130	0.302	0.273
Lawrence & Sides (2014)	Unemployment rate	0.096	0.120	0.177
Lawrence & Sides (2014)	Black	0.120	0.200	0.202
Lawrence & Sides (2014)	Poverty rate	0.130	0.200	0.211
Lawrence & Sides (2014)	Hispanic	0.150	0.200	0.230
Lawrence & Sides (2014)	4 year college degree	0.270	0.300	0.324
Lawrence & Sides (2014)	White	0.650	0.550	0.575
Theiss-Morse (2003)	On welfare	0.021	0.300	0.263
Theiss-Morse (2003)	Jewish	0.025	0.260	0.271
Theiss-Morse (2003)	Asian	0.041	0.240	0.295
Theiss-Morse (2003)	Hispanic	0.119	0.330	0.352
Theiss-Morse (2003)	Black	0.128	0.390	0.356
Theiss-Morse (2003)	White	0.713	0.600	0.513
Theiss-Morse (2003)	Christian	0.830	0.610	0.551

Table 3: Data from Original Studies

Source	Group	Actual	Est. (Mean)	Estimate (SE)	Model Pred.
2016 CCES	Unemployed	0.050	0.276	0.013	0.248
2016 CCES	Asian	0.058	0.218	0.010	0.258
2016 CCES	Black	0.134	0.333	0.010	0.325
2016 CCES	Hispanic	0.181	0.317	0.009	0.353
2016 CCES	Gunowner	0.300	0.494	0.011	0.408
2016 CCES	4 year college degree	0.334	0.451	0.008	0.421
2016 CCES	Republican	0.440	0.466	0.007	0.460
2016 CCES	Democrat	0.480	0.501	0.007	0.474
2016 CCES	Owns Home	0.630	0.491	0.009	0.527
2016 CCES	White	0.766	0.571	0.009	0.584
2018 Lucid	Gunowner	0.300	0.483	0.007	0.419
2018 Lucid	4 year college degree	0.334	0.407	0.007	0.433
2018 Lucid	Clinically obese	0.400	0.524	0.006	0.457
2018 Lucid	Has a passport	0.420	0.425	0.007	0.464
2018 Lucid	Makes < $30,000	0.493	0.515	0.007	0.489
2018 Lucid	Lives east of Miss. River	0.561	0.472	0.006	0.512
2018 Lucid	Currently married	0.600	0.506	0.006	0.526
2018 Lucid	Has a car	0.633	0.641	0.006	0.538
2018 Lucid	Owns Apple product	0.640	0.474	0.007	0.541
2018 Lucid	Owns Home	0.644	0.475	0.006	0.542
2018 Lucid	Owns dishwasher	0.674	0.539	0.006	0.553
2018 Lucid	Owns clothes dryer	0.803	0.632	0.006	0.610
2018 Lucid	Owns wash. machine	0.824	0.654	0.006	0.621
2018 Lucid	Has a driver's license	0.870	0.659	0.006	0.650
2018 Lucid	Owns stove	0.914	0.807	0.007	0.684
2018 Lucid	Has a cellphone	0.950	0.815	0.006	0.726
2018 Lucid	Owns microwave	0.961	0.790	0.006	0.744
2018 Lucid	Has full indoor plumbing	0.995	0.753	0.007	0.856
2018 Lucid	0-94 years old	0.999	0.736	0.009	0.902

3 General Social Survey Question Wording

3.1 Contact

Respondents were first asked whether they personally know anyone from each group that they do not report belonging to themselves. (e.g., knowwht, knowblk). Rs were then separately asked whether they feel close to each group they personally know a person from.

- Do you personally know any [Whites, Blacks, Hispanics, Jews, Asians]
- Are any of these [Whites, Blacks, Hispanics, Jews, Asians] people that you feel close to?

3.2 Perceived Threat

3.2.1 Main Perceived Threat Index

As described in the main text, we created a mean index comprised of 8 items, which are listed below:

- **Violence:** Do the people in the following groups tend to be violence prone or do they tend not to be prone to violence.
- **Contribution to Country:** Has the group has made one of the most important positive contributions tothis country, an important contribution, some contribution, or little positive contribution to this country? (English, Italians, Chinese, Jews, Blacks, Mexicans, Vietnemese, Cubans, Irish, Puerto Ricans, Japenese)
  - Note that while this question asks about Jews and Blacks, the three remaining groups asked about in this question do not perfectly correspond to the groups we use in this study (white, Hispanic, and Asian). We combine multiple ethnic groups for these three remaining racial groups and report the Cronbach's alpha for each below. We create mean indices for each group using these ethnic groups below.
    * White: English, Italians, Irish (Cronbach's alpha = .72)
    * Hispanic: Puerto Ricans, Mexicans, Cubans (Cronbach's alpha = .87)
    * Asian: Chinese, Vietnemese, Japenese (Cronbach's alpha = .79)
- **Committment to Equal Treatment of All Groups:** Whites committed to fair and equal treatment of all: Where would you rate Whites in general on this scale? A score of 1 means that you think almost all of the people in the group have a commitment to the fair and equal treatment of all groups in society. A score of 7 means that you think almost everyone in the group lacks commitment to the fair and equal treatment of all groups in society.
- **Social Distance (Neighbor):** Would you favor living in a neighborhood where half of your neighbors were [Whites, Blacks, Hispanics, Asians, Jews]?
- **Social Distance (Family):** How would you respond to a close relative marrying a [White, Black, Hispanic, Asian, Jewish] person?
- **Intelligence:** Do people in these groups tend to be unintelligent or tend to be intelligent?
- **Committment to Strong Families:** Where would you rate Whites in general on this scale? A score of 1 means that you think almost all of the people in the group have a commitment to strong families. A score of 7 means that you think almost everyone in the group lacks a commitment to strong families.
- **Laziness:** Do the people in the following groups tend to be hard working or do they tend to be lazy?

3.2.2 Alba et al. (2005) Percieved Threat Measures

We follow Alba et al.'s (2005) operationalization of perceived threat using survey items asking specifically about African Americans and Hispanics, including questions measuring racial resentment, threat posed by Hispanic immigrants. For African Americans, the questions reflect physical, cultural, and economic threat: respondents were asked how violence-prone African Americans are, whether they agree that African Americans should not push themselves where they are not wanted, and whether a White person would not get a job or promotion because an equally or less qualified Black person got one instead. While the GSS does not directly measure perceptions of threat posed by Hispanics, Alba et al. use measures of the perceived threat of immigrants to measure perceptions of threat posed by Hispanics. Respondents were asked whether more immigration makes it harder to keep the country united, leads to higher crime rates, and causes native-born Americans to lose their jobs. We took the mean of these three items to create an index of perceived threat posed by Hispanics (Cronbach's $\alpha = .77$).[1] Following Alba and colleagues, we also include items measuring whether there should be more immigrants from Spanish-speaking countries and how violence-prone Hispanics are.

- **Blacks Shouldn't Push Themselves:** Blacks/ African-Americans shouldn't push themselves where they're not wanted (original coding: 1 = agree strongly, 4 = disagree strongly) (RACPUSH)
- **Black Violence:** How violence prone are Blacks? (original coding: 1 = violent, 7 = not violent) (VIOLBLKS)
- **Black Job Threat:** What do you think the chances are these days that a white person won't get a job or promotion while an equally or less qualified black person gets one instead? (original coding: 1 = very likely, 3 = not very likely) (DISCAFF)
- **Hispanic Violence:** How violence prone are Hispanic Americans? (original coding: 1 = violent, 7 = not violent) (VIOLHSPS)
- **Immigrant Threat Index:** What do you think will happen as a result of more immigrants coming to this country?
    1. Make it harder to keep the country united (IMMUNITE)
    2. Higher crime rates (IMMCRMUP)
    3. People born in the U.S. losing their jobs (IMMNOJOB)
- **Let in More/Less Hispanic Immigrants:** What about the number of immigrants from Latin America (that is, Spanish-speaking countries of the Americas)? Should it be increased a lot, increased a little, left the same as it is now, decreased a little, or decreased a lot? (original coding: 1 = increased a lot, 5 = decreased a lot) (LETINHISP)

[1] While Alba et al. use the GSS item that measures preferences for increased immigration from all foreign countries, we use the GSS item that measures preferences for increased immigration from Latin America specifically.

4 National Estimate Distributions

Figures 4 and 5 in the main text report the distribution of respondents' estimates of the size of local out-groups and in-groups, along with predictions from each of the models. In the case of national estimates, each estimated group has only one true value (e.g., the proportion of the U.S. population that is Hispanic was .13). Therefore, in Figure 6 in the main text we report the mean of respondents' estimates alongside model predictions. Here we report the full distributions of respondents' estimates of national in-groups and out-groups.

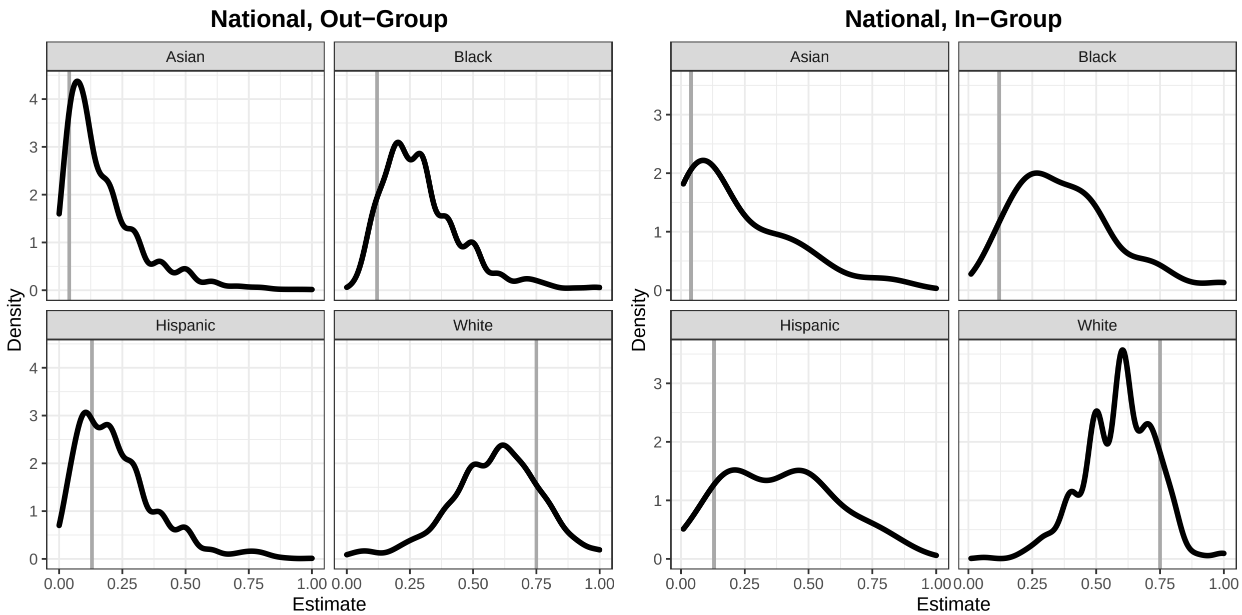


Figure 1: Distribution of National Estimates
Response distributions for estimates of out-group and in-group estimates at the national level. In each plot the vertical line indicates the true size of the group being estimated.

5 Parameter Estimates for All Models

Table 4: Local Estimate Models

	Out-Group				In-Group	
	Baseline	Threat/Contact	Rescaling	Full	Baseline	Rescaling
Age	-0.014	-0.017	-0.017	-0.015	0.029	0.042
	(0.004)	(0.005)	(0.004)	(0.005)	(0.008)	(0.011)
Female	0.022	0.021	0.02	0.022	-0.005	-0.014
	(0.006)	(0.006)	(0.008)	(0.008)	(0.013)	(0.025)
Education	-0.003	0.001	-0.008	-0.01	-0.001	0.008
	(0.003)	(0.004)	(0.003)	(0.004)	(0.008)	(0.01)
Conservatism	-0.009	-0.01	-0.009	-0.009	0.014	0.014
	(0.004)	(0.004)	(0.004)	(0.004)	(0.008)	(0.01)
Income	-0.005	-0.005	-0.008	-0.009	-0.008	0.000
	(0.006)	(0.006)	(0.005)	(0.005)	(0.012)	(0.013)
Married	-0.001	-0.002	-0.013	-0.012	0.006	0.031
	(0.006)	(0.006)	(0.007)	(0.007)	(0.014)	(0.027)
Perceived Threat		0.01		0.003		
		(0.004)		(0.004)		
Contact		-0.008		0.013		
		(0.004)		(0.004)		
Prior (δ_{odds})*			0.442	0.418		1.519
			(0.025)	(0.029)		(0.199)
Weight (γ)			0.444	0.437		0.35
			(0.027)	(0.03)		(0.049)
N	3313	3313	3313	3313	1169	1170
R^2	.09	.09	.26	.27	.07	.17
BIC	-2124.659	-2125.99	-3187.412	-3191.126	235.51	-170.951

Note: Parameter estimates and standard errors for models of local out-groups and in-groups. Parameter estimates for the prior (δ) are reported in odds space and correspond to .187 (out-group Rescaling model), .175 (out-group Full model), and .655 (in-group Rescaling model) in probability space. Variables for age, education, conservatism, income, perceived threat, and contact are standardized to have a mean of 0 and standard deviation of 1. Measures of gender and marital status are indicator variables, where 1 equals female and married, respectively.

Table 5: National Estimate Models

	Out-Group				In-Group	
	Baseline	Threat/Contact	Rescaling	Full	Baseline	Rescaling
Age	-0.001	-0.002	-0.003	-0.004	0.000	0.013
	(0.004)	(0.004)	(0.004)	(0.005)	(0.007)	(0.006)
Female	0.124	0.07	0.07	0.122	-0.046	-0.002
	(0.007)	(0.01)	(0.009)	(0.008)	(0.011)	(0.014)
Education	-0.02	-0.03	-0.028	-0.017	-0.015	0.004
	(0.004)	(0.005)	(0.004)	(0.005)	(0.008)	(0.006)
Conservativism	0.000	0.000	0.000	-0.002	0.004	0.007
	(0.005)	(0.004)	(0.004)	(0.005)	(0.007)	(0.006)
Income	-0.007	-0.009	-0.009	-0.008	-0.033	-0.015
	(0.006)	(0.006)	(0.006)	(0.007)	(0.01)	(0.008)
Married	0.056	-0.005	-0.005	0.051	-0.061	0.013
	(0.007)	(0.011)	(0.009)	(0.009)	(0.012)	(0.014)
Perceived Threat		0.00		0.02		
		(0.005)		(0.005)		
Contact		0.006		-0.001		
		(0.005)		(0.004)		
Prior (δ_{odds})*			0.697	0.708		1.028
			(0.05)	(0.044)		(0.071)
Weight (γ)			0.435	0.44		0.292
			(0.027)	(0.026)		(0.029)
N		3328	3328	3328	1154	1154
R^2	.02	.02	.52	.52	.04	.24
BIC	-1833.291	-1850.435	-2812.135	-2800.448	-162.422	-1018.75

Note: Parameter estimates and standard errors for models of national out-groups and in-groups. Parameter estimates for the prior (δ) are reported in odds space and correspond to .345 (out-group Rescaling model), .351 (out-group Full model), and .510 (in-group Rescaling model) in probability space. Variables for age, education, conservatism, income, perceived threat, and contact are standardized to have a mean of 0 and standard deviation of 1. Measures of gender and marital status are indicator variables, where 1 equals female and married, respectively.

Table 6: White's Estimates of Local Black Population

	White's Estimates of Local Black Population			
	Baseline	Threat & Contact	Rescaling	Full
Age	-.02	-0.022	-0.017	-0.021
	(0.009)	(0.009)	(0.009)	(0.009)
Female	0.064	0.064	0.05	0.049
	(0.013)	(0.014)	(0.015)	(0.02)
Education	-0.013	-0.11	-0.016	-0.011
	(0.007)	(0.008)	(0.007)	(0.009)
Ideology (Conservative)	-0.017	-0.018	-0.011	-0.014
	(0.007)	(0.007)	(0.007)	(0.008)
Income	-0.034	-0.033	-0.02	-0.02
	(0.009)	(0.01)	(0.01)	(0.013)
Married	-0.004	-0.004	-0.037	0.037
	(0.013)	(0.013)	(0.017)	(0.023)
Black Americans Threaten White Jobs		0.003		0.011
		(0.008)		(0.008)
Black Americans Push Where Unwanted		-0.002		0.008
		(0.009)		(0.01)
Black Americans are Violent		0.00		-0.002
		(0.009)		(0.009)
Contact		-0.01		-0.006
		(0.008)		(0.01)
Prior (δ_{odds})*			0.334	0.325
			(0.057)	(0.079)
Weight (γ)			0.277	0.263
			(0.078)	(0.128)
N	503	503	503	503
Rˆ2	0.09	0.094	0.261	0.267
BIC	-359.587	-336.512	-478.093	-459.344

Note: Parameter estimates and standard errors for models of Whites' estimates of local Black and Hispanic populations. Parameter estimates for the prior (δ) are reported in odds space and correspond to .180 (Rescaling model) and .179 (Full model). Variables for age, education, conservatism, income, perceived threat, and contact are standardized to have a mean of 0 and standard deviation of 1. Measures of gender and marital status are indicator variables, where 1 equals female and married, respectively.

Table 7: White's Estimates of Local Hispanic Population

	White's Estimates of Local Hispanic Populations			
	Baseline	Threat & Contact	Rescaling	Full
Age	-.018	-0.017	-0.015	-0.013
	(0.006)	(0.006)	(0.006)	(0.006)
Female	0.037	0.036	0.022	0.025
	(0.009)	(0.009)	(0.01)	(0.01)
Education	-0.019	-0.013	-0.014	-0.01
	(0.005)	(0.005)	(0.005)	(0.005)
Ideology (Conservative)	-0.009	-0.012	-0.01	-0.012
	(0.005)	(0.005)	(0.005)	(0.005)
Income	-0.027	-0.029	-0.023	-0.025
	(0.008)	(0.08)	(0.008)	(0.008)
Married	0.007	0.008	-0.016	-0.014
	(0.009)	(0.009)	(0.01)	(0.01)
Immigrant Threat Index		0.003		0.009
		(0.008)		(0.006)
Decrease Hispanic Immigration		0.01		0.01
		(0.007)		(0.006)
Hispanics are Violent		0.01		-0.007
		(0.007)		(0.006)
Contact		0.002		0.011
		(0.005)		(0.005)
Prior (δ_{odds})*			0.506	0.479
			(0.049)	(0.042)
Weight (γ)			0.495	0.496
			(0.047)	(0.038)
N	769	769	769	769
R^2	0.069	0.082	0.172	0.188
BIC	-771.147	-775.906	-852.275	-840.566

Note: Parameter estimates and standard errors for models of White's estimates of local Black and Hispanic populations. Parameter estimates for the prior (δ) are reported in odds space and correspond to .206 (Rescaling model) and .188 (Full model) in probability space. Variables for age, education, conservatism, income, perceived threat, and contact are standardized to have a mean of 0 and standard deviation of 1. Measures of gender and marital status are indicator variables, where 1 equals female and married, respectively.

6 Modelling Aggregate estimates From Past Studies

To model the group size estimates from a variety of past surveys, as plotted in Figure 3 of the Main Text, we implemented our two-parameter model (Equation 4) in the JAGS programming language for Bayesian modeling, using the rjags package in R. Since each survey contributed multiple observations, the implementation was hierarchical, clustered at the survey level; model parameters γ and δ for each individual survey were assumed to be normally distributed. This allowed us to account for between-survey variability.

We modeled the 'best case scenario' where individuals' underlying information about demographic proportions is uncertain but unbiased and correct; thus, when generating explicit estimates of demographic proportions, they engage in Bayesian rescaling of their uncertain knowledge. For each estimate, Y_{ij}, of the size of a demographic group i, estimated on survey j, we assumed that respondents knew the true size, p_{ij}, expressed in log-odds:

$$r_{p_{ij}} = \log\Big(\frac{p_{ij}}{1 - p_{ij}}\Big) \tag{1}$$

To generate a numerical estimate in log-odds, this belief about the group size, $r_{p_{ij}}$, is then rescaled toward a prior belief about typical group sizes (γ_j):

$$\Psi'(r_{p_{ij}}) = \gamma_j r_{p_{ij}} + (1 - \gamma_j)\delta_j \tag{2}$$

This rescaled estimate is finally expressed as a probability:

$$Y_{ij} = e^{\Psi'(r_{p_{ij}})} e^{\Psi'(r_{p_{ij}})} + 1 \tag{3}$$

Equation 4 then combines Equations 1, 2, and 3:

$$Y_{ij} = e^{\delta_j(1-\gamma_j)} p_{ij}^{\gamma_j} e^{\delta_j(1-\gamma_j)} p_{ij}^{\gamma_j} + (1 - p_{ij})^{\gamma_j} \tag{4}$$

We thus modeled demographic proportion estimates, Y_{ij}, as a function of the true value, p_{ij}, with normally-distributed survey-specific parameters γ_j and δ_j.

7 Multilevel Model Comparing Bayesian Rescaling to Perceived Threat and Social Contact

We conducted a robustness check using a multi-level model to ensure that the stronger performance of the Bayesian rescaling models relative to the threat & contact models is not a product of the nested structure of the GSS data. Respondents on the GSS are drawn multiple counties in the U.S., with some counties represented more than others. To confirm that the success of our model is not driven by particular individuals or regions, we use a hierarchical linear model of group size estimates, clustered at the level of both individuals and counties. This model has no intercept and only two predictors: the predictions of our Rescaling model, and the predictions of the Threat & Contact model. These two predictors were included as fixed effects and as by-county random effects (since respondents were clustered within counties). When respondents gave multiple estimates (e.g., when estimating the size of local out-groups), we also included by-respondent random effects.

We used this model to compare our Bayesian rescaling model with the threat/contact model on five datasets: (1) estimates of the size of local out-groups, (2) estimates of the size of respondents' own group, (3) estimates of the local size of the Black population, (4) estimates of the size of the local Hispanic population, and (5) estimates of the size of national out-groups. Models where implemented in R, using the lme4 package. Since our model does not include intercepts, the model's fixed parameters indicate which model does a better job of accounting for respondents' group size estimates, taking into account the hierarchical structure of the data (multiple responses from each individual, and multiple individuals within each county).

Table 8: Parameter Estimates from Multilevel Model

	Fixed Effects	
	Threat	Rescaling
Estimates of Local Out-Groups	0.05 (0.04)	1.00 (0.06)
Estimates of Local In-Groups	-0.02 (0.08)	1.02 (0.09)
Whites' Estimates of Local Black Population	-0.30 (0.13)	1.30 (0.14)
Whites' Estimates of Local Hispanic Population	0.29 (0.15)	0.79 (0.14)
Estimates of National Out-Groups	-0.03 (0.03)	1.03 (0.03)

Results are summarized in Table 8. In every case, our rescaling model performed significantly better than the threat-contact model. Note that, for all five datasets, the parameter estimates for the Threat & Contact prediction were close to 0, while the parameter estimates for our rescaling prediction were always close to 1, indicating that we the best way to predict individuals' actual responses was essentially to take the prediction of our rescaling model and ignore the prediction of the threat/contact model. This model confirms that, even after accounting for individual and county variation, the predictions of the Rescaling model were systematically related to group size estimates, while the predictions of the Threat & Contact model were not.