Bias in the Self-Knowledge of Global Communities

Eleanor Brower (erbrower@indiana.edu) **David Landy** (dlandy@indiana.edu)
Department of Psychological and Brain Sciences, Indiana University
1101 E. 10th St. Bloomington, IN 47408 USA

Abstract

A plethora of research over the past two decades has demonstrated that citizens in countries around the world dramatically overestimate the size of minority demographic groups and underestimate the size of majority groups. Researchers have concluded that this misestimation is a result of characteristics of the group being estimated, such as level of threat the group poses and the amount of exposure someone has with to the group. However, explanations of this misestimation have largely ignored theoretical models of perception and measurement, such as those developed in classic psychophysics. This has led to interpretations that are at variance with modern theories of measurement. We present a model which combines an understanding of the nature of human estimations with a conceptualization of uncertainty, which extends to accommodate bias. We apply this model to three large-scale datasets collected by the Ipsos MORI research group. Model fits from our approach suggest that to a considerable degree, the errors people make are due to uncertainty rather than bias. These biases are quite different in character from those that other groups have reported. Many of the present biases, furthermore, are shared widely across different countries.

Keywords: demographic perception, psychophysics, bias, uncertainty, proportional reasoning, numerical reasoning

Introduction

People inhabit large, complex, and networked communities. Many of these groups are far too large to be directly surveyed, and yet people have strong intuitions about their structure. People have good macro-scale intuitions; in the US, for instance, most people believe (correctly) that European-descended white Americans form the majority of the population. However, when it comes to specific proportions of people in one's national or local community who match a particular descriptor—e.g., the proportion of people who identify as Muslim, or the proportion who report being happy—people are consistently incorrect, and sometimes wildly so.

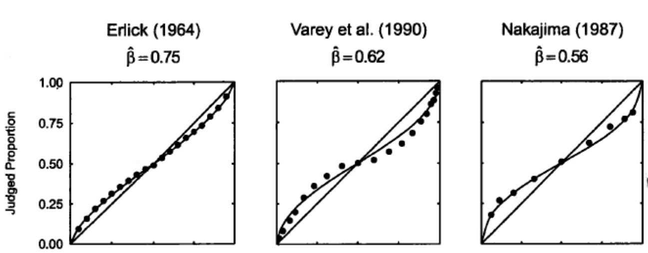

Figure 1: (reprinted from Hollands & Dyre, 2000): Classic examples of the 'over-under' pattern from laboratory experiments in psychophysics.

For many years, psychophysicists have also found errors in proportion estimation in laboratory tasks. People make similar patterns of errors when estimating how many immigrants live in their country as they do when they estimate the number of dots presented on a screen, or the length of a set of lines (Hollands & Dyre, 2000; see Figure 1).There is a general consensus in the psychophysical literature that a perfectly *unbiased* person, when uncertain of the true proportion in question, displays a pattern of responses characterized by overestimating small proportions and underestimating large proportions.

Overestimating small proportions and underestimating large proportions is exactly the pattern of behavior that people display when making estimates about demographics. For instance, US citizens at one time estimated, on average, that about a quarter of our federal budget was allocated to foreign aid; the true value is closer to 2% (Gilens, 2001). People in the US and Europe massively overestimate the proportion of immigrants (Herda, 2013), and US populations overestimate populations of Jewish, Muslim, Asian, and Black Americans (Gallagher, 2003; Wong, 2007), as well as LGBT populations (Martinez, Wald, & Craig, 2008). In the last few years, a series of studies by the Ipsos MORI group has fleshed out this general pattern by conducting a series of multinational surveys designed to help characterize worldwide self-perception. They report that people systematically misestimate a large variety of demographic facts, ranging from the proportion of atheists, to the proportion of people who report being happy.

Explanations from political scientists and sociologists for why people overestimate minority populations or rare phenomena have focused on features of the underestimated groups themselves, and have ignored the over-underestimation pattern so familiar to psychologists. For instance, many researchers have noticed that people tend to perceive smaller groups as more socially threatening (Allport, 1954; Wong, 2007), so-called "phobic innumeracy". Alternatively, it has been suggested that perhaps the media treats groups—especially small or stigmatized groups—differently, leading to biased impressions based on media exposure (Herda, 2013). Finally, simple misinformation might lead to misperceptions. For instance, if people misunderstand the medical standard of obesity, they might misclassify obese individuals as overweight, or even as normal weight, leading to biased perceptions (in this case, the prediction would be an underestimate of obesity).

Recently, Landy et al. (2017) proposed that the general misestimation in demographic proportion estimation is driven primarily by psychophysical phenomena, and not biases such as phobia, media misrepresentation, or specific innumeracy regarding percentage scales. Even in the absence of any biases, people would misestimate demographic proportions in roughly the same ways they have been found to misestimate other proportions in laboratory settings. Nonetheless, it is very plausible a person estimating demographic proportions is operating under biased information. It is very plausible, in fact, that media misrepresentation, social fears, and innumeracy *do* exist, and may impact numerical judgments. Empirically, we find variation in how much overestimation there is. This could certainly be due to bias. Current models of demographic perception cannot explain the variation in overestimation. Therefore, the current manuscript sets out to address how to formally characterize the structured information, including biased information, contained in the beliefs people have about their large communities.

Model

In their 2017 paper, Landy et al. made the connection that psychophysics has observed the same kind of pattern—overestimation of small values and underestimation of large values. This paper suggests an approach to explain this pattern which includes bias, but to be clear, we do not claim that this is a uniquely successful conceptualization, nor the only model that might capture the broad patterns in the data.

At its simplest, the approach we describe suggests that the over-under pattern of errors is due to how people respond under *uncertainty*. Imagine a situation in which a person had no information at all about a proportion or regarded their information as completely unreliable: in Bayesian terms, a person's belief would be uniform across the probabilities. On average, the rational way to minimize error in such a situation is to guess that the proportion is one half—if you have no information and/or no certainty, guessing 50% is sensible (in that it is the mean of the uniform values across the range). In contrast, if a person has perfect confidence in their information, then the sensible thing is to use that information to formulate an estimate. This idea of guessing one half suggests that if one has limited or imperfect information, it is sensible to do something in the middle: to guess a number between what is signaled by one's information and 50%. This exactly reduces to overestimating small values and underestimating large ones. We will call this process *hedging one's bets*, or just *hedging*. The darker lines reflect more bet hedging.

The upper left panel of Figure 2 displays results from a simulation showing this pattern: the darkest line corresponds to complete uncertainty, the lightest to complete certainty. The curved lines show the results from intermediate models. The darker lines reflect more bet hedging. This method has been formalized many times, including by Huttenlocher and colleagues (Huttenlocher, Hedges, & Duncan, 1991), and more recently by Bayesian cognitive scientists (Lee & Danilieiko, 2014).

Models of proportions tend to take for granted that people become more accurate when estimating very extreme proportions. Ours is no exception. It turns out (see Landy, Guay, & Marghetis, 2017; Petzschner, 2012) that the best way to reconcile the fact that people both hedge intermediate proportions and accurately estimate extreme ones is to assume that people intuitively experience not the proportion itself, but the logarithm of the odds. That is, if one sees 20% black circles, one converts this into an internal scale something like $log(20/80)$. Not only does this model produce sensible data fits, it actually reduces precisely to the most common psychophysical models based on Steven's power laws (e.g., Hollands & Dyre, 2000).

We start from the assumption that people encode proportions in terms of ratios representing the odds of a random item belonging to the group. We assume that the prior expectation for an unknown demographic proportion is 50%. This is reasonable, since the average size of all subsets of the population must be 50%. Our bias-free model is

$$\psi(p) = \frac{p^{\gamma}}{p^{\gamma} + (1-p)^{\gamma}}$$

Here, γ is the certainty of the responder, realized as the relative precisions of the prior and the data and p is the proportion of the demographic group in question. This baseline model is incapable of accounting for misinformation (*bias*), as it presumes accurate input data. In the next section, we develop a model that accounts for one simple sort of bias.

How does bias fit in?

We also posit that the action of *uncertainty* on proportion perceptions is distinct from the action of *bias* (Kuklinski et al., 2000). We conceive of bias not simply as misestimation, but as misestimation as a result of misinformation of one sort or another. We can consider two kinds of bias: overperception and underperception, depending on whether one relies on misinformation that skews responses upwards or downwards. This misinformation could be caused by media misrepresentation, sampling errors, or internal affective state or beliefs (e.g., threat), or by anything else. Regardless of its origin, this bias has the effect of transforming the typical "S-shape" curve, sometimes bending it out of shape altogether.

Conceptually, bias in the model that occurs in the source information is different from endogenously generated bias. Furthermore, bias could come from many sources, and could be highly dependent on the frequency of the source event. Here, we consider a very simple kind of bias that assumes a 'fixed effect' of bias on the perceived information—that is, we consider bias that altered the apparent frequency of a group by a multiplier, b, while leaving all other apparent frequencies identical.

$$\psi_b(p, b) = \frac{b^{\gamma} p^{\gamma}}{b^{\gamma} p^{\gamma} + (1-p)^{\gamma}}$$

Unfortunately, this model cannot directly distinguish bias from uncertainty from a single data point. Intuitively, any particular response results from a combination of adjustment toward the mean (λ) and biased sampling (b), and the relative contributions cannot be uniquely distinguished. To compensate, we ask the same question of multiple people in different countries and treat bias and uncertainty as normally distributed random variables.

A visualization of the impact of differing amounts of bias on a person who is completely certain of their responses (i.e., who does not hedge their bets at all) can be found in the upper right panel of Figure 2. The lightest blue line shows a person who is radically over-perceiving, while the darkest blue line shows a person who is radically under-perceiving. Each of the remaining panels of Figure 2 shows predicted response patterns under different combinations of bias and uncertainty. Notice that in both cases of under- and over-perception, sometimes people overestimate and sometimes people underestimate: overestimation does not result purely from bias, but rather from some combination of bias, 'certainty', and the true value in question.

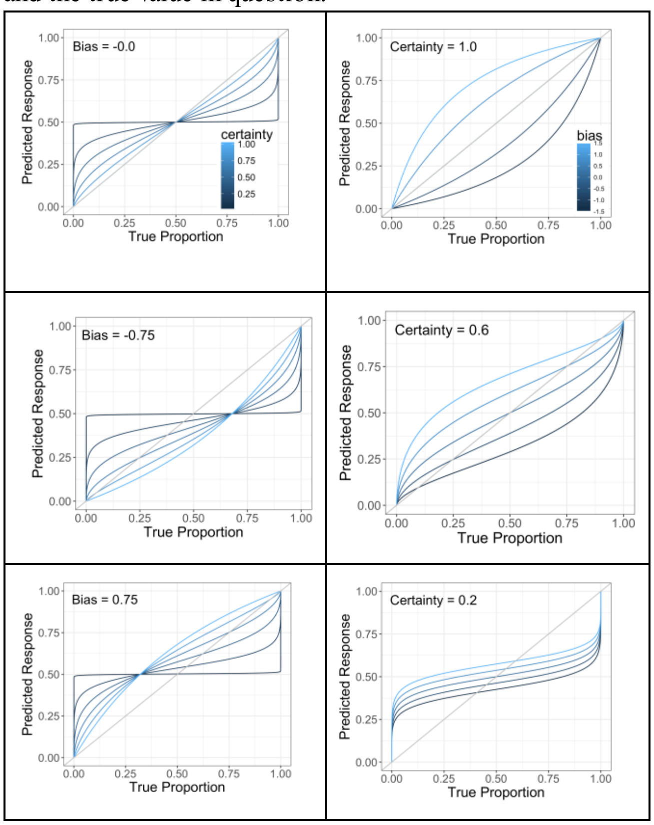

Figure 2: Simulations of the model under different parameter settings. Each shows the predicted response under different kinds of certainty and bias. In the left column, the lower two panels show the effects of bias on varying levels of uncertainty. Bias of -0.75 shifts the pattern of responses such that they are systematically under-perceiving the true values, while +0.75 indicates over-perception. In the right column, the lower two panels show the effects of uncertainty on varying levels of bias. Reducing certainty to 0.6 shifts the pattern of responses to have a flatter curve over the intermediary proportions—evidence of hedging. Certainty of 0.2 shows even more hedging.

Thus, this model correctly captures the idea that people can overestimate when very certain of the true value (Kuklinski et al., 2000).

This model provides a very different, and very clear interpretation of cross-national data. In particular, we describe our interpretation the data collected by Ipsos MORI as part of the "Perils of perception" project conducted yearly since 2013. The model also helps clarify substantial literatures coming from political science (Citrin & Sides, 2008; Herda, 2013; Kuklinski et al., 2000; Wong, 2007)

Analysis Plan

To estimate parameters for the data sets, we used multilevel Bayesian model fitting. We separately estimated parameters for each question, and for each country within each question. We estimate means and 95% highest posterior densities. In addition to the bias and certainty parameters, we estimated a *within-nation variability* parameter, governing the precision of responses across participants: this parameter combines variability due to uncertainty within each individual, and variability caused by heterogeneity between individuals within one country. Because of the small number of items answered by each participant, we were unable to estimate parameters for each participant separately. See the Ipsos MORI Perils of Perception Report 2014 for a full description of the data set.

Bias/Uncertainty Perspective vs. Error Perspective

The perspective we present here is, we believe, standard in psychophysics and measure theory (Hollands & Dyre, 2000; Landy et al., 2017; Petzschner, 2012; Shepard, 1981). On the other hand, it contrasts sharply with the interpretation of proportion estimation data that has been dominant in the popular press, the political science literature (Herda, 2013; Wong, 2007), and indeed in the interpretation Ipsos MORI has given of their own results. These interpretations have usually relied on what we might call *bias* to explain the whole pattern: for example, if people estimate a larger value for the Muslim population than is true of their country, one would interpret that as bias—people have some reason for overestimating, whether it be disproportionate media representation, threat, or something else.

On our account, this kind of bias *also* plays a role, but a secondary one. In essence, we first account for systematic errors in responses by making standard psychophysical assumptions about responses, and then account for residual deviations by invoking bias. As a result of this process, the conclusions we reach often differ from those that a more common (but, in our view, less well-founded) analysis might. For example, it is true that people across many nations overestimate the population of Muslims. However, with two exceptions, the overestimation across countries is well explained by psychophysical estimates compatible with either no bias or a slight negative bias (i.e., under-perception of Muslims, see Fig. 3, Left Panel).

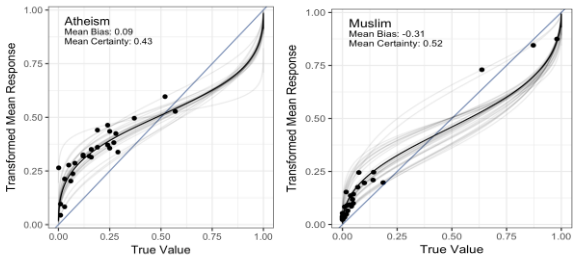

Figure 3: (Left panel) Mean estimate of the proportion of Muslims (mean across subjects taken in log odds space) plotted against actual proportion of Muslims in that country, from the 2015 Ipsos MORI data set. (Right Panel) the same analogous plot of the estimates and actual self-reported atheists in each country.

Estimates of atheism show a very similar pattern: traditional analyses show that nearly all countries overestimate the proportion of their populations that are atheist. Our analysis suggests that this overestimation is primarily caused by people hedging their bets. Once this psychophysical effect is accounted for, the residual is largely unbiased. This unbiased strategy of estimation can be seen in Fig. 3 (Right Panel), in which all countries reported fall very close to the curve predicted by the psychophysical effect. One can see that the two analytical approaches (analyzing raw error versus our psychophysical approach) yield quite different sets of implications: If one takes estimates as reflecting the truth about people's beliefs, one sees overestimation of both Muslims and atheists. If one instead includes psychophysical effects of responses, one sees no net bias in either of these cases. Furthermore, on the traditional analysis, different countries appear to differ starkly in how *much* they overestimate these groups (mostly as a function of the true prevalence). On our account, there is strong cross-national consistency in both items.

An example that reveals bias of *over*-perception, is computer access. Looking at Fig. 4, we see that most country-level polls overestimate how many of their citizens have regular computer access. However, even those who underestimate do so only slightly, and this only occurs in countries with high actual rates of computer access. This over-perception, in turn, is consistent with the fact that the Ipsos MORI polls are conducted online, providing a strong case for why an over-perception bias might be present.

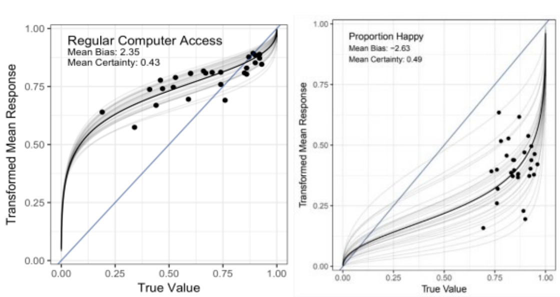

Figure 4: Mean estimates of the proportion of people (Left) with computer access and (Right) who report being happy.

Survey of Ipsos MORI Results

We now examine some of the Ipsos MORI data and share our interpretation of it given our model. For ease of exposition, we divide the Ipsos MORI data up into broad content areas, summarizing each content area separately.

These content areas are invented by the authors as an easy way to express the data patterns, and do not capture anything about the ways these questions were asked on the survey, nor do they have any direct theoretical import.

The Ipsos MORI dataset

This data set includes three years of polling conducted across several countries. The particular countries shifted from year to year, as did the total number of responses. The poll was conducted online, and in publications by Ipsos MORI was weighted to create a representative sample. Table 1 presents the basic descriptive statistics of the data sets themselves.

Table 1: Data from Ipsos MORI included in this data set.

Year	Countries	Total N	Total Responses	Proportion Estimations
2014	14	9941	56,160	7
2015	27	17888	111,684	8
2016	33	19056	122,528	9

Religion

In sharp contrast to most published reports, we do not find strong evidence for bias about religious subgroups. Instead, the errors made across all items were consistent with unbiased perception, except for a perhaps a slight under-perception of **Christianity**. Overall, estimates of **atheism** tended to be slightly less certain than estimates of other factors, but this was especially true of India. The only country to over-perceive atheism was China.

Immigration

Research on public perceptions in political science has focused particularly on immigration (Citrin & Sides, 2008; Herda, 2013). Ipsos MORI asked about immigration in their 2014 and 2015 surveys. In neither case do we find that immigration is over-perceived. Instead, although immigration is indeed overestimated in raw numeric terms, the pattern of observations is quite consistent with accurate perception (2014) or indeed provides substantial evidence for under-perception (2015). That is, people overestimate approximately as you would expect them too if they were uncertain how many immigrants there were, and under-perceived them in the environment.

The immigration results from 2014 and 2015 are quite divergent, possibly because the countries studied in 2015 (a more numerous and more heterogeneous set) are different from those studied in 2014, and these new countries more strongly underperceive immigration. Indeed, this seems even more plausible when we consider just the countries which were surveyed in both years:

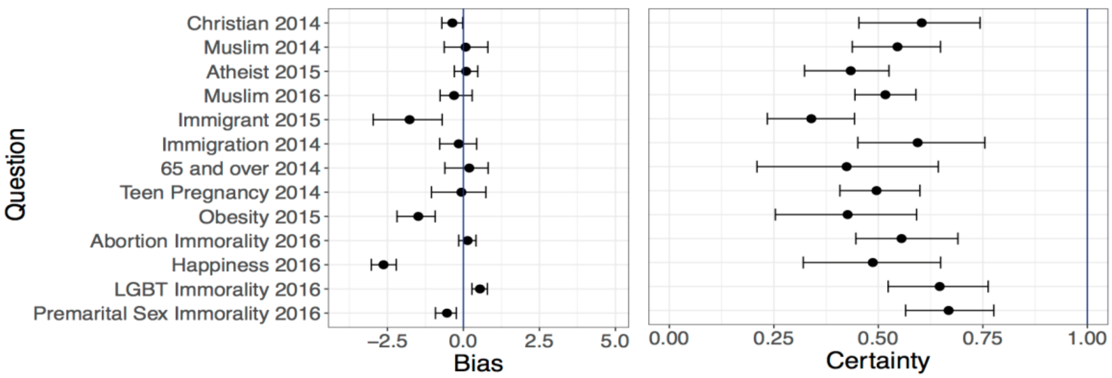

Figure 5: Estimates of the bias, degree of certainty, and variation between countries. Black error bars reflect 95% High-posterior density of the mean value. That is, black bars show how certain the model estimate is.

the mean 2014 and 2015 estimates from these countries were very strongly correlated (r=0.94), and the overall mean error is quite similar (+11% in 2014; +11.8% in 2015).

The immigration results from 2014 and 2015 are quite divergent, possibly because the countries studied in 2015 (a more numerous and more heterogeneous set) are different from those studied in 2014, and these new countries more strongly underperceive immigration. Indeed, this seems even more plausible when we consider just the countries which were surveyed in both years: the mean 2014 and 2015 estimates from these countries were very strongly correlated (r=0.94), and the overall mean error is quite similar (+11% in 2014; +11.8% in 2015).

Health and Lifestyle

One of the most dramatic effects in the entire data set is the under-perception of **obesity** in the 2015 survey data. This pattern is offset by substantial cross-national diversity, and indeed a few countries show little or no under-perception. On the whole, however, most data suggest that people in most countries are uncertain about obesity and fail to recognize it. One possibility is that people are unfamiliar with the medical definition of obesity, and so classify only a subset of medically obese people into this category.

Even more extreme than the under-perception of obesity, however, is the under-perception of self-reported **happiness** (studied in 2016, see Fig 4). Again, this under-perception is associated with huge-cross-national variability. Examination of the raw data may suggest that no work is being accomplished by our model for these data at all: people may be using different heuristics to make their happiness estimates. Alternatively, it is not implausible to assume that citizens of different countries simply differ quite widely in how knowledgeable they are about the happiness of their co-citizens.

Overall, citizens tended to over-perceive the perceived immorality of **LGBT lifestyles** (2016). That is, in most countries, people underestimated how much their compatriots accepted LGBT lifestyles as either moral or not a moral issue. However, this slight general tendency masks statistically significant variation, in which a few countries significantly underperceived opposition to LGBT lifestyle. Again, the 2016 survey data were mixed on this issue.

As mentioned above, we find that people under-perceive opposition to **premarital sex**, with little exception. Despite massive raw overestimates, we find the countries in the 2014 survey to be systematically unbiased in their estimates of the rate of **teen pregnancy** and the **elderly** population. The countries included in the 2016 survey were also found to be unbiased in their perceptions of the immorality of **abortion**.

Discussion

People do not simply report their true beliefs about numerical values. We find here that the assumptions one makes about how people respond has a dramatic impact on where one sees bias. The (traditional) perspective notes that people overestimate Muslims, atheists, immigration, teen pregnancy, and those over 65, while underestimating happiness and Christianity, and focuses on patterns of misestimation. The psychophysically informed perspective explains this overestimation in terms of uncertainty and bias, which indicates that views of Muslims, atheists, teen pregnancy, and retirees are essentially unbiased, and people may actually underperceive the levels of immigration and rates of obesity (as well as happiness and Christianity).

This is important: how we interpret errors in estimation critically informs what stands in need of explanation. Classical efforts, noting that many minorities are overestimated, have focused their efforts on explaining why people over-represent or overestimate minorities: explanations have thus been formed around the ways that society interacts with minority groups. From this perspective, differing amounts of overestimation are hard to interpret and of secondary concern: the fact of overestimation indicates bias. On the other hand, the psychophysical interpretation takes into account properties of measurement and response scales, and thus considers overestimation of minorities perfectly normal behavior, requiring no special explanation. The explanation is, in fact, in the model: people always tend to overestimate small proportions under uncertainty.

What *does* stand in need of explanation is the degree of certainty people feel about certain demographics, and the residual bias that pushes estimates up and down relative to the baseline of misestimation. This means that the complex of 'surprising' results that stand in need of explanation is different, and more heterogeneous. From our perspective, we find explanations that target minorities, such as the 'social threat' hypothesis, not to be very powerful—they reach further (with less supporting data), to explain what our model naturally takes into consideration. On the other hand, our approach reveals phenomena that the traditional approach misses. In finer detail, the two approaches differ vastly in how they look at cross-national variation. Where the traditional approach finds a very large difference between a country that overestimates a value by 20 points and one that is accurate, the psychophysical approach does not—so long as both follow the same trend (e.g. the value is small in the first country, and moderately large in the second). So, which countries exhibit similarity or even agreement differs strongly between the two perspectives.

Both approaches do agree on one thing: there is surprising cross-cultural and cross-national consensus in the error patterns. What happens in one country, is for many issues, much like what happens in another. This degree of consensus appears to vary for different items but is overall strongly present. Nearly all countries in the Ipsos MORI sample under-perceive premarital sex disapproval, despite the fact that these countries vary vastly in their true rates of premarital sex approval and treat premarital sex extremely differently. Countries with Muslim populations that range from 0.1- 20% perceive these populations in a very similar (unbiased) way, even though their media and cultures treat Muslims differently. These countries include Israel, India, Singapore, Denmark, the US, and Thailand—countries with different kinds of relationships between their majority and minority cultures, making a broad social argument like the "phobic innumeracy" argument less convincing. These patterns point to a more surprising uniformity across cultures and approaches globally (or at least in the particular sample Ipsos MORI collected) than the traditional approach apprehended.

The current work contributes a practical approach to psychophysics in the presence of misinformation and bias. The generality of classical psychophysical approaches to response measurement is limited by its focus on unbiased perception. We find that these models have broad applicability well beyond their classical application, and so promise to prove quite robust as a modeling framework. Beyond this, we advance the perspective that classic psychophysics can naturally be accommodated within a rational inference framework, and that doing so reconciles two classic approaches: log scaling and Steven's power law.

Results showing misestimation are being read, often at face value, by lawmakers, scientists, politicians, and the public, as well as forming a key application area and testbed for social science theorizing. We believe that it is vitally important to use the best available behavioral science to understand how people respond to numerical scales: we know that people do not in general have numerical beliefs, nor do they somehow uniquely give numbers that correspond to those beliefs. At the same time, people *do* have structured perceptions, and we can access information about that structure. This suggests that political scientists and pollsters would be well-served by exploring alternative tasks that can access structure without misleading. One alternative is to ask people to rank order different proportions. Our lab is exploring this alternative (Haussecker & Landy, 2018), and finding that it is often possible to reconstruct metric properties of beliefs from rank ordering tasks in practical contexts of general public interest. Rank ordering may not be the best way to access metric beliefs about proportions, but we believe it is a better one than simply asking for estimates, since it does not provide the convenient illusion of a direct response.

This is an exciting time to be doing behavioral science. As we explore and collect the abundance of new data available to us, it is important to remember the advances made by scientists working directly in the lab. In the case of response and measure theory, there is a real risk that a new generation of tech-savvy and data-sophisticated scientists will fall into errors that scientists of the past foresaw and avoided.

References

Allport, G. (1954). Chapter 3: Formation of In-Groups. In *The Nature of Prejudice* (25th Anniv. Edition, pp. 29–46).

Citrin, J., & Sides, J. (2008). Immigration and the Imagined Community in Europe and the United States. *Political Studies*, *56*, 33–56.

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

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

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

Haussecker, C. & Landy, D. (2018). Ordinal Ranking as a Method for Assessing Real-World Proportional Representations. *Proceedings of the Cognitive Science Society*.

Herda, D. (2013). Too Many Immigrants?: Examining Alternative Forms of Immigrant Population Innumeracy. *Sociological Perspectives*, *56*(2), 213–240.

Hollands, J. G., & Dyre, B. P. (2000). Bias in proportion judgments: The cyclical power model. *Psychological Review*, *107*(3), 500–524.

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, *98*(3), 352–376.

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

Landy, D., Guay, B., & Marghetis, T. (2017). Bias and ignorance in demographic perception. *PBR*, 1–13.

Lee, M., & Danilieiko, L. (2014). Using cognitive models to combine probability estimates. *Judgment and Decision Making 9*(3), 259–273.

Martinez, M. D., Wald, K. D., & Craig, S. C. (2008). Homophobic Innumeracy? *Public Opinion Quarterly*, *72*(4), 753–767.

Petzschner, F. (2012). *Magnitude Estimations in Humans*. Ludwig-Maximilians-Universität, München.

Shepard, R. N. (1981). Psychological relations and psychophysical scales *Journal of Math Psych*, *24*, 21–57.

Wong, C. (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.