The psychophysics of society: Uncertain estimates of invisible entities

Tyler Marghetis[1,2] tmarghet@indiana.edu
Brian Guay[3] brian.guay@duke.edu
A. Karlapudy[1] akarlapu@umail.iu.edu
David Landy[1] dhlandy@gmail.com

[1] Psychological and Brain Sciences, Indiana University, Bloomington, Indiana
[2] School of Public and Environmental Affairs, Indiana University, Bloomington, Indiana
[3] Political Science, Duke University, Durham, North Carolina

Abstract

Large-scale societies are impossible to perceive directly. Unsurprisingly, lay demographic estimates are wildly inaccurate. How should we interpret these errors? Most accounts assume these errors are evidence of topic-specific biases and prejudices. (e.g., "People overestimate immigration because immigrants threaten the status quo.") But this glosses over the distortions that are introduced whenever underlying perceptions are translated into explicit numerical estimates. For instance, estimates are typically hedged, or 'rescaled,' toward an expected value — a perfectly rational strategy when information is uncertain. We show that uncertainty-based rescaling accounts for most error in individual demographic estimates. Residual errors were not even always in the same direction; populations that appeared to have been over-estimated (e.g., Asian-Americans) now appear to be under-estimated. The amount of rescaling engaged in by an individual was proportional to their uncertainty (about politics or about numbers). Perceptions of society are surprisingly good; the psychophysics of estimation gets in the way.

Keywords: numerical cognition; Bayesian estimation; confidence; demographics; psychophysics;

Introduction

People appear to have massively warped perceptions of the macrostructure of society. They overestimate the size of minority groups potentially seen as threatening to the status quo, but underestimate the size of dominant populations (Kuklinski et al, 2000; Wong, 2007; Sigelman and Niemi, 2001; Lawrence and Sides, 2014). For instance, U.S. residents estimate that more than 15% of the national population is Muslim; the correct percentage is a mere 1% (Ipsos Social Research Institute, 2014). By contrast, they estimate that Christians are approximately 60% of the national population; the correct percentage is over 70%. The fact that we struggle to estimate the makeup of our own societies is perhaps unsurprising: Large-scale societies are distributed across space and time in ways that make them impossible to perceive directly. It is no small wonder that we have any sense at all of our society's demographic structure.

These errors, however, have been taken as evidence of widespread public ignorance, misinformation, and prejudice (Kuklinski et al, 2000; Gilens, 2001; Lupia, 2015; Lawrence and Sides, 2014; Lodge and Taber, 2013). Most accounts of these errors have been piecemeal and topic-specific, invoking targeted mechanisms that range from personal prejudice to media bias. Some explanations posit that members of majority groups overestimate minority or out-group prevalence through a combination of ignorance, bias, and fear (Kuklinski et al, 2000; Lawrence and Sides, 2014; Lodge and Taber, 2013; Sigelman and Niemi, 2001). For instance, overestimation of immigrants is associated with negative attitudes toward immigration (Sides and Citrin, 2007). To explain why minorities overestimate other minorities, authors have invoked the greater presence of minorities in social networks of minorities (Wong, 2007). In addition, media-based accounts note that people may infer a demographic group's size from the amount of associated media coverage, but some groups are overrepresented in the media relative to their actual size. For instance, there has been a recent explosion in LGBTQ coverage by US media. If media coverage of LGBTQ people is disproportionate to their prevalence of, then this might cause overestimates.

These explanations are specific to particular questions (e.g., xenophobia affects estimates of immigrants) or respondents (e.g., minorities know more minorities). But when we zoom out to consider a wide range of demographic issues, a unified pattern appears to govern estimates across the board: Aggregate estimates of small populations are reliably overestimated, while aggregate estimates of large populations are reliably underestimated. While errors in demographic estimates have been explained by a heterogeneous set of distinct population- and issue-specific biases, the presence of this systematic pattern invites a unified explanation (Landy, Guay, & Marghetis, 2017).

Aggregate errors and uncertainty-based rescaling

We have proposed an alternative explanation of this general pattern of overestimation and underestimation: people are acting sensibly under uncertainty by combining their perceptions with domain-general prior expectations (Landy, Guay, & Marghetis, 2017). On this account, explicit numerical estimates are not a direct, unfiltered reflection of an individual's underlying beliefs or perceptions (Fechner, 1860; Stevens, 1957). When we estimate a demographic proportion, we must transform our internal perceptions into whatever format is demanded by the task. But this process of translation is unlikely to be linear — and, in fact, there is a mountain of evidence that it is not (Tversky and Kahneman, 1992; Gonzalez and Wu, 1999; Huttenlocher et al, 1991; Hollingworth, 1910; etc.).

An immediate corollary is that demographic estimates are not a direct window onto an individual's perceptions of their social world. When demographic *estimates* are systematically wrong, this does not necessarily reflect systematic biases in people's underlying *perceptions*. On the contrary: The pattern of over- and under-estimation that is typically reported for demographic estimation might reflect the way information is translated psychologically into an explicit numerical response, rather than systematic biases in what people actually believe. Said otherwise, an individual with perfect information might *appear* to have systematic biases; it all depends on how they translate their underlying information into explicit numerical estimates.

Specifically, Landy et al (2017) proposed that much of the error in demographic estimates reflects two processes that are common to all proportion estimations, not just demographic ones:

First, when thinking about proportions, people represent them mentally as log-odds — an unbounded scale — rather than as raw proportions ranging from 0 to 1 (Shepard, 1981; Stevens, 1957; Tversky & Kahneman, 1992; etc.).

Second, individuals should follow the generically Bayesian strategy of combining new information with their prior expectations, with new information weighted more when people are more confident in it. In practice, this means that estimates for a particular demographic should reflect a combination of two things: an individual's underlying perception of that demographic, and their expectation of what a 'typical' demographic proportion would be. We refer to this process as '***domain-general uncertainty-based rescaling.'*** *Rescaling,* because, in generating an explicit numerical estimate, the net result is that individual's perceptions are rescaled toward their prior expectations; *uncertainty-based*, because the amount of rescaling should depend on the amount of uncertainty; and *domain-general*, because this process makes no assumptions about the particular population being estimated (e.g., LGBTQ vs. Muslim vs. White Americans). When formalized as a mathematical model of an individual's demographic estimations, this requires only two parameters: their prior expectation of a 'typical' demographic proportion, and the amount they rescale their perceptions toward this expectation.

Taken together, these two assumptions suffice to generate the S-shaped curve that is typical of demographic estimations: systematic overestimation of smaller groups and underestimation of larger groups. We have reported previously that this model of demographic estimation can account for much of the error in aggregate estimates from a large, multinational survey (Landy et al, 2017). There, we found that the pattern of over- and under-estimation in average demographic estimates follows the pattern predicted by domain-general psychophysical rescaling. We concluded that much-ballyhooed errors in demographic estimates have been over-interpreted. Those errors have been taken as evidence for topic-by-topic bias, prejudice, and misinformation (e.g., Muslims are overestimated because of Islamaphobia). Instead, aggregate estimates are close to what we would expect if individuals had near-perfect, unbiased information about the macrostructure of society — but, adopting a rational Bayesian strategy, they hedged their explicit estimates toward a more typical value.

The present study

Two critical questions remain unaddressed. The first is whether uncertainty-based rescaling accounts for errors in *individual* estimates, not just aggregates. Our proposal is an account of *individual* psychological processing — but so far we have only analyzed aggregate estimates that average thousands of individual responses (Landy et al, 2017). To address this question, we conducted an online study. Our approach was to generate estimates from perfect (i.e., unbiased) perception of the demographic structure of society, but represented and rescaled as described above. By comparing these predictions to actual responses, we can estimate error that remains to be explained *after accounting for the psychological processes involved in generating proportion estimates.* Said otherwise, by assuming that perceptions were unbiased, we can distinguish between misestimates that are in line with unbiased perception, from those that suggest biased perception.

The second question is whether this rescaling truly is 'uncertainty-based.' If so, then individual uncertainty should predict individual differences in the amount of rescaling (Tversky and Kahneman, 1992; Gonzalez and Wu, 1999; Huttenlocher et al, 1991; Hollingworth, 1910)—whether that uncertainty is domain-specific (Fennell and Baddeley, 2012) or due to domain-general innumeracy (Petrova, Pligt, and Garcia-Retamero, 2014). While an individual's sense of certainty for a particular question likely integrates a variety of sources of information, here we measured two: political knowledge and numeracy. If rescaling is a reflection of uncertainty, then it should be more pronounced in less certain individuals— whether because they know little about politics, or because they are uncertain about numbers.

Methods

Participants

Participants reporting to be U.S. residents and citizens were recruited through Amazon Mechanical Turk, an online labor market (Buhrmester, Kwang, and Gosling, 2011). According to criteria established during piloting (an independent sample of $N = 40$), participants were eliminated when they either left some entries blank (4 participants) or reported percentages less than 0% or greater than 100% (9 participants). Recruitment continued until we reached our target sample size ($N = 400$).

Demographic Proportion Estimation

A range of demographic populations ($N = 24$; Table 1 at the end of this manuscript) were selected to probe participants' knowledge of the demographic structure of the United States. Participants had to estimate the proportion of the US population that belonged to each population. The size of these populations had true "benchmark" values derived from high-quality data, sourced from US Census, or established consulting companies such as Gallup Inc. and the Pew Research Center.

Items spanned the range from 0 to 1, with over-representation at both extremes — close to 0 and close to 1 — to better estimate the S-shaped curve that is typical of demographic estimation (Landy et al, 2017). We included demographic populations that have been proposed to elicit topic-specific bias (e.g., Islamophobia leading to overestimation of Muslims). Past work has focused almost exclusively on these 'bias-eliciting' populations. In addition, we included a range of items that, *a priori*, should not elicit systematic errors, *if misestimations reflect the kind of topic-specific biases that have been invoked in past work*. Examples include the US population that is aged between 0 to 94 years old, or that lives east of the Mississippi river.

Procedure

Participants first answered the *Demographic Proportion Estimation* questions (ordered randomly by-participant). Participants were instructed to enter their answers as numerical percentages with as many decimal places as they deemed appropriate. Explicit encouragement to use decimals was included to discourage excessive rounding.

These were followed by *Personal Characteristic* questions that probed participants' socio-demographic characteristics, political knowledge (Delli Carpini and Keeter, 1996), and numeracy (Cokely et al, 2012). These were included to investigate sources of individual variability in rescaling. Socio-demographic items included: age, gender, political party identification, ethnic background, education, income, and level of political activity. General political knowledge was measured using the five-item Political Knowledge battery (Delli Carpini and Keeter, 1996), known to correlate strongly with a larger political knowledge battery. Numerical understanding of risk was measured using the four-item multiple-choice version of the Berlin Numeracy Test (Cokely et al, 2012).

Finally, an instructional manipulation check was included to ensure that participants were paying attention. In this multiple-choice question, participants were asked to select which word appeared in a sentence on the current page. Roughly 90% of participants answered correctly. The pattern of results was identical when we excluded participants who failed this check.

No other measures were collected.

Analysis

We assumed participants had unbiased information about each demographic group. That is, we predicted how participants *would* respond if they had perfect perceptions of the demographic structure of society, but engaged in uncertainty-based rescaling when generating their estimates.

We first converted *estimated* and *true* proportions to odds and then log-transformed them. For a demographic proportion p, that gives us:

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

To avoid infinite values, values of 0 and 1 were recoded to .001 and .999. We then assumed that, to generate an explicit numerical estimate, participants rescale this value toward a more typical value, δ (i.e., their prior expectation). This was formalized as a linear interpolation (in log-odds space) of the underlying perception, r_p, and the expected value, δ:

(Eq. 2) $$\psi'(r_p) = \gamma r_p + (1-\gamma)\delta$$

Combining equations 1 and 2 gives us a model of the psychological process by which implicit information about a proportion, p, is transformed into an explicit numerical estimate of that proportion (cf. Gonzalez & Wu, 1999):

(Eq. 3) $$\psi(p) = \frac{\delta^{(1-\gamma)} p^{\gamma}}{\delta^{(1-\gamma)} p^{\gamma} + (1-p)^{\gamma}}$$

This model has two parameters: δ, the prior expectation in log-odds space; and γ, the rescaling parameter, which is equal to 1 when an individual gives no weight to their prior expectation, and gets closer to 0 as participants give less weight to their own perceptions and more weight to their prior expectation.

We implemented Eq. 2 as a linear mixed effects model, with by-participant random intercepts and slopes:

$$\textit{log-odds(estimate)} \sim \beta_0 + \gamma_1\, \textit{log-odds(actual)} + (\beta_{0j} + \gamma_{1j}\, \textit{log-odds(actual)} \mid \textit{subject})$$

The key coefficients in this model are the fixed slope term, γ_1, an estimate of overall rescaling; and the random by-subject slopes, γ_{1j}, an estimate of how much more or less subject j rescaled compared to the rest of the population. For reasons of space, we do not analyze participants' prior expected value (i.e., $\delta = \beta_0/(1-\gamma_1)$).

Results

Every participant produced estimates that were significantly incorrect (absolute error: M = 13.8 percentage points, 95% CI [13.2, 14.3], all ts > 2.6, all ps < .02).

Looking at individual items, every demographic group that made up less than half the US population was overestimated (14/14, error: M = +7.4 percentage points, 95% CI [4.6, 10.1], t_{13} = 5.9, p < .001), while every demographic group that made up more than half the US population was underestimated (10/10; error: M = –15.1 percentage points, 95% CI [–17.6, –12.7], t_9 = –14.2, p < .001), giving the S-shaped curve that is typical of proportion estimation in general and demographic estimation in particular (Figure 1A). Said otherwise, raw errors varied systematically with the actual size of the demographic group, such that smaller proportions were systematically overestimated, while larger proportions were systematically underestimated (Figure 1B).

Does rescaling explain demographic misestimation?

We next investigated whether this pattern of errors could be explained by our account of demographic proportion estimation. We compared participants' estimates to the predictions of the model. The model cut root-mean-squared error (RMSE) in half, compared to RMSE calculated relative to the true values (RMSE relative to model predictions: M = 0.017, 95% CI [0.016, 0.019]; RMSE relative to true values: M = 0.043, 95% CI [0.029, 0.047]), a highly significant reduction in error (t_{399} = 13.4, p < .001). In short, assuming that individuals engage in uncertainty-based rescaling can explain much of their apparent 'misperception' of demographic proportions.

Moreover, the aggregate pattern of over- and underestimation appeared to be driven largely by systematic rescaling of proportions by individuals (Fig. 2). We estimated the size of this relationship with a linear mixed effects model of signed errors, with a fixed effect for the true proportion, and random intercepts and slopes by participants. Before accounting for rescaling, there was a large and reliable relation between the *actual* size of the demographic group and the signed error, b = -0.29 ± 0.01 SEM, p < .001. After accounting for uncertainty-based rescaling, however, this relation was an order of magnitude smaller, although still significant, b = -0.05 ± 0.004 SEM, p < .001. The systematic pattern of over- and under-estimation that characterizes demographic estimations, therefore, can be explained almost entirely by general psychological processes involved in estimating proportions.

Individual uncertainty predicts amount of rescaling

Some people appeared to hedge their estimates toward a typical value a lot, while others appeared to hedge hardly at all. We estimated the amount of rescaling performed by each individual participant, using the random effects from the mixed effects model of demographic estimation. When this measure of rescaling is closer to 1, an individual gives more weight to their own underlying perception and largely ignores their prior expectation; as it approaches 0, an individual gives increasing weight to their prior expectation (i.e., increased rescaling). We analyzed these individual differences in rescaling with a multiple linear regression. This model included predictors for all socio-demographic measures. Critically, we also included predictors for our two measures of uncertainty: political knowledge and numeracy.

As predicted, rescaling was greater in less-certain individuals. Worse political knowledge and numeracy were both associated with more rescaling (political knowledge: β = 0.04 ± 0.01 SE, p < .0001; numeracy: β = 0.04 ± 0.01 SE, p < .0001). None of the socio-demographic predictors were associated significantly with rescaling (all ps > .15, except for income, p = .08). Thus, as predicted by general Bayesian considerations, rescaling appears to reflect individual variability in certainty, whether domain-specific (i.e., political knowledge) or due to general innumeracy.

Residual error after accounting for rescaling

Finally, we investigated the residual error after accounting for uncertainty-based rescaling. Since rescaling is a domain-general feature of proportion estimation, people will generate estimates that reflect rescaling, *even if they have*

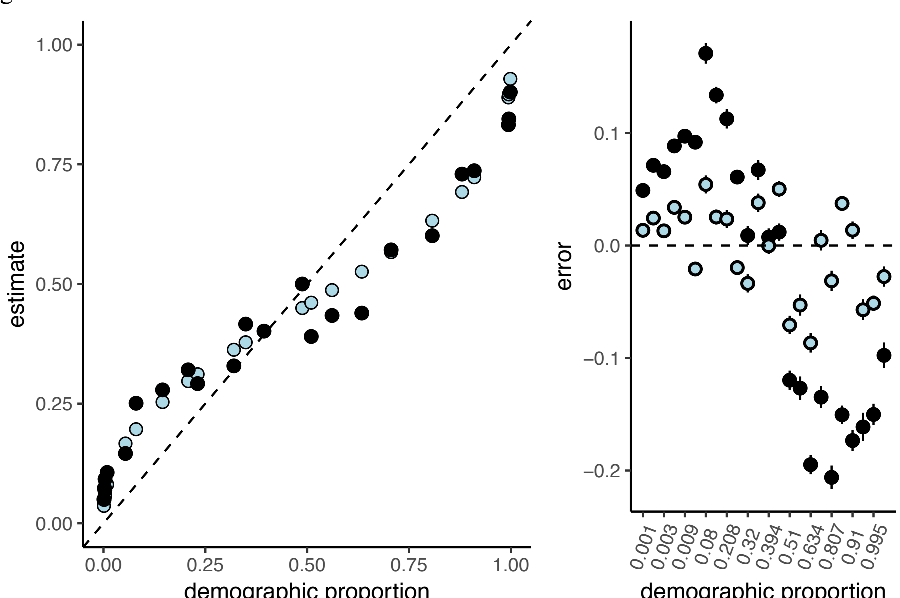

Figure 1. Systematic errors in demographic estimates were explained by uncertainty-based rescaling. (**Left**) Mean estimates (black) deviated systematically from the correct value (dashed diagonal line), but these deviations followed the pattern predicted by uncertainty-based rescaling (blue). (**Right**) Errors in estimates (black) varied systematically with the size of the demographic group (in increasing order along the x-axis). Small groups were overestimated; large groups were underestimated. This largely disappeared after accounting for rescaling (blue). Error bars indicate SEM, but are too small to display for most items.

perfect information about demographic proportions. For each trial, we thus calculated error relative to the demographic estimate that we would expect people to generate for a demographic group of that size (i.e., directional comparison between actual estimates and model predictions).

Accounting for uncertainty-based rescaling completely changed the pattern of errors (Table 1). A quarter of the items exhibited residual error that was in the opposite direction from pre-rescaling raw error (6/24).

For instance, according to the US census, 88% of American adults have high school diplomas. On average, respondents gave an estimate of 73%—a large apparent *underestimate*. But if participants were perfectly informed about the size of this group, but rescaled their information before generating explicit numerical estimates, then we should expect a mean estimate of 69%. This implies that perceptions of the high-school graduation rate are actually too *high*. Similarly for premarital sex: The raw estimate of 74% appears to massively underestimate its true prevalence of 91%. But if participants had perfect knowledge of premarital sex but rescaled this information in the way they appear to rescale all their perceptions, then we should expect a mean estimate of 72%. The actual mean estimate was higher, suggesting that perceptions of unwed sexual activity are perhaps a bit overheated.

Conversely, while Asian-Americans make up only 5% of the US population, they were estimated to make up 15% —a large overestimation. But once we take into account the fact that small proportions will, in general, be rescaled upwards when generating a numerical estimate, the direction of the error changes entirely. Specifically, if participants were to generate estimates for a group that they thought comprised 5% of the population — the true size of the Asian-American population — then we should expect a mean estimate of 17%. Relative to this, it appears that participants actually *underestimate* the size of the Asian-American population.

Discussion

We investigated whether individual misestimation of the demographic structure of society could be explained by a simple, domain-general model of proportion estimation. The model formalized two simple assumptions: proportions are represented mentally as log-odds; judgments reflect Bayesian rescaling of new information toward prior expectations. These assumptions sufficed to explain most error, and in particular accounted for the systematic over- and under-estimation that characterizes demographic estimates. Critically, the amount of rescaling was predicted by individual differences in uncertainty: Greater numeracy and political knowledge were associated with less rescaling.

Our ability to make sense of the macrostructure of society has both applied and theoretical importance. From an applied perspective, understanding how people do — and do *not* — misperceive society is a critical precursor to deciding how we can improve public decision making about critical issues, including which political policies to support. Indeed, while this S-shaped pattern of misjudgment is not new to cognitive scientists (e.g., Tversky & Kahneman, 1992; Gonzalez and Wu, 1999), its implications for political behavior and knowledge has not been appreciated (Landy et al, 2017). Theoretically, it provides a case study in numerical reasoning beyond the human-scale typically studied in cognitive science (e.g., estimating the number of dots on a screen; mapping small numbers to a line; etc.)

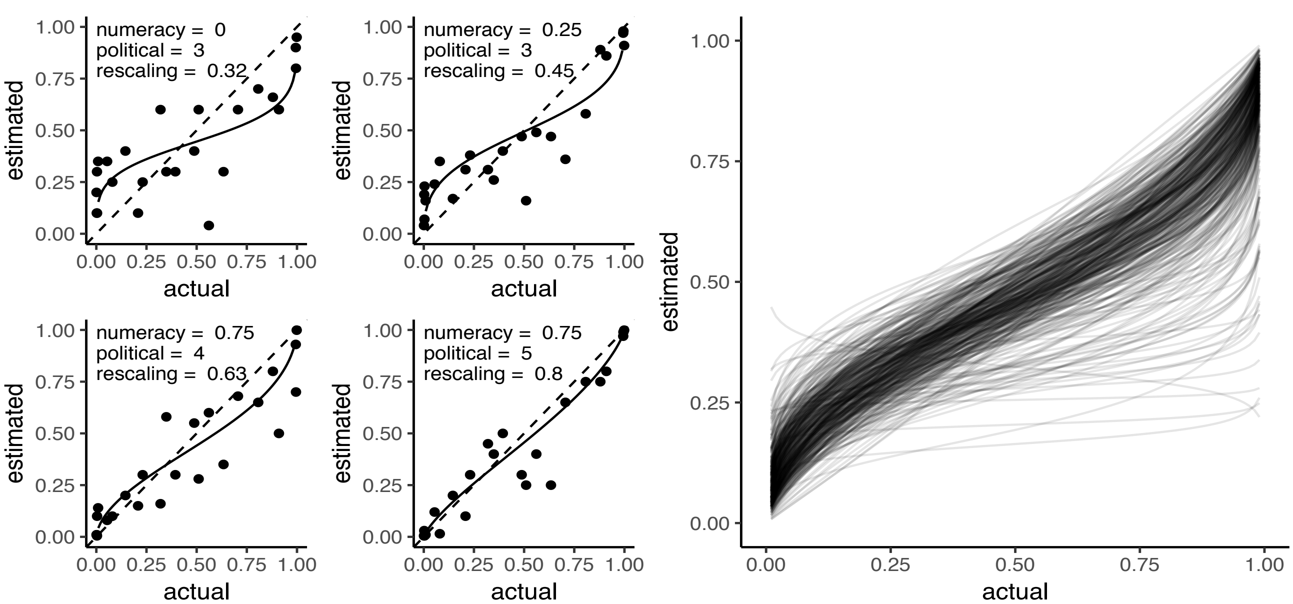

Figure 2. (**Left**) Sample individual estimates and model fits, illustrating how uncertainty related to rescaling. More uncertainty due to poor numeracy (worst = 0, best = 1) or poor political knowledge (worst = 0, best = 5) was associated with more rescaling (max rescaling = 0; no rescaling = 1). (Black dots = individual's estimates; black lines = predictions of rescaling model, assuming perfect underlying perceptions; dashed lines = predictions if estimates were direct reflections of underlying perceptions (i.e., without rescaling). (**Right**) Superimposed model fits for all participants (N = 400), revealing the systematic pattern of over- and underestimation that is typical of proportion estimation under uncertainty.

References

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk? *Persp on Psych Sci, 6,* 3-5

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin Numeracy Test. *JDM, 7,* 25-47.

Delli Carpini, M., & Keeter, S. (1996). *What the American public knows and why it matters.* New Haven: Yale UP

Fechner, G. T. (1860). *Elemente der Psychophysik.* Leipzig: Breitkopf & Härtel

Fennell J & Baddeley R (2012) Uncertainty plus prior equals rational bias. *Psych Review, 119,* 878-887.

Gilens, D. (2001). Political ignorance and collective policy preferences. *American Politi Sci. Rev., 95,* 379-396.

Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cog Psych, 38,* 129-166

Hollingworth, H. L. (1910). The central tendency of judgment. *J. Phil, Psych & Sci. Methods, 7*, 461-469.

Huttenlocher, D. J., Hedges, & L. V., Duncan, S (1991). Categories and particulars. *Psych Review, 98,* 352–376.

Kuklinski, J. F., Quirk, P. J., Jerit, J., Schwieder, D., & Rich, R-F. (2000). Misinformation and the currency of democratic citizenship. *J. of Politics, 62,* 790-816.

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

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

Lodge, M., & Taber, C. S. (2013). *The Rationalizing Voter*. (New York: Cambridge University Press)

Lupia, A. (2015). *Uninformed.* New York: Oxford UP.

Petrova, D., Pligt, J. & Garcia-Retamero, R. (2014). Feeling the numbers. *J. of Behav Decision Making, 27,* 191-199

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

Sides, J., & Citrin, J. (2007). European opinion about immigration. *British J. of Political Science 37*, 477-504.

Sigelman, L., Niemi, R-G. (2001). Innumeracy about minority populations. *Pub. Opinion Quart., 65,* 86-94.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review, 64,* 153-181.

Tversky, A., Kahneman, D. (1992). Advances in prospect theory. *Journal of Risk & Uncertainty 5,* 297-323.

Wong, C. J. (2007). “Little” and “big” pictures in our heads. *Public Opinion Quarterly, 71,* 392.

Table 1. Mean errors before and after accounting for uncertainty-based rescaling.

Demographic subgroups (in USA)	True Value (%)	Raw Error	Error after rescaling	Change in sign of error?
Pakistani	0.1	4.9	1.4	
Japanese	0.2	7.1	2.4	
Colombian	0.3	6.6	1.3	
Aged 13 and older living with an HIV infection	0.4	8.9	3.4	
Muslim	0.9	9.7	2.5	
Asian	5.4	9.2	-2.1	*
Served in the Armed Forces	8.0	17	5.0	
Live below the poverty line	14.5	13.4	2.5	
Catholic	20.8	11.3	2.3	
Aged 15 and over who have never been married	32	0.9	-3.4	*
Aged 20 and over clinically classified as obese	34.9	6.7	3.8	
Hold at least a 2-year college degree	39.4	0.8	0.0	
Aged 15 and over and currently married	48.8	1.2	5.0	
Workers who make less than $30,000 per year	51	-12.0	-7.1	
Live east of the Mississippi River	56.1	-12.4	-5.3	
Adults who own homes	63.4	-19.4	-8.7	
Christian	70.6	-13.5	0.5	*
Live in urban areas	80.7	-20.6	-3.1	
Aged 25 and over and own a high school diploma	88	-15.0	3.7	*
Have had premarital sex	91	-17.3	1.4	*
Over 6 months old	99.4	-16.1	-5.7	
Homeowners who have full indoor plumbing	99.5	-15.0	-5.1	
Aged between 0 to 94 years old	99.9	-9.8	-2.8	