# Measuring Perceptions of Shares of Salient and Stereotypical Groups

## Online Appendix

Douglas J. Ahler* Gaurav Sood†

# 1 Study 1

Study 1 utilizes data from Ahler and Sood (2016). The results presented in SI I are also part of the appendix in Ahler and Sood (2016), and provided here only for convenience.

## 1.1 Study 1: MTurk Sample Demographics

*Doug is a post-doctoral fellow at the Center for the Study of Democratic Institutions, Vanderbilt University, and Assistant Professor of Political Science starting Fall 2017 at the Department of Political Science, Florida State University, doug.ahler@gmail.com.

†Gaurav can be reached at gsood07@gmail.com.

**Table 1:** Characteristics of the MTurk Sample

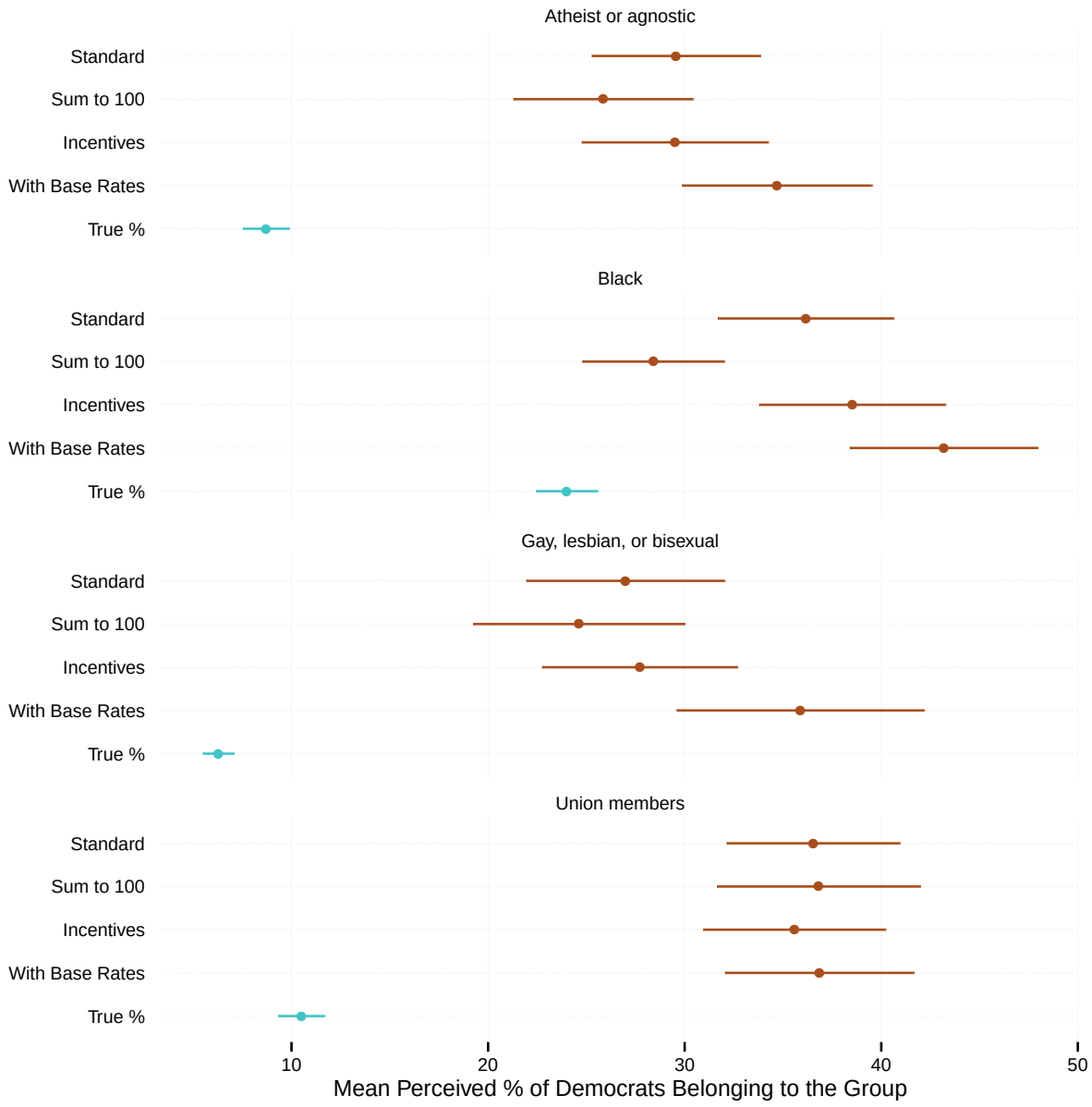
| | Sample | Population |
|---|---|---|
| **Partisanship** | | |
| Democratic (inc. leaners) | 58.9% | 49.0% |
| Republican (inc. leaners) | 22.3% | 39.0% |
| Non-leaning Independent | 18.9% | 11.9% |
| **Gender** | | |
| Female | 50.5% | 50.9% |
| Male | 49.5% | 49.1% |
| **Race/Ethnicity** | | |
| White/Caucasian | 76.2% | 63.7% |
| Black/African-American | 8.1% | 12.2% |
| Asian-American/Pacific Islander | 10.1% | 4.8% |
| Native American/Native Alaskan | 1.6% | 1.1% |
| Latino/Hispanic | 9.7% | 16.4% |
| **Education** | | |
| Less than high school | 0.5% | 8.9% |
| High school diploma (or equiv.) | 9.7% | 31.0% |
| Some college | 46.1% | 28.0% |
| 4-year degree | 34.8% | 18.0% |
| Advanced degree | 8.9% | 9.3% |
| **Age** | | |
| 18-39 | 79.1% | 39.1% |
| 40-64 | 19.1% | 43.7% |
| 65+ | 1.8% | 17.2% |

NOTE: Population estimates come from the 2010 US Census, except for partisanship, which comes from the 2012 ANES.

## 1.2 Study 1 Results by Item

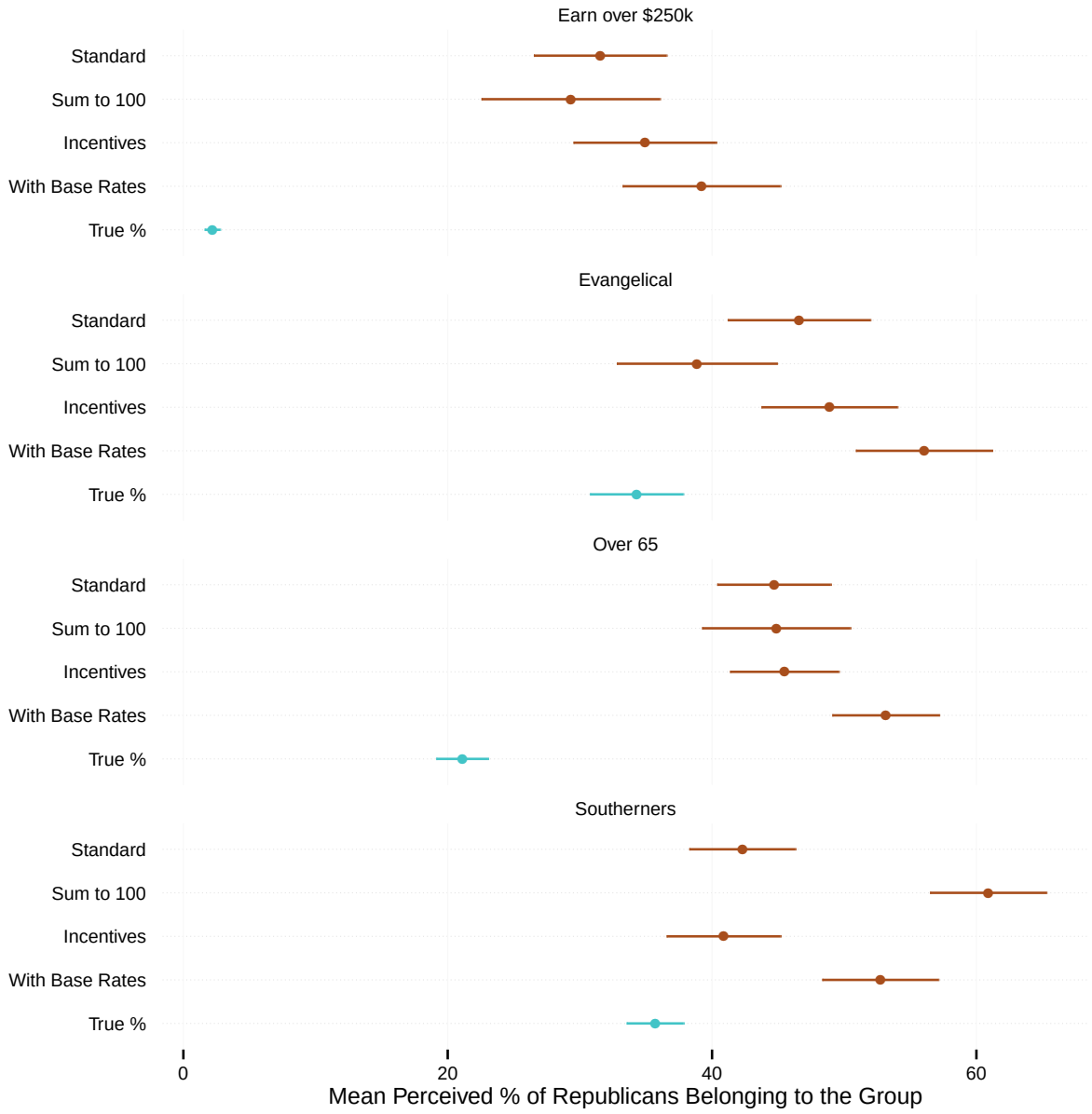
**Figure 1:** Expressive Responding, Innumeracy, and Poor Knowledge of Base Rates Fail to Explain Apparent Perceptual Errors

**(a)** Perceived Composition of Democratic Identifiers, by Experimental Condition




NOTE: 95% confidence intervals depicted.

**(b)** Perceived Composition of Republican Identifiers, by Experimental Condition




NOTE: 95% confidence intervals depicted.

## 1.3 Kolmogorov-Smirnov Tests

**Table 2:** Kolmogorov-Smirnov tests of Distributions

| | Incentives | | Sum-to-100 | | Base Rates | |
|---|---|---|---|---|---|---|
| Party-Group | D | p(Same dist.) | D | p(Same dist.) | D | p(Same dist.) |
| Dem.-Ath./ag. | 0.14 | 0.25 | 0.11 | 0.78 | 0.17 | 0.12 |
| Dem.-Black | 0.08 | 0.85 | 0.20 | 0.12 | 0.23 | 0.01* |
| Dem.-LGB | 0.10 | 0.64 | 0.19 | 0.64 | 0.19 | 0.06 |
| Dem.-Union | 0.09 | 0.80 | 0.11 | 0.8 | 0.09 | 0.74 |
| Rep.-$250k+ | 0.11 | 0.57 | 0.09 | 0.9 | 0.18 | 0.07 |
| Rep.-Evangelical | 0.12 | 0.48 | 0.21 | 0.08 | 0.20 | 0.03 |
| Rep.-Age 65+ | 0.07 | 0.96 | 0.49 | 0.00*** | 0.19 | 0.06 |
| Rep.-Southern | 0.07 | 0.93 | 0.07 | 0.93 | 0.19 | 0.05 |

NOTE: Statistics are based on K-S tests comparing the distribution of responses under the condition named at the top of each column to the distribution under the "standard estimation" condition. Asterisks denote statistical significance under the Benjamini and Hochberg (1995) method for controlling the false discovery rate. (Family defined as each set of eight K-S tests comparing responses under two distinct conditions.) * = $p < \alpha^*$ when $\alpha = 0.1$, ** = $p < \alpha^*$ when $\alpha = 0.05$, *** = $p < \alpha^*$ when $\alpha = 0.01$.

## 1.4 Affect Toward Groups is a Weak Predictor of Perceptions of Prevalence of Groups in Parties

Expressive benefits (rather than misperception) are a potential alternative explanation for the apparent misperceptions we find. Under this explanation, respondents dislike particular social groups and also dislike the out-party and, thus, declare the out-party to be excessively composed of groups they dislike. If expressive responding explained our findings, we would expect perceptual errors to be associated with how much a respondent likes a group. At the start of the extremity perceptions experiment we asked respondents to rate the eight groups related to party prototypes on a 101-point feeling thermometer. (We separated these feeling thermometers and the experimental content with a lengthy demographics and political information battery.) If dislike of the groups and the out-party explains our primary descriptive finding, we should find a significant correlation between respondents' feeling thermometer ratings for group $g$ and the reported perceptions of how prevalent $g$ is in party $p$. However, as the table below shows, we fail to find relationships that are distinguishable from zero for any of the eight group-party dyads. (Note that in this study we swapped one party-stereotypical group, "people over 65" for Republicans, for a counter-stereotypical group, "people under 35.")

**Table 3:** Feeling Thermometer Ratings for Groups Fail to Predict Perceptions of Group Prevalence

| | Reg. coefficient | Std. error | 95% conf. interval | $n$ |
|---|---|---|---|---|
| **Democratic Party Groups** | | | | |
| Blacks | -0.01 | 0.05 | [-0.11., 0.10] | 297 |
| Union members | -0.04 | 0.05 | [-0.15, 0.06] | 297 |
| Gay, lesbian, & bisexual | -0.07 | 0.05 | [-0.17, 0.03] | 297 |
| Atheist/Agnostic | -0.03 | 0.05 | [-0.13., 0.06] | 297 |
| **Democratic Party Groups** | | | | |
| The rich/earn over $250,000 | 0.01 | 0.04 | [-0.07., 0.10] | 659 |
| Evangelicals | -0.02 | 0.04 | [-0.09, 0.05] | 659 |
| Southerners | -0.03 | 0.03 | [-0.09, 0.04] | 659 |
| The young/people under 35 | 0.02 | 0.02 | [-0.03, 0.07] | 659 |

Note: The coefficient is from the regression of response to the question, "What percentage of supporters of party $p$ do you think are members of group $g$?" on feeling thermometer rating of $g$.

# 2 Study 2

## 2.1 MTurk Sample Demographics

**Table 4:** Characteristics of the MTurk Sample

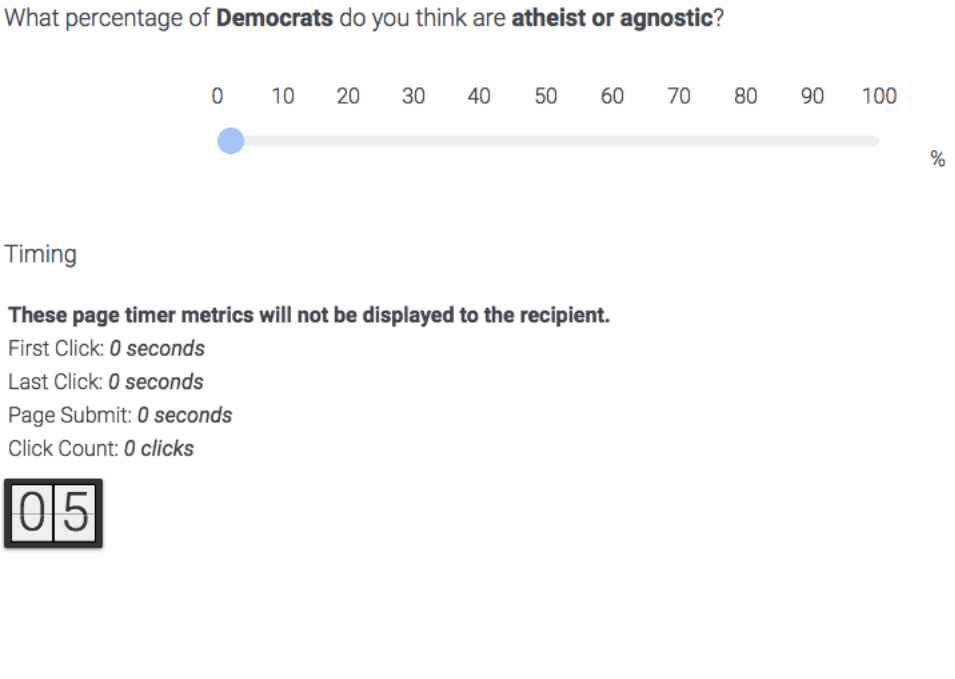
| | Sample | U.S. Population |
|---|---|---|
| **Partisanship** | | |
| Democratic (inc. leaners) | 56.0% | 49.0% |
| Republican (inc. leaners) | 28.0% | 39.0% |
| Non-leaning Independent | 16.0% | 11.9% |
| **Gender** | | |
| Female | 50.5% | 50.9% |
| Male | 49.5% | 49.1% |
| **Race/Ethnicity** | | |
| White/Caucasian | 78.0% | 63.7% |
| Black/African-American | 6.7% | 12.2% |
| Asian-American/Pacific Islander | 7.8% | 4.8% |
| Latino/Hispanic | 19.4% | 16.4% |
| **Age** | | |
| 18-39 | 66.8% | 39.1% |
| 40-64 | 30.9% | 43.7% |
| 65+ | 2.4% | 17.2% |

NOTE: Population estimates come from the 2010 US Census, except for partisanship, which comes from the 2012 ANES.

## 2.2 Depiction of Study 2 Treatments

**Figure 2:** Timing Experiment

**(a)** "Time Pressure" Condition




**(b)** "Time Requirement" Condition

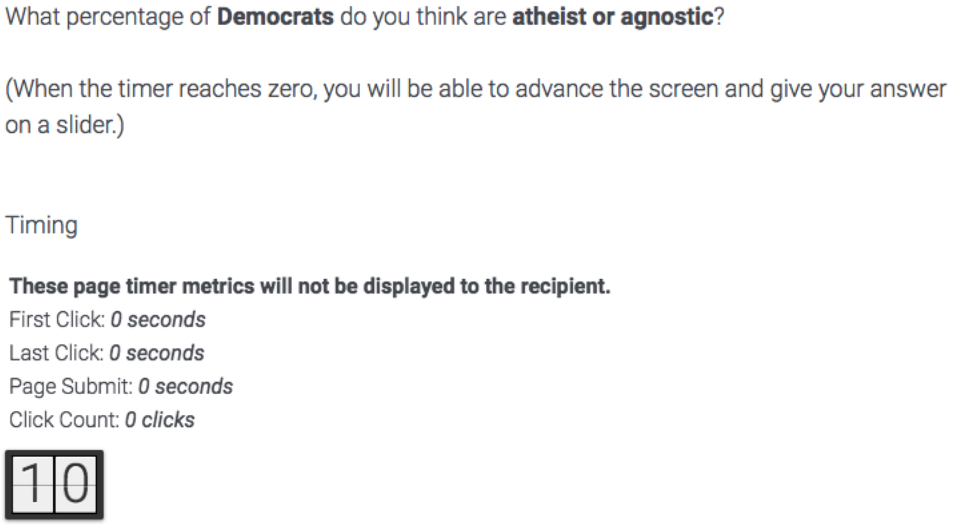
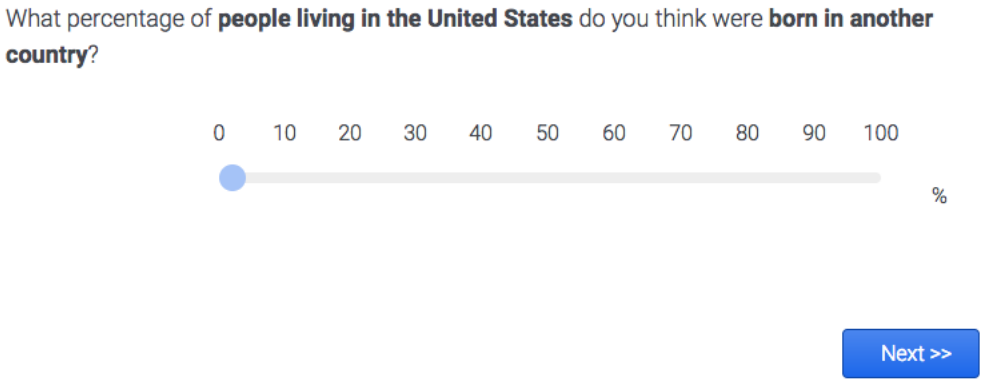


**Figure 3:** Wording Experiment

**(a)** "Percentage" Condition




**(b)** "Out of 100" Condition

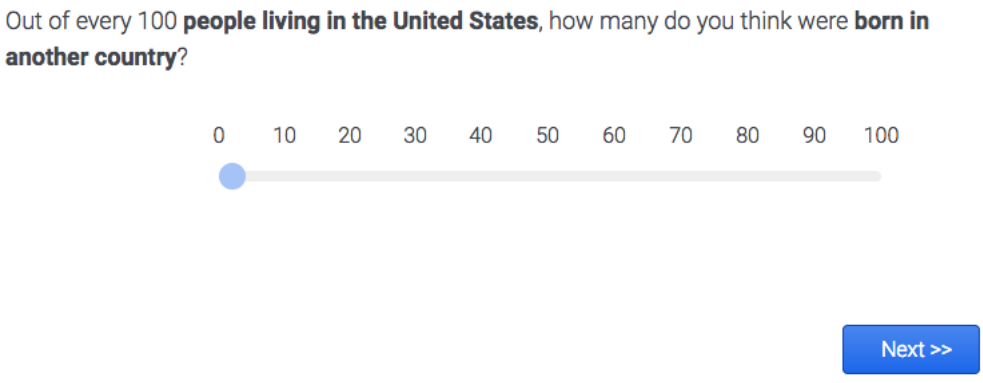
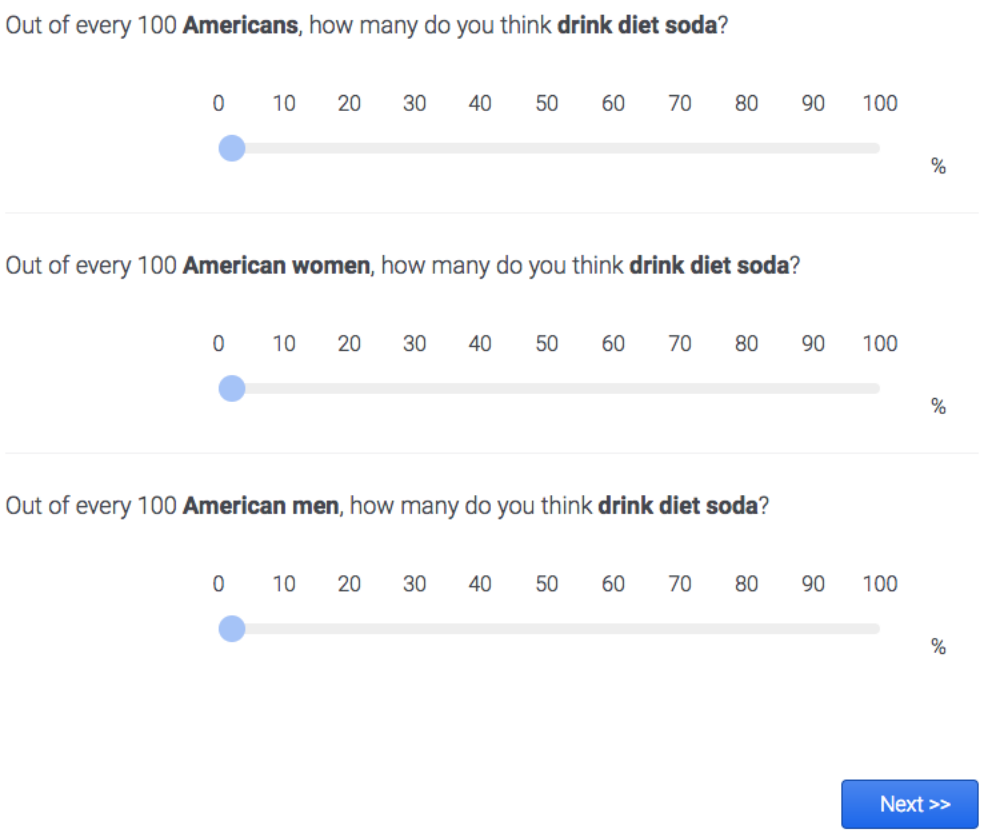


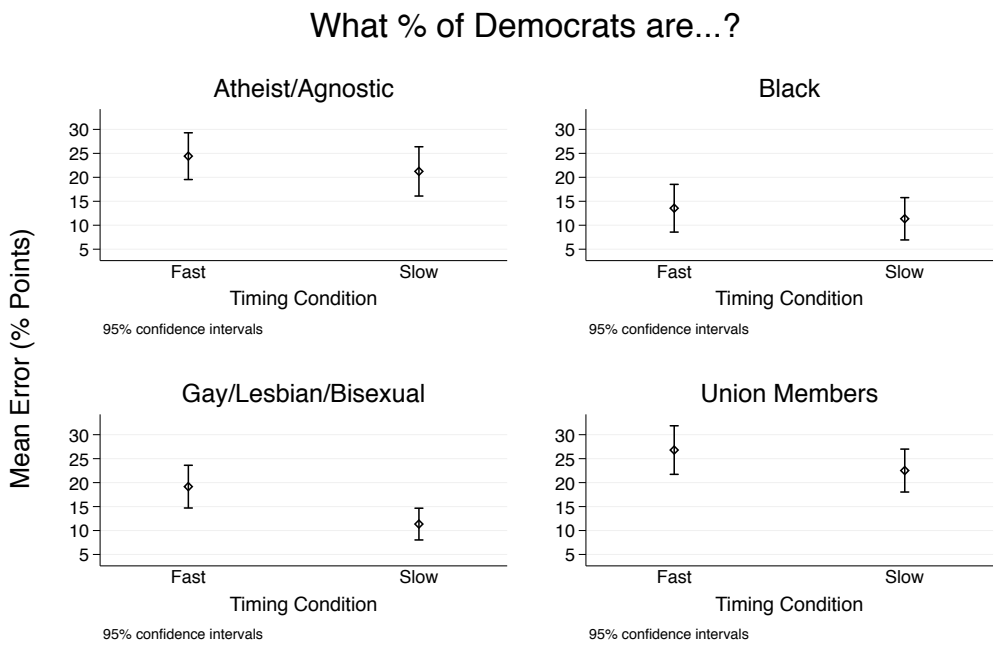
**Figure 4:** "Base rates and diet soda" experiment—"Control" condition omits top item



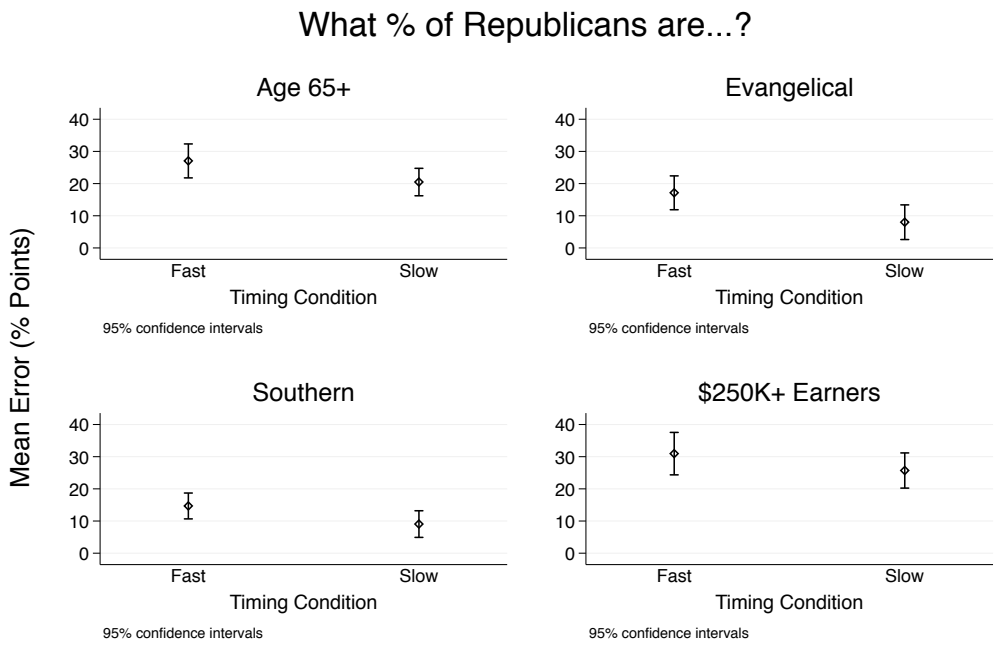
## 2.3 Timing Experiment: Average Estimates by Item and Condition

**Figure 5:** Average Responses (and Perceptual Error) Tend to Be Lower in the "Slow" Condition

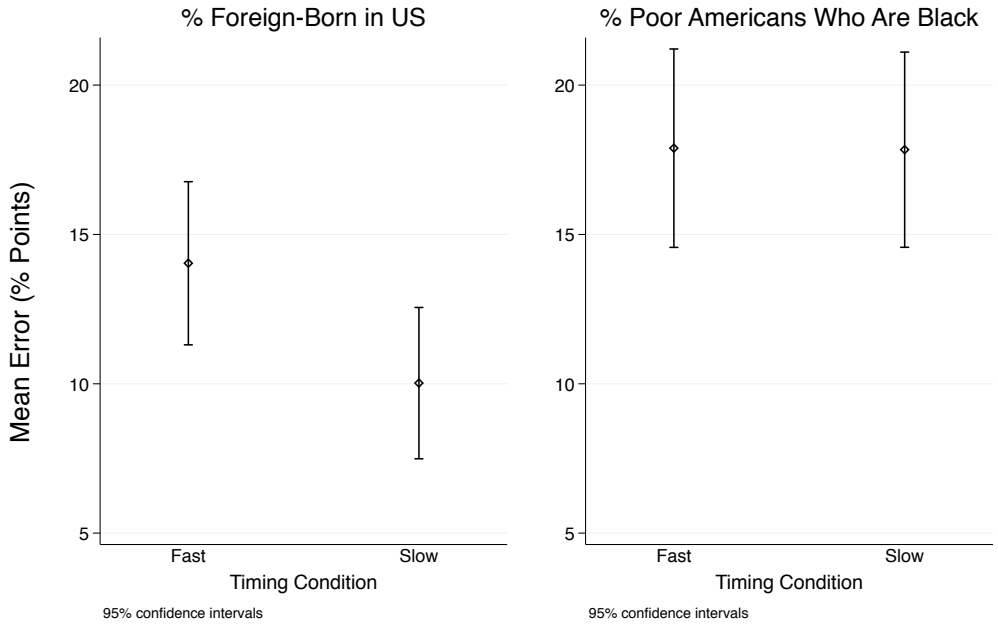
**(a)**



**(b)**



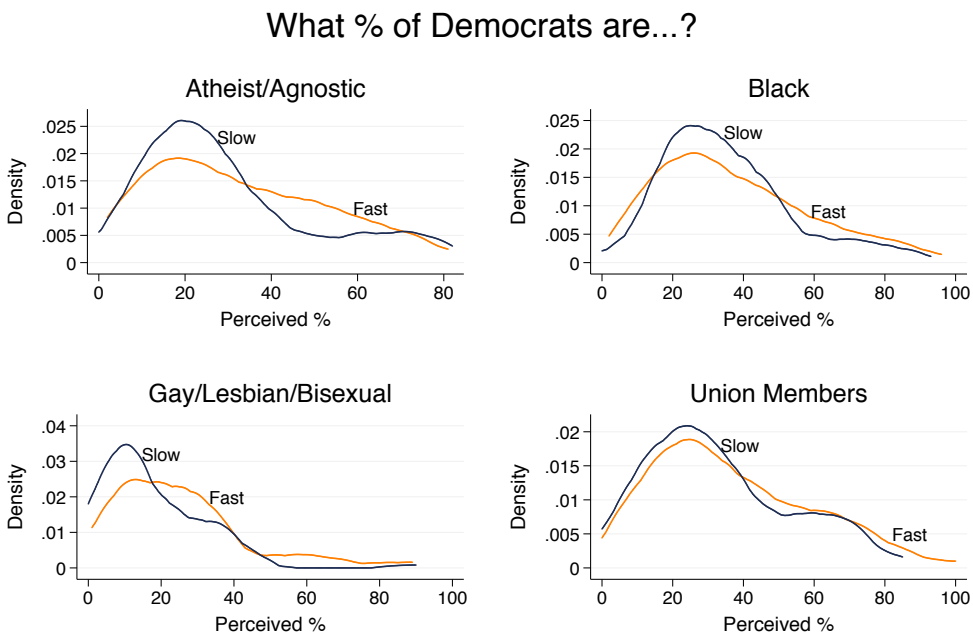

**(c)** Estimates for the Immigration and Race/Poverty Items



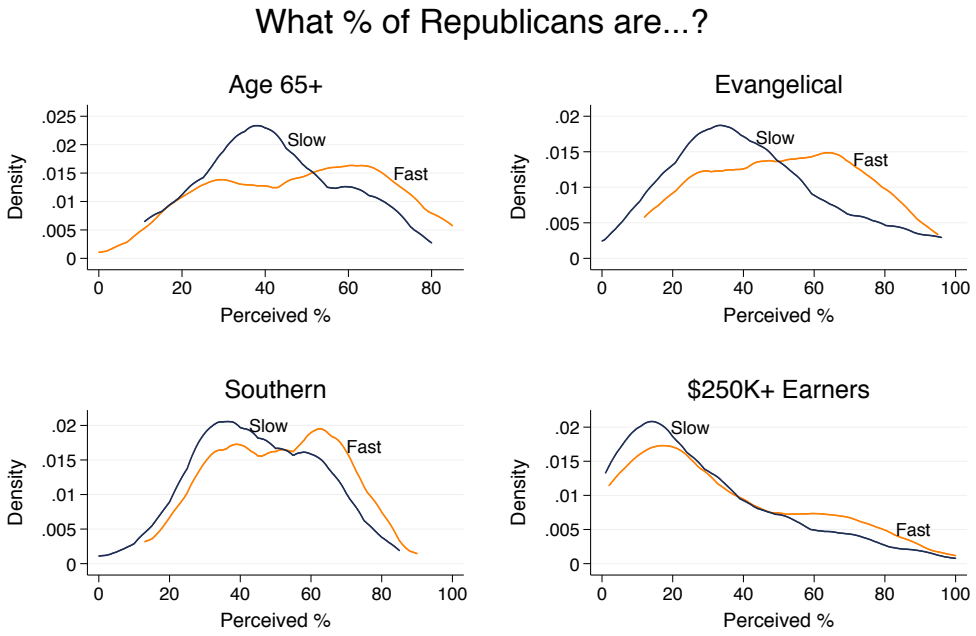
## 2.4 Timing Experiment: Density Plots by Item and Condition

**Figure 6:** Average Responses (and Perceptual Error) Tend to Be Lower in the "Slow" Condition

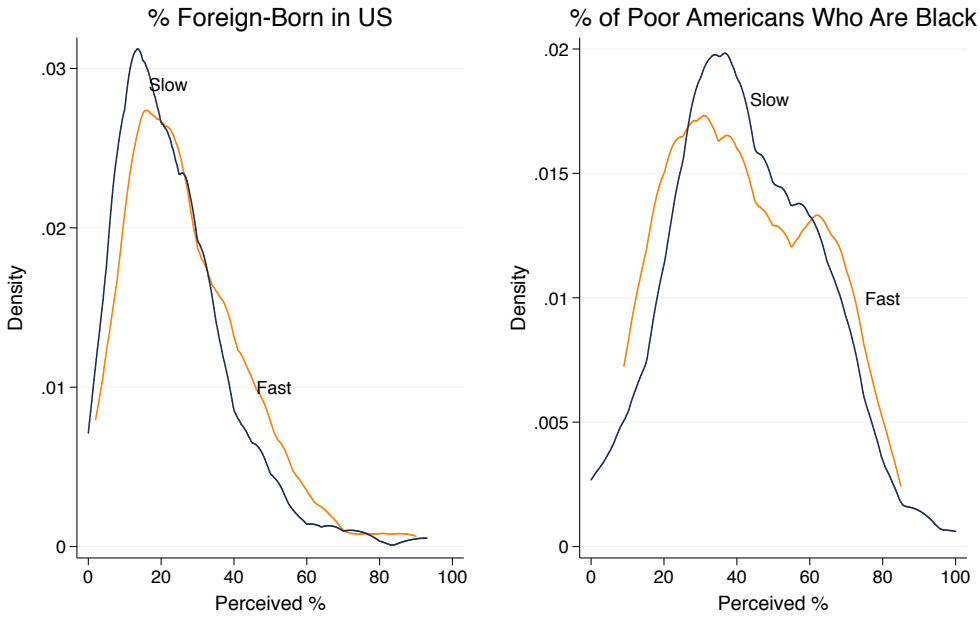
**(a)**



**(b)**




**(c)** Estimates for the Immigration and Race/Poverty Items



## 2.5 Timing Experment: A Statistical Test Showing No Evidence of Cheating

One concern in the timing experiment is that respondents, given more time to consider their responses, may use that time to consult outside references. This is undesirable because we want to gauge actual perceptions that people generally hold. To assess this, we calculate the percentage of respondents in each condition (slow, or "time requirement," versus fast, or "time pressure") responding with either integer adjacent to the correct response. (E.g., if someone said that 8 or 9 percent of Democrats are atheist or agnostic, they would be "correct" since the correct response is 8.7.) As the table below shows, few significant differences emerge on individual items. Only in the case of responses to the percentage of Republicans who are from the U.S. South are responses significantly more accurate ($p < .1$) in the "slow" condition. Aggregating across items (including item fixed effects and clustering standard errors by respondent) lends more power, and as the bottom line of the table indicates, we do detect a significant difference across conditions. 1.3% more responses are "on the nose" in the "slow" condition. However, interpreting this result ameliorates concerns. First, the difference is substantively very small. Even in the "slow" condition, very few responses are close-to-correct—just 3.5% (compared to 2.2% in the "fast" condition). This suggests that even if there is cheating, only 1% of respondents are doing so. Further, the increase is not consistent across all items—three items show a decrease in the "slow" condition, which provides evidence against systematic cheating. More likely, respondents simply became less biased in the "slow" condition as a result of having time to consider the numbers they called to mind.

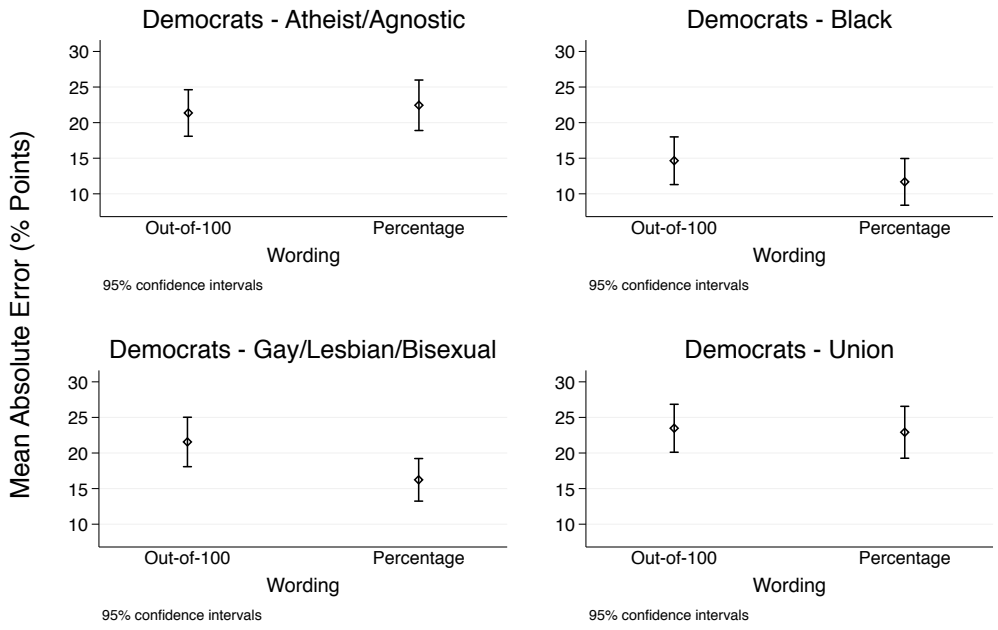
**Table 5:** Few Respondents Answer Perceptual Items Exactly Correct, and Differences are Small Across Timing Conditions

| | Fast | Slow | Difference | Std. Error of Dif. | $n$ | $P > \|t\|$ |
|---|---|---|---|---|---|---|
| % Dems. – Ath./Ag. | .02 | .01 | -.01 | .01 | 284 | .32 |
| % Dems. – Black | .04 | .06 | .03 | .03 | 284 | .27 |
| % Dems. – LGB | .02 | .06 | .04 | .02 | 284 | .13 |
| % Dems. – Union | .02 | .04 | .03 | .02 | 284 | .48 |
| % Reps. – Age 65+ | .01 | 0 | -.01 | .01 | 284 | .32 |
| % Reps. – Evang. | .02 | .03 | .01 | .02 | 284 | .70 |
| % Reps. – Southern | .01 | .05 | .04 | .02 | 284 | .09 |
| % Reps. – $250K+ | .02 | .01 | -.01 | .01 | 284 | .32 |
| % U.S. foreign-born | .02 | .05 | .03 | .02 | 284 | .20 |
| % U.S. poor – Black | .04 | .06 | .02 | .02 | 284 | .40 |
| Across all items | .02 | .04 | .01 | .01 | 2840 | .05 |

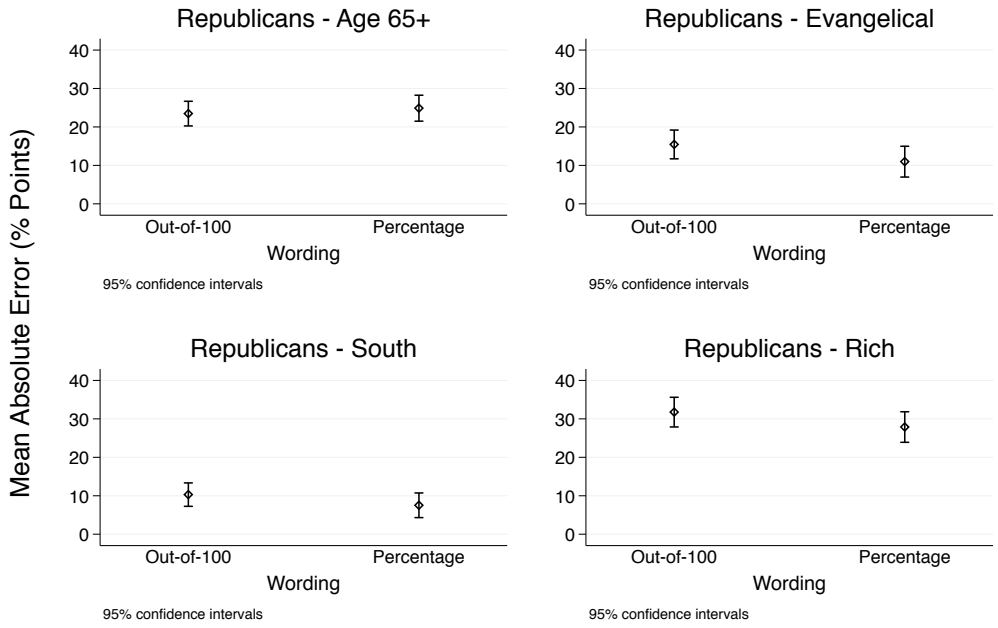
## 2.6 Wording Experiment: Average Estimates by Item and Condition

**Figure 7:** Little Apparent Relationship Between Perceptual Error and "Out of 100" or "What Percentage" Wording

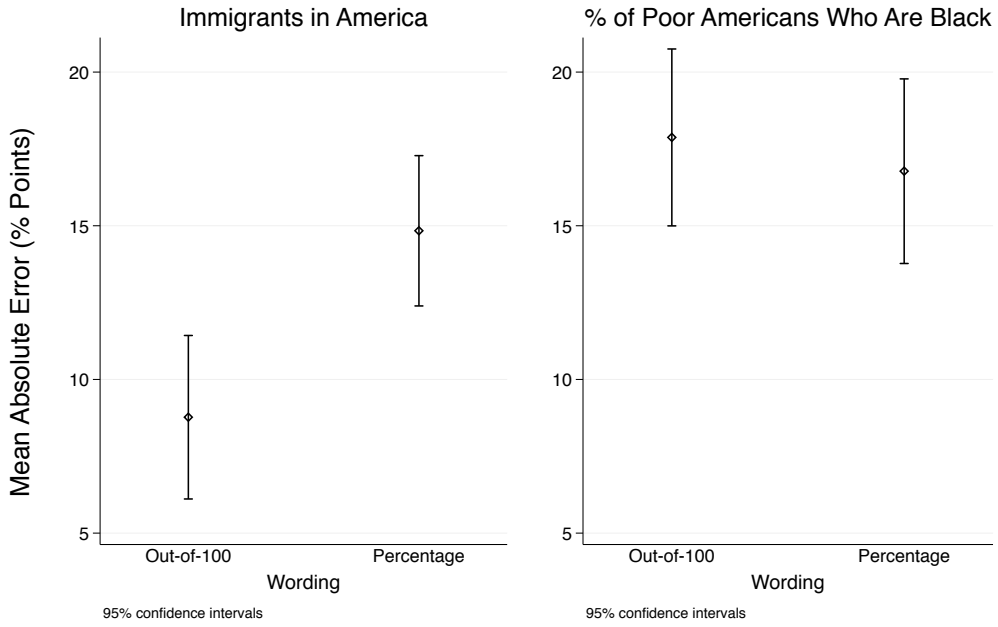
**(a)**



**(b)**




**(c)** Estimates for the Immigration and Race/Poverty Items



## 2.7 Perceptions of Base Rates of Party-Stereotypical Groups in the Population at Large

In the Study 1's `standard estimation` condition, after measuring party-specific perceptions, we asked respondents to estimate the percentage of the US adult population that belongs to a randomly-assigned subset of the eight party-stereotypical groups. As the table below shows, consistent with previous work (e.g., Wong 2007), respondents tend to overestimate the prevalence of these groups. However, misperceptions do appear to be party-specific: perceptions of the prevalence of these groups in their associated parties are significantly higher than those for the population writ large. And, importantly, the substantive difference between these party-specific and base rate perceptions tend to be quite large.

**Table 6:** Comparison of Party-Specific Perceptions to Perceptions of Population Base Rates of Party-Stereotypical Groups

| Group | Mean Perceived Base Rate | Mean Perceived Party Rate | Difference |
|---|---|---|---|
| Southerners | 32.74% | 41.94% | -9.20** |
| Over 65 | 30.36% | 46.54% | -16.18*** |
| Evangelical | 35.5% | 49.98% | -14.48*** |
| Earning Over $250K | 11.4% | 28.6% | -17.19*** |
| Black | 31.38% | 35.96% | -4.58 |
| Atheists or Agnostics | 22.93% | 28.04% | -5.11+ |
| Union Members | 25.74% | 33.52% | -7.78** |
| LGBT | 14.86% | 27.33% | -12.47** |

+ $p < .1$, *$p < .05$, **$p < .01$, ***$p < .001$

### Difference-in-Differences Estimator

To better understand misperceptions about subpopulation composition, we can compute a difference-in-differences:

$$(p(\text{group}|\text{representative party})_P - p(\text{group}|\text{representative party})_A) - (p(\text{group})_P - p(\text{group})_A)$$

where $P$ indexes perceived quantities and $A$ indexes actual quantities. While the results in the above table confirm that base rates alone do not drive these perceptions, and that they reflect something about party stereotypes, the D/D estimator can tell us whether one of those factors—misperceived base rates or party stereotyping—more heavily influences the apparent beliefs we observe. If the D/D estimate is significantly negative, then the misperception is more a function of base rate error, in the aggregate. By contrast, if the D/D estimate is significantly positive, then the base rate error contributes relatively little to the misperception in the aggregate. The table below presents the results, showing that base rate error contributes less than apparent party stereotyping for five of the eight party-group dyads, and contributes more for none of them.

Note that the quantity being estimated (and presented in the table) is the raw error of the party-specific perception minus the raw error of the base rate perception.

**Table 7:** Difference-in-Differences Estimates: Do Misperceptions of Group Base Rates Contribute Significantly More/Less Than Beliefs About Parties to the Misperceptions We Observe?

| Party-Group Dyad | D/D Estimate |
|---|---|
| Republicans—Southerners | 7.04 ** |
| Over 65 | 7.88 ** |
| Republicans—Evangelical | -0.82 |
| Republicans—Earning Over $250K 16.99 *** | |
| Democrats—Black | 4.56 |
| Democrats—Atheists or Agnostics | 2.10 |
| Democrats—Union Members | 8.58 *** |
| Democrats—LGBT | 9.97 ** |

+ $p < .1$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

## 2.8 Alternative Specifications for Respondent Error

In the paper, the main dependent measure is *perceptual bias*,by which we mean average signed (or directional) error in the respondents' reports across items. We can also assess the effect of the treatments on absolute error, average absolute error in respondentsâĂŹ reports across items, and percentage bias, average percentage by which respondents err (signed error) across items. As Table 8 shows, results are consistent across these different specifications of respondent perceptual error.

**Table 8:** Results are consistent across different specifications of the dependent measure

| | Timing Experiment | | | Wording Experiment | | |
|---|---|---|---|---|---|---|
| | Perceptual bias | Percentage bias | Absolute error | Perceptual bias | Percentage bias | Absolute error |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Timing: slow | -4.34*** | -48.1** | -3.58*** | | | |
| | (1.43) | (20.2) | (1.24) | | | |
| Wording: percentage | | | | -1.25 | -23.7 | -1.10 |
| | | | | (1.26) | (18.0) | (1.05) |
| Constant | 19.7 | 235.0 | 22.0 | 18.9 | 264.6 | 22.2 |
| | (1.12) | (15.5) | (1.24) | (1.03) | (13.8) | (0.86) |
| $R^2$ | 0.08 | 0.47 | 0.07 | 0.08 | 0.49 | 0.05 |
| $n$ perceptions | 1696 | 1696 | 1696 | 3119 | 3119 | 3119 |
| $n$ respondents | 284 | 284 | 284 | 312 | 312 | 312 |

# References

Ahler, Douglas J and Gaurav Sood. 2016. "The Parties in our Heads: Misperceptions About Party Composition and Their Consequences.".

Benjamini, Yoav and Yosef Hochberg. 1995. "Controlling the false discovery rate: a practical and powerful approach to multiple testing." *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 289–300.

Wong, Cara 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.