# ONLINE APPENDIX TO
# "PARTISAN BIAS IN FACTUAL BELIEFS ABOUT POLITICS"

John G. Bullock

Alan S. Gerber

Seth J. Hill

Gregory A. Huber

This online appendix has seven parts:

# PART I: DETAILS ABOUT THE EXPERIMENT 1 SAMPLE

The CCES is an Internet survey of U.S. citizens that was conducted by YouGov/Polimetrix. YouGov/Polimetrix uses sampling and matching techniques to generate a sample that approximates the demographic composition of the adult U.S. population. The full sample for the 2008 CCES is based on the 2005-06 American Community Study, November 2008 Current Population Survey, and the 2007 Pew Religious Life Survey. Thus, this target sample is representative of the general population on a broad range of characteristics including a variety of geographic (state, region and metropolitan statistical area), demographic (age, race, income, education and gender), and other measures (born-again status, employment, interest in news, party identification, ideology and turnout). Polimetrix invited a sample of their opt-in panel of 1.4 million survey respondents to participate in the study. Invitations were stratified based on age, race, gender, education and by simple random sampling within strata. For more detailed information on this type of survey and sampling technique see Vavreck and Rivers (2008). More broadly, see Baker et al. (2010) for a report on the potential strengths and limitations of online panels.

The experiment sample was part of a private module on the 2008 CCES, with a target sample population of 1,800 individuals. These questions were asked of a subset, drawn at random, of 626 of the 1,800 individuals in the full sample. Of the 419 partisans used in our analysis, 81% were white, 7% were black, 8% were Hispanic and 54% were female. Their mean age was 48 years old, their median level of educational attainment was "some college," and 67% were married or in a domestic partnership.

Respondents in online samples often know more about politics and have more interest in politics than respondents in other surveys. It is not possible to establish whether this pattern holds with respect to knowledge in Experiment 1: the 2008 CCES includes few conventional knowledge questions (and none that have been used in recent ANES studies). But the data show that the pattern does hold with respect to political interest. For example, 65% of partisans in the 2008 CCES report being "very much interested" in politics; the corresponding percentage in the 2008 ANES is 38%. (No question in the 2008 ANES perfectly corresponds to the CCES political interest question. The closest ANES question, which we use here, is V085073a.)

Nonrepresentativeness on baseline characteristics does not necessarily imply that the treatment effects reported in Table 2 are different from those that we would find with a more representative sample (Druckman and Kam 2011). But it is easy to imagine ways in which the over-representation of politically interested people in our sample may cause us to overestimate—or to underestimate—the average effects of incentives. For example, even after conditioning on strength of partisanship, more interested people may be more likely to know the correct answers to our "partisan" questions. They may therefore be more likely to change their answers in response to payments that we offer for correct answers. If so, our estimates of the effects of incentives, while valid for our sample, overstate the effectiveness of such payments among ordinary partisans.

On the other hand, one may imagine that, even after conditioning on party identification, more interested respondents will issue more extreme answers to the "partisan" questions that we ask, or that they will hold to their answers more strongly (regardless of whether they know the correct answers). In either case, our estimates of the effects of payments for correct answers are likely to understate the effects that we would observe in a more representative sample.

We began to examine these possibilities by estimating models in which payments for correct answers are interacted with political interest. The relevant results appear in the third column of Table 2 and are discussed on pages 12-13. We find that the responses of politically interested subjects are more polarized, under ordinary conditions, than the responses of others. But interest does not moderate our estimated treatment effect. (The estimated coefficient on the relevant interaction term, –.23, is half the size of its standard error.) If anything, then, the overrepresentation of the interested makes our results conservative: a less interested population would be less polarized under ordinary survey conditions, and because the effect of incentives would be similar in magnitude, it would bring about a greater proportional reduction of the "distance" between the answers of members of different parties.

We can further consider the issue by considering how the results change when we weight the data to account for sample nonrepresentativeness. Table OA1 reports these results. The analyses are identical to those reported in Table 2, except that those in the "weighted analysis" columns incorporate the sample weights that are provided with the 2008 CCES. The critical coefficient in the table, "Payment for correct response ×

Political interest × Democrat," is again small and approximately half the size of its standard error (−.034, SE .065). This result further suggests that overrepresentation of the interested makes little difference to our results.

Finally, we note that we obtain similar results in our Mechanical Turk sample, which does seem to be representative of the population of U.S. partisans in terms of political interest. See the next part of this online appendix for details.

*Coding of Correct Answers*

The text of each Experiment 1 question is shown in Table 1, as is the response option that we deem "correct" for each question. We provide information about correctness to satisfy readers' curiosity: our analysis is about partisan divisions in responses to factual questions, not about correctness *per se.* Even so, a few additional words about some of the questions are in order.

One question asks about casualties of U.S. soldiers in Iraq in the second half of 2007 and the first half of 2008. The "surge" of U.S. troops in Iraq occurred during this period, and it corresponded to a widely reported decline in U.S. casualties: there were 37% fewer U.S. casualties in the first half of 2008 than in the second half of 2007 (Iraq Coalition Casualty Count 2014). Accordingly, we have coded "lower" (i.e., casualties fell) as the correct answer to the question. The response options to this question ("lower," "about the same," and "greater") were chosen because they have often been used in ANES retrospection items. See Experiment 2 for items that permit a wider range of responses.

Two of the questions were about the ages of John McCain and Barack Obama. Had McCain won the election, he would have been the oldest first-term president in history. His age was a particular concern to voters in 2008 (e.g., Benen 2008, Alonso-Zaldivar 2008, Pew 2008b), especially among the elderly (Pew 2008a). Obama's age was a lesser issue, although the concern that he was "too young for the no. 1 job" did surface (e.g., Calabresi 2008).

## PART II: DETAILS ABOUT THE EXPERIMENT 2 SAMPLE AND STUDY, INCLUDING A REPLICATION ON THE 2012 CCES

We recruited 1,506 participants for the Mechanical Turk study over the web from March 29, 2012 to April 16, 2012. Subjects for the experiment were recruited with an advertisement for "A quick survey to see what you know and how you learn." Because Mechanical Turk samples tend be more Democratic than the general population, we invited equal numbers of Democrats and Republicans who had previously taken our unrelated surveys to participate in this study. We invited 115 each strong Democrats and Republicans, 208 each Democrats and Republicans, and 111 each weak Democrats and Republicans, in an attempt to attract more Republicans than are ordinarily found in Mechanical Turk samples. Of the 795 partisans in our sample, 65% were Democrats, 89 were assigned to the control group, 327 to the pay-for-correct-response group, and 379 to the pay-for-correct-and-"don't know" group. For this group, age ranged from 19 to 75 with a mean of 33, 54 percent were female, and 46 percent had at least a four-year college degree.

We only extended invitations to people who had previously identified themselves as U.S. residents. As a further check on the residence of our subjects, we geocoded the IP addresses that they used to participate in the experiment. Of the 1,506 participants, only 38 (2.5%) had IP addresses that we located outside of the United States, and an additional three participants had IP addresses that we could not geocode. The 38 outside-the-US participants were distributed among 22 different countries. Of course, many of these participants may have been U.S. residents who were connecting to our web site during temporary travels abroad.

For all of the questions asked in this experiment, we used a novel graphical input device to measure participants' attitudes. Part VII of this online e appendix displays examples of the "sliders" that we used to gather answers to each of the questions we asked. After we trained participants to use this interface (complete instructions appear below), we asked them to respond to each question by manipulating the slider. Additionally, in the conditions in which participants were paid for correct responses, subjects were informed that a response would be scored as correct if the slider overlapped the correct answer.

The experiment had three conditions: a control condition, the pay-for-correct condition, and the pay-for-correct-and-"don't know" condition. (It also had a fourth condition that we do not analyze here: see footnote 16.)

*Instructions in the control condition:* "Once again, your answers will be timed. By answering these questions, you will earn an additional 50 cent bonus."

*Instructions in the pay-for-correct condition:* "Once again, your answers will be timed. Additionally, we are now going to give you a [X] cent bonus for each question you answer correctly. We'll tell you how many questions you get right at the end of the survey. You'll get credit for answering a question correctly if the thick horizontal bar underneath your arrow covers the correct answer. So, for example, in the picture below, the arrow is at 5. If the correct answer were 5.25, which is under the bar, you would earn the bonus. If the correct answer was 7, however, you would not earn the bonus."

*Instructions in the pay-for-correct-and-"don't know" condition:* "Once again, your answers will be timed. Additionally, we are now going to give you a X cent bonus for each question you answer correctly. We'll tell you how many questions you get right at the end of the survey. You'll get credit for answering a question correctly if the thick horizontal bar underneath your arrow covers the correct answer. So, for example, in the picture below, the arrow is at 5. If the correct answer were 5.25, which is under the bar, you would earn the bonus. If the correct answer was 7, however, you would not earn the bonus. As an alternative to being paid for a correct answer, you can instead earn a $X \times Y$ cent bonus for each question you tell us you don't know the answer to. We'll pay you for saying 'don't know' if you click the check box next to 'don't know,' but when you do so, the location of your arrow, whether correct or incorrect, does not affect your payment. Because the payment for 'don't know' is $(Y \times 100)\%$ of the payment for getting an answer correct, you will on average earn more by selecting don't know than your best guess if you are less than $(Y \times 100)\%$ sure that the bar underneath the arrow covers the correct answer."

*Analysis of consultation of outside references:* After the survey was over, we asked participants if they had looked up the answers to each question that they were asked, noting explicitly that "Your bonus is already determined, and we won't change your bonus in any way on the basis of your answer to these

questions." Of our 795 partisan participants, only 20 (2.5 percent) reported looking up the answer to 41 questions (0.74 percent of all questions asked). The percentages of user-questions by treatment assignment are 0.32 percent (control), 0.96 percent (pay for correct), and 0.64 percent (pay for correct and pay for don't know).

*Sample representativeness.* As with the Experiment 1 sample, one may be concerned about nonrepresentativeness of the Mechanical Turk sample that we use in Experiment 2. The Experiment 2 sample is far more diverse, and representative of the population of American partisans, than most samples that are used in studies of incentives: the large majority of those studies continue to be composed chiefly of undergraduates, and Mechanical Turk samples tend to be both more diverse and more representative than undergraduate samples (e.g., Berinsky, Huber, and Lenz, 355-65). Even so, one might fear that the sample overrepresents the interested or the knowledgeable, or those who are highly responsive to incentives, in ways that make the results unlike those that would be found in a more representative sample.

Consider first the concern about political interest. The finding that political interest does not moderate the effects of incentives should temper this concern: it suggests that overrepresentation of the interested would make little difference to the results. (See pages 12-13 and the discussion in the previous part of this online appendix.) Perhaps even more to the point, the 2012 Mechanical Turk sample does not seem to overrepresent those who have a great deal of interest in politics. Our Mechanical Turk subjects were asked

> Some people seem to follow what's going on in government and public affairs most
>
> of the time, whether there's an election going on or not. Others aren't that interested.
>
> Would you say that you follow what's going on in government and public
>
> affairs…[most of the time / some of the time / only now and then / hardly at all]?

Only 28% of subjects responded "most of the time." In the 2008 ANES, which used an identical question (V085072), the corresponding percentage was 32%. (The question that we used to measure interest had been used for decades by the ANES, but it was dropped after the 2008 ANES time series study.)

Although highly interested people do not seem to be overrepresented in the Mechanical Turk sample, it remains possible that the sample overrepresents those who know a lot about politics. And

overrepresentation of the knowledgeable might limit the generalizability of our results. For example, if the Mechanical Turk sample contains an unusually large number of knowledgeable subjects, the effects of incentives may be larger in the sample than in ordinary populations: all else equal, knowledgeable partisans will be more able to converge to the same (correct) answer after being offered an incentive to do so.

Our Mechanical Turk sample includes the political knowledge item:

> Do you happen to know how much of a majority is required for the United States
> Senate and House to override a Presidential veto?

The response options to this question were "a majority (fifty percent plus one vote)," "two-thirds (sixty-seven percent)," "three-fourths (seventy-five percent)," "ninety percent," and "don't know." The question has not been asked in the ANES for decades, but it was asked in a 1999 RDD survey of Tallahassee residents that had an unusually high completion rate (Mondak and Davis 2000, esp. 221). We find that 72% of partisans answer the question correctly—a figure that is very close to the 74% that Mondak and Davis (2000, 213) find, albeit with a question that had slightly different response options.

Amazon.com's Mechanical Turk is an electronic forum in which "workers" offer to complete tasks—typically quite brief tasks—in exchange for money. Mechanical Turk subjects are thus those who are actively seeking small and immediate payments, and one might therefore worry that they are unusually responsive to the financial incentives that we offer for correct and "don't know" responses to knowledge questions. We begin to explore this possibility by nothing that the results that we obtained from Mechanical Turk subjects are similar to those that we obtained from a very different sample of participants—the 2008 CCES sample that we used in Experiment 1. Although CCES subjects are rewarded for their participation (as most survey subjects are), they are not included in the CCES sample on the basis of their willingness to perform small tasks in exchange for immediate payments. Even so, we find similar results across the two samples.

Of course, Experiment 2, which uses the Mechanical Turk sample, includes several innovations that do not appear in Experiment 1, including payments for "don't know" responses. (See pages 14-15 for details.) To more precisely replicate the Experiment 2 results, we included a one-question experiment in the 2012 CCES. The question was

How did the unemployment rate in the country change between January 2009, when

President Obama took office, and September 2012?

We offered seven response options: "decreased 2%" (coded 1), "decreased 1%" (.83), "no change" (.67), "increased 1%" (.5), "increased 2%" (.33), "increased 3%" (.17), "increased 4%" (0). As with our analysis of Experiment 2, respondents who selected "don't know" in the pay correct and don't know condition were assigned the mean (average) response among those in the control condition, regardless of their party. All other variable coding is consistent with Experiment 2. There were 573 subjects in the experiment.

The results are reported in Table OA2, and they are similar to those that we obtained in Experiment 2. Relative to the control condition, payments for correct responses and payments for correct and "don't know" responses both reduced partisan divergence. The effect of paying for both correct and don't know responses was larger than the effect of just paying for correct responses, but the difference was not statistically significant (F-test p-value = .16, one-tailed).[1]


## PART III: RESULTS INCLUDING PARTISAN LEANERS

See Tables OA3 and OA4.


## PART IV: ROBUSTNESS TO WITHIN, COLLAPSED, AND EXCLUDING CHEATERS ANALYSIS

See Tables OA5, OA6, and OA7.

---

[1] One may also expect that our estimates are too conservative because of something like panel conditioning: respondents may have taken so many surveys before ours that they have tired of surveys or otherwise become inured to the survey setting. Responses to a question at the end of our survey suggest that this may not be the case. Only 21% of our subjects reported taking at least six Mechanical Turk surveys (counting our own) in the previous month. By contrast, 56% of subject reported taking no more than two Mechanical Turk surveys in the previous month.

## PART V: PREVIOUS RESEARCH ON PARTISAN DIVERGENCE IN FACTUAL ASSESSMENTS

A long line of research has noted partisan differences in evaluation of factual matters relating to politics. The questions in our experiments were chosen based on prior research documenting partisan divisions for similar topics. Here, we list the motivating research for our different questions. In Experiment 1, we asked questions about performance during the American invasions of Iraq and Afghanistan (see Jacobson 2010), economic performance during President Bush's tenure (see Bartels 2002; Evans and Andersen 2006; Kinder and Mebane 1983), and Obama and McCain's age during the 2008 campaign (see Pew 2008, documenting partisan divisions over whether McCain was too old to be president). Given stark partisan differences in assessments of president popularity, we also asked examined whether partisans differed in their assessments of Bush's overall and within-party popularity. In Experiment 2, for similar reasons we included questions about economic performance, the Iraq war, and Obama's election performance. The presence of partisan divides on preferences for government spending on health care and defense, the TARP (bailout) program, global warming, and attitudes toward immigrants led us to act factual questions in those areas too.


## PART VI: REFERENCES TO WORKS CITED IN THE ONLINE APPENDIX

Alonso-Zaldivar, Ricardo. 2008 July 14. "Born Before Computers, McCain Is 'Too Old' to Some." *Associated Press.* http://usatoday30.usatoday.com/news/politics/2008-07-14-1535993616_x.htm. Accessed on 2014 February 23.

Baker, Reg, Stephen J. Blumberg, J. Michael Brick, Mick P. Couper, Melanie Courtright, J. Michael Dennis, Don Dillman, Martin R. Frankel, Philip Garland, Robert M. Groves, Courtney Kennedy, Jon Krosnick, Paul J. Lavrakas, Sunghee Lee, Michael Link, Linda Piekarski, Kumar Rao, Randall K. Thomas, and Dan Zahs. 2010. "Research Synthesis: AAPOR Report on Online Panels." *Public Opinion Quarterly* 74 (4): 711-81.

Bartels, Larry M. 2002. "Beyond the Running Tally: Partisan Bias in Political Perceptions." *Political Behavior* 24(2): 117-150.

Benen, Steve. 2008 August 24. "The Dreaded Septuagenarian Issue." http://www.washingtonmonthly.com/archives/individual/2008_08/014389.php. Accessed on 2014 February 23.

Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20: 351-68.

Calabresi, Steven G. 2008 July 22. "Too Young for the No. 1 Job?" *Chicago Tribune.* http://articles.chicagotribune.com/2008-07-22/news/0807210340_1_average-life-expectancy-constitution-barack-obama. Accessed on 2014 February 23.

Druckman, James N., and Cindy D. Kam. 2011. "Students as Experimental Participants: A Defense of the 'Narrow Data Base'." In *Cambridge Handbook of Experimental Political Science,* ed. James N. Druckman and Donald P. Green and James H. Kuklinski and Arthur Lupia. New York: Cambridge University Press.

Evans, Geoffrey and Robert Andersen. 2006. "The Political Conditioning of Economic Perceptions." *Journal of Politics* 68(1): 194-207.

Iraq Coalition Casualty Count. "Iraq Coalition Casualties: Fatalities by Year and Month." http://icasualties.org/Iraq/ByMonth.aspx. Accessed 2014 February 23.

Jacobson, Gary. 2010. "A Tale of Two Wars: Public Opinion on the U.S. Military Interventions in Afghanistan and Iraq." *Presidential Studies Quarterly* 40(4): 585–610

Kinder Donald R, and Walter Mebane. 1983. "Politics and economics in everyday life." In *The Political Process and Economic Change*, ed. KR Monroe, pp. 141–180. New York: Agathon.

Mondak, Jeffery J., and Belinda Creel Davis. 2001. "Asked and Answered: Knowledge Levels When We Will Not Take "Don't Know" for an Answer." *Political Behavior* 23 (September): 199-224.

Pew Research Center for the People and the Press. 2008 May 21. "McCain's Age Problem." http://www.pewresearch.org/daily-number/mccains-age-problem/. Accessed on 2014 February 23.

Pew Research Center for the People and the Press. 2008 October 21. "Growing Doubts about McCain's Judgment, Age, and Campaign Conduct." http://www.people-press.org/files/legacy-pdf/462.pdf. Accessed on 2014 February 23.

Vavreck, Lynn, and Douglas Rivers. 2008. "The 2006 Cooperative Congressional Election Study." *Journal of Elections, Public Opinion and Parties* 18 (4): 355-66.

PART VII: COMPLETE INSTRUCTIONS AND SCREENSHOTS FOR EXPERIMENT 2

(Begins following appendix tables and figures.)

Table OA1: Experiment 1: Effect of Payment for Correct Responses on
Partisan Differences in Scale Scores (Weighted and Unweighted Analyses)

| | Weighted Analysis | | | Unweighted Analysis | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Democrat (1=Yes, 0=Republican) | 0.128 | 0.116 | 0.112 | 0.118 | 0.105 | 0.082 |
| | [0.022]*** | [0.022]*** | [0.039]*** | [0.015]*** | [0.016]*** | [0.022]*** |
| Political interest × Democrat | | | 0.033 | | | 0.059 |
| | | | [0.044] | | | [0.030]** |
| Payment for correct response × Democrat | -0.063 | -0.057 | -0.045 | -0.065 | -0.059 | -0.057 |
| | [0.030]** | [0.025]** | [0.057] | [0.022]*** | [0.022]*** | [0.037] |
| Payment for correct response × Political interest × Democrat | | | -0.034 | | | -0.023 |
| | | | [0.065] | | | [0.046] |
| Payment for correct response | 0.035 | 0.023 | 0.031 | 0.038 | 0.032 | 0.045 |
| | [0.020]* | [0.017] | [0.041] | [0.016]** | [0.016]* | [0.029] |
| Payment for correct response × Political interest | | | 0.005 | | | -0.005 |
| | | | [0.046] | | | [0.035] |
| Political interest (0,1) | | | 0.002 | | | -0.034 |
| | | | [0.028] | | | [0.021] |
| Constant | 0.277 | 0.249 | 0.276 | 0.239 | 0.163 | 0.261 |
| | [0.033]*** | [0.072]*** | [0.041]*** | [0.021]*** | [0.060]*** | [0.024]*** |
| Observations | 3321 | 3299 | 3305 | 3321 | 3299 | 3305 |
| R-squared | 0.354 | 0.369 | 0.355 | 0.398 | 0.407 | 0.400 |
| Includes additional controls? | No | Yes | No | No | Yes | No |

Note: Source: 2008 CCES. Includes only Democrats and Republicans. Robust standard errors, clustered by respondent. Question fixed effects not reported. The "unweighted analysis" results are the same as those that are reported in Table 2. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table OA2: Replication of Experiment 2 on 2012 CCES

| | Obama Unemployment Performance (Higher Values Indicate Unemployment Decreased) |
|---|---|
| Democrat (1=Yes, 0=Republican) | 0.366 |
| | [0.050]*** |
| Democrat * Pay Correct | -0.132 |
| | [0.074]* |
| Democrat * Pay Correct and Don't Know | -0.222 |
| | [0.072]*** |
| Pay Correct | 0.130 |
| | [0.053]** |
| Pay Correct and Don't Know | 0.161 |
| | [0.053]*** |
| Constant | 0.235 |
| | [0.035]*** |
| Observations | 593 |
| R-squared | 0.109 |
| F-test, Pay Correct * Dem. > Pay DK and Correct * Dem. | 0.110 |

Note: Source: 2012 CCES. Includes only Democrats and Republicans. Robust standard errors. F-test p-values are one-tailed. * indicates significant at 10%; ** significant at 5%; *** significant at 1%.

Table OA3: Experiment 1 Including Partisan Leaners: Effect of Payment for Correct Responses on Partisan Divergence in Scale Scores

| | (1) | (2) | (3) |
|---|---|---|---|
| | Mean Scale Score (0 to 1) | | |
| | (Pooled for 8 questions with partisan gap, p<.10, among control cases) | | |
| Democrat (1=Yes, 0=Republican) | 0.115 | 0.104 | 0.088 |
| | [0.013]*** | [0.015]*** | [0.021]*** |
| Interest * Democrat | | | 0.042 |
| | | | [0.027] |
| Payment for Correct Response * Democrat | -0.061 | -0.056 | -0.060 |
| | [0.020]*** | [0.019]*** | [0.034]* |
| Pay Correct * Interest * Democrat | | | -0.007 |
| | | | [0.042] |
| Payment for Correct Response | 0.032 | 0.029 | 0.039 |
| | [0.014]** | [0.014]** | [0.026] |
| Pay Correct * Interest | | | -0.007 |
| | | | [0.032] |
| Knowledge (0,1) | | -0.001 | |
| | | [0.011] | |
| Race: White (1=yes) | | -0.030 | |
| | | [0.017]* | |
| Race: Black (1=yes) | | -0.045 | |
| | | [0.026]* | |
| Race: Hispanic (1=yes) | | -0.020 | |
| | | [0.025] | |
| Female (1=yes) | | 0.012 | |
| | | [0.010] | |
| Age (Years) | | 0.002 | |
| | | [0.002] | |
| Age-squared/100 | | -0.003 | |
| | | [0.002] | |
| Region: Northeast | | 0.029 | |
| | | [0.015]** | |
| Region: Midwest | | 0.028 | |
| | | [0.014]** | |
| Region: South | | 0.002 | |
| | | [0.013] | |
| Income (1=<10k; 14=>150k; 15=RF/Missing) | | 0.004 | |
| | | [0.002]** | |
| Income Missing | | -0.036 | |
| | | [0.022] | |
| Education (1=No HS; 6=Post-grad) | | -0.004 | |
| | | [0.005] | |
| Education: No HS | | -0.003 | |
| | | [0.024] | |
| Education: Some college | | 0.022 | |
| | | [0.013]* | |
| Education: 2-year college | | 0.020 | |
| | | [0.020] | |
| Education: 4-year college | | 0.008 | |
| | | [0.016] | |
| Married/Domestic Partnership (1=yes) | | -0.008 | |
| | | [0.011] | |
| Religious Attendance (1-6) | | -0.001 | |
| | | [0.003] | |
| Political Interest (0,1) | | | -0.026 |
| | | | [0.019] |
| Constant | 0.231 | 0.205 | 0.248 |
| | [0.018]*** | [0.053]*** | [0.022]*** |
| Observations | 4229 | 4199 | 4213 |
| R-squared | 0.405 | 0.414 | 0.407 |

Note: Source: 2008 CCES. Includes only Democrats and Republicans (with leaners). Robust standard errors, clustered by respondent. Question fixed effects not reported. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table OA4: Experiment 2 Including Partisan Leaners: Effect of Payment for Correct Responses on Partisan Divergence in Scale Scores

| | (1) | (2) | (3) |
|---|---|---|---|
| Sample | All 10 non-placebo questions with partisan-gaps (p<.10) pre-treatment | | |
| Specification | OLS | Tobit | OLS |
| Democrat (1=Yes, 0=Republican) | 0.111 | 0.116 | 0.111 |
| | [0.024]*** | [0.025]*** | [0.024]*** |
| Payment Correct * Democrat | -0.056 | -0.057 | |
| | [0.026]** | [0.028]** | |
| Payment DK and Correct * Democrat | -0.076 | -0.079 | |
| | [0.025]*** | [0.027]*** | |
| Payment for Correct Response | 0.013 | 0.011 | |
| | [0.020] | [0.021] | |
| Payment for DK and Correct Response | 0.039 | 0.038 | |
| | [0.020]** | [0.020]* | |
| Amount correct = 0.10 * Democrat | | | -0.053 |
| | | | [0.029]* |
| Amount correct = 0.25 * Democrat | | | -0.062 |
| | | | [0.029]** |
| Amount correct = 0.50 * Democrat | | | -0.073 |
| | | | [0.029]** |
| Amount correct = 0.75 * Democrat | | | -0.007 |
| | | | [0.032] |
| Amount correct = 1.00 * Democrat | | | -0.083 |
| | | | [0.035]** |
| Prop. payment for DK=.20 * Democrat | | | -0.021 |
| | | | [0.018] |
| Prop. payment for DK=.25 * Democrat | | | -0.020 |
| | | | [0.020] |
| Prop. payment for DK=.33 * Democrat | | | -0.015 |
| | | | [0.019] |
| Amount correct = 0.10 | | | 0.014 |
| | | | [0.023] |
| Amount correct = 0.25 | | | 0.019 |
| | | | [0.022] |
| Amount correct = 0.50 | | | 0.023 |
| | | | [0.023] |
| Amount correct = 0.75 | | | -0.024 |
| | | | [0.025] |
| Amount correct = 1.00 | | | 0.037 |
| | | | [0.028] |
| Prop. payment for DK=.20 | | | 0.021 |
| | | | [0.014] |
| Prop. payment for DK=.25 | | | 0.028 |
| | | | [0.018] |
| Prop. payment for DK=.33 | | | 0.019 |
| | | | [0.016] |
| Constant | 0.625 | 0.632 | 0.626 |
| | [0.020]*** | [0.021]*** | [0.020]*** |
| Observations | 5880 | 5880 | 5880 |
| R-squared | 0.176 | | 0.178 |
| F-test, Pay Correct * Dem. > Pay DK and Correct * Dem. | 0.080 | 0.080 | |

Source: Mechanical Turk, March-April 2012. The dependent variable is the mean scale score for the ten questions on which we observed pre-treatment partisan gaps of $p < .10$. It ranges from 0 to 1. The analysis includes only Democrats and Republicans (with leaners). Cell entries are coefficients with robust standard errors, clustered by respondent. Question fixed effects are not reported. * significant at 10%; ** significant at 5%; *** significant at 1% (two-tailed tests).

Table OA5: Experiment 2 Within person analysis

| | (1) | (2) |
|---|---|---|
| | Post-treatment cases asked in pre, all questions with partisan gap among pre-treatment cases, p<.10 | |
| Pre (Lagged) directed slider response | 0.636 | |
| | [0.015]*** | |
| Democrat (1=Yes, 0=Republican) | 0.077 | 0.160 |
| | [0.014]*** | [0.029]*** |
| Payment Correct * Democrat | -0.066 | -0.098 |
| | [0.016]*** | [0.033]*** |
| Payment DK and Correct * Democrat | -0.090 | -0.127 |
| | [0.017]*** | [0.031]*** |
| Payment for Correct Response | 0.032 | 0.022 |
| | [0.011]*** | [0.027] |
| Payment for DK and Correct Response | 0.057 | 0.056 |
| | [0.012]*** | [0.026]** |
| Constant | 0.198 | 0.608 |
| | [0.016]*** | [0.027]*** |
| Observations | 3275 | 3275 |
| R-squared | 0.638 | 0.190 |
| F-test, Pay Correct * Dem. > Pay DK and Correct * Dem. | 0.030 | 0.060 |

Source: Mechanical Turk, March-April 2012. The dependent variable is the mean scale score for the ten questions on which we observed pre-treatment partisan gaps of $p < .10$. It ranges from 0 to 1. The analysis includes only Democrats and Republicans answering questions they also answered pre-treatment. Cell entries are coefficients with robust standard errors, clustered by respondent. Question fixed effects are not reported. * significant at 10%; ** significant at 5%; *** significant at 1% (two-tailed tests).

Table OA6: Experiment 2 Excluding cheaters from the analysis

|  | (1) | (2) |
|---|---|---|
|  | Post-treatment cases, all questions with partisan gap among pre-treatment cases, p<.10 | Excluding people who report any cheating. |
| Democrat (1=Yes, 0=Republican) | 0.145 | 0.149 |
|  | [0.028]*** | [0.028]*** |
| Payment Correct * Democrat | -0.087 | -0.090 |
|  | [0.030]*** | [0.030]*** |
| Payment DK and Correct * Democrat | -0.117 | -0.123 |
|  | [0.029]*** | [0.029]*** |
| Payment for Correct Response | 0.018 | 0.018 |
|  | [0.025] | [0.025] |
| Payment for DK and Correct Response | 0.049 | 0.050 |
|  | [0.024]** | [0.024]** |
| Constant | 0.614 | 0.613 |
|  | [0.026]*** | [0.026]*** |
| Observations | 4608 | 4492 |
| R-squared | 0.179 | 0.179 |
| F-test, Pay Correct * Dem. > Pay DK and Correct * Dem. | 0.020 | 0.010 |

Source: Mechanical Turk, March-April 2012. The dependent variable is the mean scale score for the ten questions on which we observed pre-treatment partisan gaps of p < .10. It ranges from 0 to 1. The analysis includes only Democrats and Republicans. Cell entries are coefficients with robust standard errors, clustered by respondent. Question fixed effects are not reported. * significant at 10%; ** significant at 5%; *** significant at 1% (two-tailed tests).

Table OA7: Experiment 2 Collapsed analysis to one observation per participant

| | Post-treatment cases, all questions with partisan gap among pre-treatment cases, p<.10 |
|---|---|
| Democrat (1=Yes, 0=Republican) | 0.146 |
| | [0.023]*** |
| Payment Correct * Democrat | -0.091 |
| | [0.026]*** |
| Payment DK and Correct * Democrat | -0.118 |
| | [0.025]*** |
| Payment for Correct Response | 0.023 |
| | [0.021] |
| Payment for DK and Correct Response | 0.050 |
| | [0.021]** |
| Constant | 0.546 |
| | [0.030]*** |
| Observations | 795 |
| R-squared | 0.175 |
| F-test, Pay Correct * Dem. > Pay DK and Correct * Dem. one-tailed | 0.050 |

Source: Mechanical Turk, March-April 2012. The dependent variable is the mean scale score for that respondent across all the questions on which we observed pre-treatment partisan gaps of p < .10. It ranges from 0 to 1. The analysis includes only Democrats and Republicans. Cell entries are coefficients with robust standard errors. Question fixed effects are not reported. * significant at 10%; ** significant at 5%; *** significant at 1% (two-tailed tests).

You are being asked to complete an online research survey that will take approximately 7-9 minutes. The survey is conducted by researchers at REDACTED to study how people learn. This page describes your consent.

Findings from this study may be reported in scholarly journals, at academic seminars, and at research association meetings. The data will be stored at a secured location and retained indefinitely. No identifying information about you will be made public and all of your choices will be kept completely confidential. Your participation is voluntary. You are free to stop the survey at any time without penalty.

There are no known risks associated with this study beyond those associated with everyday life. Although this study will not benefit you personally, we hope that our results will add to the knowledge about how people learn. You will receive $0.50 for completing the survey, paid through Amazon Mechanical Turk. You will also have the opportunity to earn a bonus of $0.50 or more, although not everyone will receive a bonus.

To participate in the study, you must be at least 18 years old and a United States resident. JavaScript must be activated on your browser so that the graphics in the survey will work properly. The next page will test your browser.

If you have any questions about the research, you can contact REDACTED. If you have any questions about your rights as a research participant or concerns about the conduct of this study, you may contact the REDACTED Human Subjects Committee, Box REDACTED, REDACTED, REDACTED, REDACTED, REDACTED@REDACTED.edu.

When you are ready to begin, please elect to participate and press the Submit button. You will then be taken to the first page of the survey.

- ◉ I agree to participate.
- ○ I do not agree to participate.

Submit

To confirm that our survey graphics will work with your browser, please follow the instructions in the graphic below. You have 20 seconds to complete this task. After 20 seconds, your browser will automatically proceed to the next page.

Please drag the arrow as far as you can to the right. You can move the arrow by clicking on the arrowhead and dragging.

Arrow

You have 16 seconds to submit your answer before your current answer is automatically submitted.

# Please read carefully and answer the following questions.

Here are two personality traits that may or may not apply to you. Please check the box to indicate the extent to which *you agree or disagree with each statement*. You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other. To demonstrate that you've read this much, just go ahead and select both disagree strongly and agree strongly for both questions below, no matter how you would actually answer each question. In other words, to confirm that you are paying attention, for each question please check both of these two boxes.

### *I see myself as:* Dependable, self-disciplined.

- ☐ Agree strongly.
- ☐ Agree moderately.
- ☐ Agree a little.
- ☐ Neither agree nor disagree.
- ☐ Disagree a little.
- ☐ Disagree moderately.
- ☐ Disagree strongly.

### *I see myself as:* Disorganized, careless.

- ☐ Agree strongly.
- ☐ Agree moderately.
- ☐ Agree a little.
- ☐ Neither agree nor disagree.
- ☐ Disagree a little.
- ☐ Disagree moderately.
- ☐ Disagree strongly.

Next

# Please read carefully and answer the following questions.

What is the highest level of education that you have achieved?

- ○ No high school diploma.
- ○ High school diploma or equivalent.
- ○ Some college.
- ○ Two year degree.
- ○ Four year college graduate.
- ○ Post-graduate.

What is the year of your birth?

What is your gender?

- ○ Female.
- ○ Male.

What is your state of residence?

Generally speaking, do you usually think of yourself as a Democrat, a Republican, an Independent, or what?

- ○ Democrat.
- ○ Republican.
- ○ Independent.
- ○ Other.

Next

# Please read carefully and answer the following questions.

Some people seem to follow what's going on in government and public affairs most of the time, whether there's an election going on or not. Others aren't that interested. Would you say you follow what's going on in government and public affairs...?

- ○ Most of the time.
- ○ Some of the time.
- ○ Only now and then.
- ○ Hardly at all.

We are interested in the kinds of things people do when they use the internet. What kinds of things have you used the internet for in the LAST SEVEN DAYS? (Choose as many as apply to you)

- ☐ Get directions.
- ☐ Plan vacations.
- ☐ Keep in touch with friends.
- ☐ Look at sports highlights.
- ☐ Find restaurants.
- ☐ Pay bills.
- ☐ Look up movie times.
- ☐ Shopping.
- ☐ Read the news.
- ☐ Read about politics.

Do you happen to know how much of a majority is required for the United States Senate and House to override a Presidential veto?

- ○ A majority (fifty percent plus one vote).
- ○ Two-thirds (sixty-seven percent).
- ○ Three-fourths (seventy-five percent).
- ○ Ninety percent.
- ○ Don't know.

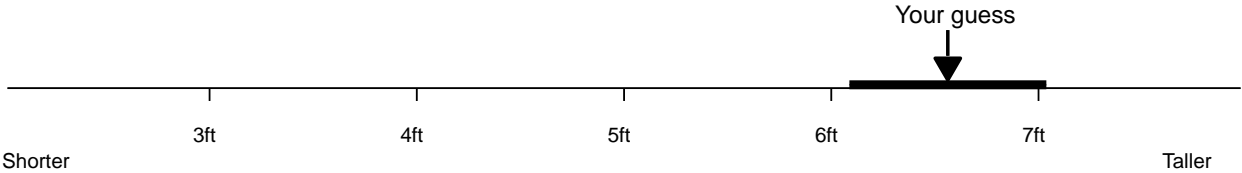Do you think most professional athletes are good role models for children today?

- ○ Yes.
- ○ No.
- ○ Don't know.

Next

In this study, we'd like you to tell us **what you think the correct answer is** to a number of questions. We don't want you to look up those answers or talk to someone else, so even if you don't know please just give us your best guess. For each question, we'll give you a short period of time -- 30 seconds -- to answer the question before we automatically take you to the next question.

To indicate your answer, we will ask you to slide the arrow on a line like that below to the point that is closest to your answer. You can slide that arrow by clicking your mouse on the arrowhead and dragging it to the left or right.

How tall is the average NBA player?

Your guess

| | | | | |
|3ft|4ft|5ft|6ft|7ft|

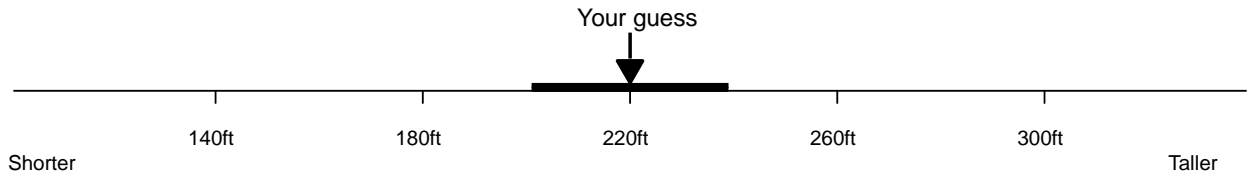Shorter                                                                                                          Taller

For example, in the above graphic, if you though the correct answer was 6 feet 6 inches, you would slide the arrow to the point midway between the lines marked 6 and 7 ft.

Give it a try! Once you are happy with where the arrow is located, you can click "Next." On the next page, we'll give you a timed example with another question.

Next

How tall is the Statue of Liberty, in feet, from the base of the feet to the top of the torch?

Your guess

| 140ft | 180ft | 220ft | 260ft | 300ft |

Shorter                                                                                          Taller

In this example, we are asking you to indicate your best guess as to how tall the Statue of Liberty is. You can also see how the countdown timer works -- you have 45 seconds to indicate your answer (see below). After you've indicated your best guess, click "Next" or just wait to go to the next page. When the timer is up, you will automatically be routed to the next page.

You have 45 seconds to submit your answer before your current answer is automatically submitted.

Next

We're almost ready to begin. Before we proceed, we just want to make sure we've been clear about what we are asking you to do.

> Dave has two dozen apples. He eats half of them, and then eight more. How many apples are left?
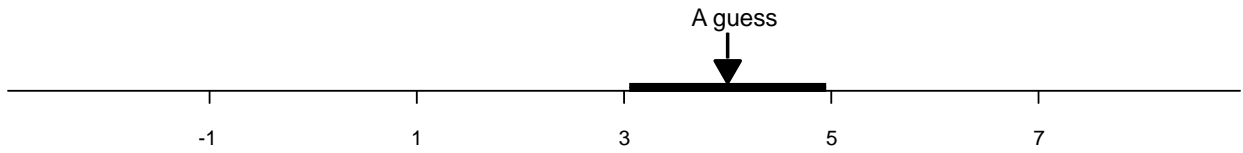>
> A guess
>
> -1    1    3    5    7
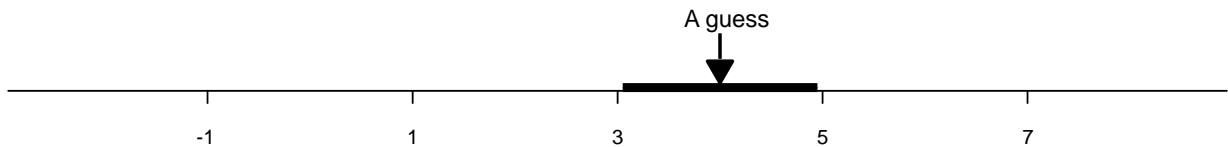
In the graph above, we've placed the arrow at a certain point to indicate somebody's response to the question. Which of the following has that person indicated is their best guess?

Their best guess is...

○ 1.

○ 2.

○ 3.

○ 4.

○ 5.

○ None of the above.

Next

Dave has two dozen apples. He eats half of them, and then eight more. How many apples are left?

A guess

-1          1          3          5          7

The arrow is located midway between 3 and 5, so the person's response is 4.

Next

We will now send you to the actual survey. On the next screen, you will be presented with your first question and will only have **a limited amount of time to respond**.

Please **do not use the back button in your browser** during this survey. Any questions your answer a second time by using the back button will not be recorded. When you are ready, please click Next.

Next

Please drag the slider to your best guess to the following

About how many U.S. soldiers were killed in Iraq between the invasion in 2003 and
the withdrawal of troops in December 2011?
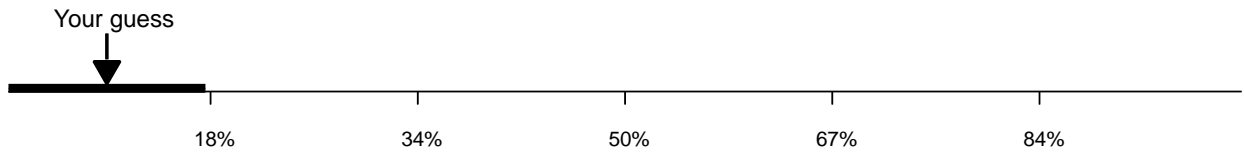
Your guess

| 2000 | 3000 | 4000 | 5000 | 6000 |
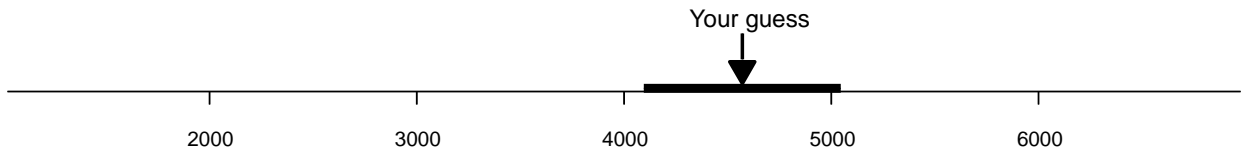
You have 27 seconds to submit your answer before your current answer is automatically submitted.

Next

Please drag the slider to your best guess to the following

According to the Census Bureau, in 2010 what percentage of the total population of the United States was born outside of the United States (foreign-born)?

Your guess

| 18% | 34% | 50% | 67% | 84% |

You have 28 seconds to submit your answer before your current answer is automatically submitted.

Next

Thank you for answering those questions, we'd now like you to answer a few more questions.

Once again, your answers will be timed.

**By answering these questions, you will earn an additional 50¢ bonus.**

Again, please **do not use the back button in your browser**. Any questions your answer a second time by using the back button will not be recorded. When you are ready to proceed, please click Next.

Next

Please drag the slider to your best guess to the following

In the 2008 Presidential Election, Barack Obama defeated his Republican challenger John McCain. In the nation as a whole, of all the votes cast for Obama and McCain, what percentage went to Obama?

Your guess

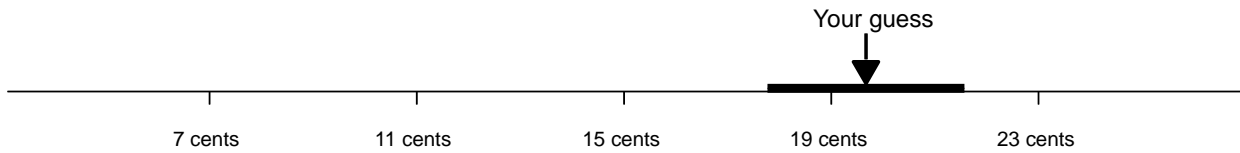52.0%          54.0%          56.0%          58.0%          60.0%
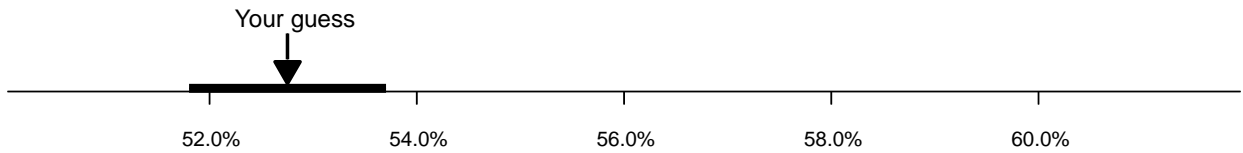
You have 28 seconds to submit your answer before your current answer is automatically submitted.

Next

Please drag the slider to your best guess to the following

For every dollar the federal government spent in fiscal year 2011, about how much went to the Department of Defense (US Military)?

Your guess

7 cents 11 cents 15 cents 19 cents 23 cents

You have 26 seconds to submit your answer before your current answer is automatically submitted.

Next

## Thank you for your participation!

Your bonus is already determined, and we won't change your bonus in any way on the basis of your answer to these questions. For research purposes...

**Did you look up the answer to this question?**

In the 2008 Presidential Election, Barack Obama defeated his Republican challenger John McCain. In the nation as a whole, of all the votes cast for Obama and McCain, what percentage went to Obama?

- ○ No, I did not look up th answer to this question.
- ○ Yes, I did look up the answer to this question.

**Did you look up the answer to this question?**

For every dollar the federal government spent in fiscal year 2011, about how much went to the Department of Defense (US Military)?

- ○ No, I did not look up th answer to this question.
- ○ Yes, I did look up the answer to this question.

**Did you look up the answer to this question?**

About how many U.S. soldiers were killed in Iraq between the invasion in 2003 and the withdrawal of troops in December 2011?

- ○ No, I did not look up th answer to this question.
- ○ Yes, I did look up the answer to this question.

**Did you look up the answer to this question?**

According to the Census Bureau, in 2010 what percentage of the total population of the United States was born outside of the United States (foreign-born)?

- ○ No, I did not look up th answer to this question.
- ○ Yes, I did look up the answer to this question.

**Did you look up the answer to this question?**

Compared to January 2001, when President Bush first took office, how had the level of unemployment, as measured using the unemployment rate, in the country changed by the time he left office in January 2009?

- ○ No, I did not look up th answer to this question.
- ○ Yes, I did look up the answer to this question.

**Did you look up the answer to this question?**

The Treasury Department initiated TARP (the first bailout) during the financial crisis of 2008. TARP involved loans to banks, insurance companies, and auto companies. Of the $414 billion spent, what percentage had been repaid, as of March 15, 2012?

- ○ No, I did not look up th answer to this question.
- ○ Yes, I did look up the answer to this question.

**Did you look up the answer to this question?**

Medicaid is a jointly funded, Federal-State health insurance program for low-income and needy people. For every dollar the federal government spent in fiscal year 2011, about how much went to Medicaid?

- ○ No, I did not look up th answer to this question.
- ○ Yes, I did look up the answer to this question.

Next

# Thank you for your participation!

What is the total number of Mechanical Turk surveys you have taken about current events or politics?

What is the total number of Mechanical Turk surveys you have taken about current events or politics **in the last month**?

If you would like to be contacted when we have future studies, please leave your email here. If not, leave blank:

If you would like to leave any comments or feedback, please do so here (up to 500 characters):

Pleast continue to the next page to retrieve your code for payment!

Next

Thank you for your participation!

You have now completed the survey.

If you have any questions, please contact REDACTED@REDACTED.edu. If you have any questions about your rights as a research participant or concerns about the conduct of this study, you may contact the REDACTED Human Subjects Committee at REDACTED@REDACTED.edu.

For the purposes of getting paid on Mechanical Turk, please enter the following code into the box on the survey's Mechanical Turk HIT page. You must enter this code to get your bonus.

# vuhtkwysobinecs

If you are curious about the sources we used to score your answers, please contact us through the Mechanical Turk interface and we are happy to provide references to you. Thank you!