Acquiescence Bias Inflates Estimates of Conspiratorial Beliefs and Political Misperceptions[*]

Seth J. Hill[†]

Margaret E. Roberts[‡]

October 25, 2021

Abstract: Citizen beliefs about the political world interest scholars, pundits, and politicians. Political scientists use opinion surveys to study beliefs about facts, such the current unemployment rate, and more conspiratorial beliefs, such as whether Barack Obama was born abroad. We find that these studies often ignore acquiescence-response bias, the tendency for survey respondents to endorse any assertion made in a survey question regardless of content. Replicating recent scholarship we show that acquiescence bias inflates estimated incidence of conspiratorial beliefs and political misperceptions among survey respondents in the U.S. and China by up to 50%. Acquiescence bias is disproportionately prevalent among more ideological respondents, inflating correlations between political ideology such as conservatism and endorsement of conspiracies or misperception of facts. We conclude that more attention ought to be paid to survey measurement, and propose and demonstrate two methods to correct for acquiescence bias.

Keywords: Political beliefs; misperception; rumors and conspiracies; acquiescence-response bias; survey methodology.

Social scientists, public intellectuals, and politicians have long been interested in citizen beliefs about the political world. Scholars have measured citizen beliefs about political rumors, political conspiracies, and politically-relevant facts, the relationship of such beliefs and misperceptions to the contemporary political information environment, and the etiology and correlates of these beliefs.[1]

Questions about conspiratorial beliefs and factual misperceptions have recently increased in importance with the segmentation of traditional news media, the expansion of digital sources of information, the rise of social media, selective censorship, and the spread of explicitly-partisan news and disinformation. Social scientists and journalists express increasing concern that the public is misinformed about public affairs. In many cases, investigations of beliefs reflect poorly on the citizenry's ability to distinguish truth from innuendo generating news headlines such as "Even If It's 'Bonkers,' Poll Finds Many Believe QAnon And Other Conspiracy Theories" (National Public Radio, 2020) and "Half of Americans Believe in 9/11 Conspiracy Theories" (Gohse, 2016).

The empirical evidence generating these troubling conclusions, however, come from opinion surveys, a measurement tool that existing research finds poses many challenges to studying political beliefs. Survey responses are influenced by measurement challenges including partisan cheerleading, shirking, lack of incentives, and idiosyncratic error (e.g., Bullock et al., 2015; Prior and Lupia, 2008; Krosnick, 1991; Ansolabehere, Rodden, and Snyder, 2008; Westwood et al., 2021).

We demonstrate an additional and first-order challenge to measuring conspiratorial and political beliefs. We return to an old literature on "acquiescence-response bias" – the phenomenon where survey respondents are more likely to answer True, Agree, and Yes than False, Disagree, or No regardless of the question asked. We find that acquiescence-response bias can have a profound effect on not only measurement of beliefs but on the correlation between beliefs and characteristics of the survey respondents such as education and political ideology.

The current literature studying conspiracies, misperceptions, and disinformation could benefit from attention to acquiescence bias. Many recent surveys measure beliefs by asking respondents whether they endorse a false belief or a conspiracy. Does the respondent "agree or disagree" that 9/11 was a conspiracy by global elites? Is the rumor that Barack Obama was born abroad "true or false?" "Yes or no, is the following statement correct?" In the presence of acquiescence bias these common instrumentation choices inflate estimates of endorsement regardless of the content of the question (e.g., Schuman and Presser, 1981; Watson, 1992; Krosnick, 1999; Billiet and McClendon, 2000). In the case of political beliefs, if most questions are written so that the acquiescent response indicates endorsement of a false belief or conspiracy, the bias would cause overestimates of the population rate of endorsement. If, on the other hand, survey questions were asked such that the agreeable response indicated rejection of the conspiracy, in the presence of acquiescence bias the instrument would underestimate level of endorsement. In other words, acquiescence bias can influence estimates in either direction depending only on construction of the survey question.

Acquiescence bias can cause an additional problem. Acquiescence bias is likely to be a dispositional trait driven by personality, education, or life experience (Schuman and Presser, 1981). Some respondents might default to the acquiescent option as they try to answer survey questions that they have not before encountered or about which they are uncertain. This means that acqui-

[1]For some recent examples, see Jerit and Zhao (2020), Guess, Nyhan, and Reifler (2020), Pennycook and Rand (2019), Lazer et al. (2018), Vosoughi, Roy, and Aral (2018), Allcott and Gentzkow (2017), Berinsky (2017), Flynn, Nyhan, and Reifler (2017), Huang (2017), Oliver and Wood (2014), Nyhan and Reifler (2010).

escence bias varies across the population in ways that correlate with individual characteristics – for example, with partisanship, ideology, or education. Correlation between acquiescence bias and individual characteristics then biases estimates of correlations between individual characteristics and conspiratorial and false beliefs.

We first present evidence that acquiescence bias can impact the inferences scholars draw from survey evidence on political beliefs, sometimes causing overestimates of misperceptions and conspiratorial beliefs of 40 and 50 percentage points. We fielded six surveys that replicated questions from recent studies about political rumors and facts in addition to our own new question. Our surveys cover a range of topics, samples, elicitation instruments, and types of misperceptions and conspiracies. We fielded surveys to thousands of respondents in the U.S. and China sampled from different online survey firms (Lucid, Qualtrics, and NORC) as well as Mechanical Turk. We asked about beliefs on matters of objective statistics, such as GDP growth and currency exchange rates, as well as about beliefs of conspiratorial nature such as airplane contrails being government-sponsored chemicals. We used different instruments to elicit beliefs: binary True/False, agreement Likert scales, and continuous subjective probabilities, with and without monetary incentives for accuracy. We find evidence of acquiescence bias across all factors. The variability in features across our studies allows greater confidence that acquiescence bias contaminates inferences for many scholars across many settings, notably including comparative scholars of opinion in China, and that our results are not driven by one or two idiosyncratic wordings or samples.

Our research design is simple. For each question on each survey, we fielded both the original version of the question and an alternative version of the same question that flips the meaning of the agreeable response so that in the alternative version acquiescence bias works in the opposite direction. We wrote the alternative version to be logically equivalent to the original wording (as best we could). Where we were successful attaining logical equivalence, absent acquiescence bias respondents should return the same rate of endorsement to each version of the question.

We find that survey respondents do not return the same rate of endorsement to the two versions of each question on the majority of questions across samples, topics, instruments, and types of belief. The magnitude of difference can be substantively large. This shows that inferences about the magnitude of conspiracy endorsement or misperception of political facts can depend importantly on whether the question is worded with affirmative belief measured with an acquiescent response. We find particularly large acquiescence bias on nearly all questions fielded to respondents in China and on questions about democratic norms and the transition of power following the 2020 American presidential election (asked in Clayton et al., 2021).

We then show how acquiescence bias obscures estimates of correlations between characteristics of individuals and conspiratorial thinking. For example, we find that acquiescence bias magnifies the partisan differences of beliefs in conspiracy theories. We replicate a highly-cited work on fake news and rumors (Allcott and Gentzkow, 2017) and find that partisan differences for some beliefs are halved when elicited using a non-acquiescence framing.

In recent years, many prominent studies and news stories have reported on troubling magnitudes of conspiracy beliefs by ideological conservatives (e.g., Garrett and Bond, 2021). We find that subjects who identify as very conservative exhibit larger acquiescence bias than both those who identify as very liberal and those with less ideological identification. Our results suggest that existing conclusions, at least in part, are driven by greater acquiescence bias by survey respondents with conservative leanings rather than greater endorsement of conspiracies by conservatives. We also find greater acquiescence bias by strong liberals relative to less ideological subjects. These

correlations suggest opportunities for future research.

After documenting bias in both extent and correlation, we present two simple and easy-to-use methods to correct for acquiescence bias. The first method generates estimates of both the population rate of the belief purged of acquiescence bias and of the average magnitude of bias. The second method generates estimates of population correlations between characteristics and conspiratorial beliefs purged of acquiescence bias. Each method rests on a simple assumption that acquiescence bias is on average symmetric between positive- and negative-keyed questions. To implement the fix, scholars simply field two versions of each question, one positive- and one negative-keyed, and apply the statistical correction.

We also examine the use of alternative multiple-choice versions of questions but find ambiguous support for claims that these approaches ameliorate bias. Finally, we conclude with speculation on the characteristics of questions that seem to correlate with greater and lesser acquiescence bias, suggesting pathways for future theory and research.

1   Acquiescence-response bias

Acquiescence bias – "the tendency to endorse any assertion made in a question, regardless of its content" (Krosnick, 1999, 552) – inflates endorsement and agreement in survey response. Acquiescence bias has long been appreciated by psychologists (e.g., Schuman and Presser, 1981; Watson, 1992; Krosnick, 1999; Billiet and McClendon, 2000). These studies note that acquiescence bias varies in magnitude across the population. Schuman and Presser (1981, ch. 8) summarize three interpretations of this heterogeneity. First, tendency to agree might be an individual personality trait generated by genes and environment. Second, acquiescence bias might follow from status differentials between participant and surveyor.[2] Third, acquiescence bias might be a heuristic response rule for participants when they don't know how to respond to a question.

Because magnitude of acquiescence bias correlates with individual characteristics such as education, inferences about the relationship between political characteristics and beliefs might also be inaccurate in the presence of acquiescence bias. This problem, in fact, was appreciated by survey researchers in political science in the 1960s. The authors of *The American Voter* found that acquiescence bias undermined conclusions others had drawn about the psychological authoritarianism scale. They showed that the existing finding that those with less education were more authoritarian was due only to their higher rate of acquiescence bias (Campbell et al., 1960, 512-514).

Acquiescence bias is a common concern in psychometrics where researchers construct psychology indices based on responses to multiple agree/disagree Likert items. To improve measurement and validity, psychologists field to respondents multiple items thought related to the underlying index. But, if all items are coded such that the agree response indicates one end of the index and disagree the other, then the subset of respondents who exhibit acquiescence bias will be improperly placed toward the agree end of the index.

Psychometricians use a variety of techniques to ameliorate acquiescence bias when scaling indices. These solutions generally suggest fielding multiple items to construct within-subject corrections. Methods applied include Item Response Theory models, factor analysis, or simple summing across within-subject randomization. For a review, see Billiet and McClendon (2000).

Unfortunately, psychometric methods for scaling indices can be less useful for social scientists

---

[2]Our surveys were fielded online which might have lessened the influence of interviewer-interviewee status differentials more common in face-to-face settings.

interested in beliefs about specific items. Further, in contrast to measuring personality traits, we might be especially concerned about acquiescence bias when eliciting beliefs about complicated or controversial quantities such as political rumors and facts.

Political scientists often ask questions about political beliefs giving respondents agreeable response options ("true," "yes," "agree," "likely so," etc.). If acquiescence arises as a heuristic response when subjects do not know how to answer, bias could be prevalent when (a) asking about topics the subject has not before encountered such as political rumors or conspiracies, or (b) eliciting beliefs about complicated matters of fact beyond the subject's personal experience. The consequences of bias increase in the rarity of the target belief. The smaller the fraction of the population that holds a belief – e.g., when asking about outlandish conspiracies – the more likely agreeable responses are due to acquiescence bias rather than actual belief.

For example, what should we conclude if we found that 30 percent of subjects respond "True" to the question "True of False: Changes to the health care system enacted by Congress and the Obama administration created 'death panels' which have the authority to determine whether or not a gravely ill or injured person should receive health care"? It might be that 30 percent did not know how to respond to the question and so used the agreeable response "True" as a default, that 30 percent truly believe the Obama administration created death panels, or that it is one of the many combinations of the two types that sum to 30 percent. Acquiescence bias poses a fundamental threat to inferences about endorsement of political conspiracies and to inferences about factual beliefs about politics.

2   Measuring acquiescence bias in studies of conspiracies, rumors, and facts

To understand the extent of acquiescence bias in studies of conspiracies, rumors, and facts, we fielded six surveys in China and the United States eliciting subject beliefs about political conspiracies and politically-relevant facts. We asked three types of questions: 1) questions fielded by other scholars on conspiracies and beliefs in rumors, 2) questions fielded by other scholars on beliefs in political facts, and 3) questions of interest to our own substantive research agendas on factual beliefs and political learning in which we elicited probabilistic beliefs. In total, we asked 53 questions across the six surveys.

We fielded two versions of each question. First, we fielded the version as written by the original study, we call this verion "positive-keyed." Second, we fielded an alternative version that was the negative of the original question so that the agreeable response had the opposite meaning, we call this version "negative-keyed." For example, Allcott and Gentzkow (2017) asked a nationally representative sample after the 2016 election "At the time of the 2016 election, would your best guess have been that this statement was true?":

> An FBI agent connected to Hillary Clinton's email disclosures murdered his wife and shot himself.

The response options were "Yes, true," "Not sure," or "No, false." We randomized whether our respondents received the original version or the alternative version

> An FBI agent connected to Hillary Clinton's email disclosures DID NOT murder his wife or shoot himself.

It is important to note that positive-keyed does not mean the content of the statement or conspiracy is positive, but rather that the agreeable response option aligns with the original version

of the question.[3] In other words, some positive-keyed questions include a "NOT" phrase, and some negative-keyed questions do not. For example, the positive-keyed question of the conspiracy surrounding Barack Obama's birth certificate is phrased, "President Barack Obama was not really born in the United States and does not have an authentic Hawaiian birth certificate," because this was the original version of the question; whereas the negative-keyed version is phrased, "President Barack Obama was born in the United States and has an authentic Hawaiian birth certificate."

Randomization occurred at the subject-question level so each subject received a mix of positive-keyed and negative-keyed questions following the psychometric literature. Creating a negated version of each item was not always trivial, but we hope that difficulties in negating cancel out across questions. We present the original and negated version of each item in Appendix Tables A1, A2, A3, and A4.

For each question, we calculated proportion of subjects agreeing with the positive-keyed statement – the original version of the conspiracy, rumor, or fact. We code "don't know" and "not sure" responses as not agreeing. For ease of comparison, we recode the negated question so that responses have the same target belief as the original. Absent bias, the two percentages should be equivalent, subject to sampling variability. Acquiescence bias, in contrast, would push responses to the two versions in opposite directions with greater estimated endorsement for the positive-keyed wording.

In Figure 1, we present the results of the 53 survey questions we fielded. On the x-axis we plot the estimated agreement with the conspiracy, rumor, or fact when the question is "positive-keyed" as in the survey we replicate. On the y-axis, we plot estimated agreement with the conspiracy, rumor, or fact when the question is "negative-keyed," the negated form of the original version. Note that the responses of the negative-keyed version have been recoded so they exist on the same scale as the positive-keyed and, therefore, absent bias the points in this plot should fall approximately on the 45 degree line.

All but one point falls below the 45 degree line, meaning that the positive-keyed version of the question has a larger estimate of endorsement than the negative-keyed version. In some questions, the points are quite close to the line; for example, question 21 fielded in the U.S. on whether the Chinese government created the coronavirus as a biological weapon does not exhibit large amounts of acquiescence bias. In other questions, the positive-keyed version of the question produces a drastically larger estimate of belief. For example, question 27 on whether or not the FBI director alerted Congress about new emails on Hillary Clinton's server on October 28, 2016 yielded an agreement of 64% when asked in the positive-keyed version, but only 22% when asked in a negative-keyed version.

In the remainder of this section we summarize our surveys and the three main types of questions we describe in Figure 1. We begin by describing the extent of acquiescence bias in our two most recent surveys, one in the U.S. and one in China, where we elicited beliefs about conspiracies and politically relevant facts drawn from previous work. We then describe four other surveys where we asked respondents in the U.S. and China about their probabilistic beliefs about facts, compensating them for being closer to the correct answer. As shown in Figure 1, in all studies we see large degrees of acquiescence bias from respondents in both countries.

---

[3]In the case of questions where we elicit probabilistic beliefs on facts, as we describe below, we are not replicating previous work. In this case, we align the "more" as positive-keyed and "less" as negative-keyed for ease of interpretation. Note that our results would not change if we were to align the "less" as positive-keyed and "more" as negative-keyed.

Figure 1: Effect of question wording on agreement with rumors and facts.
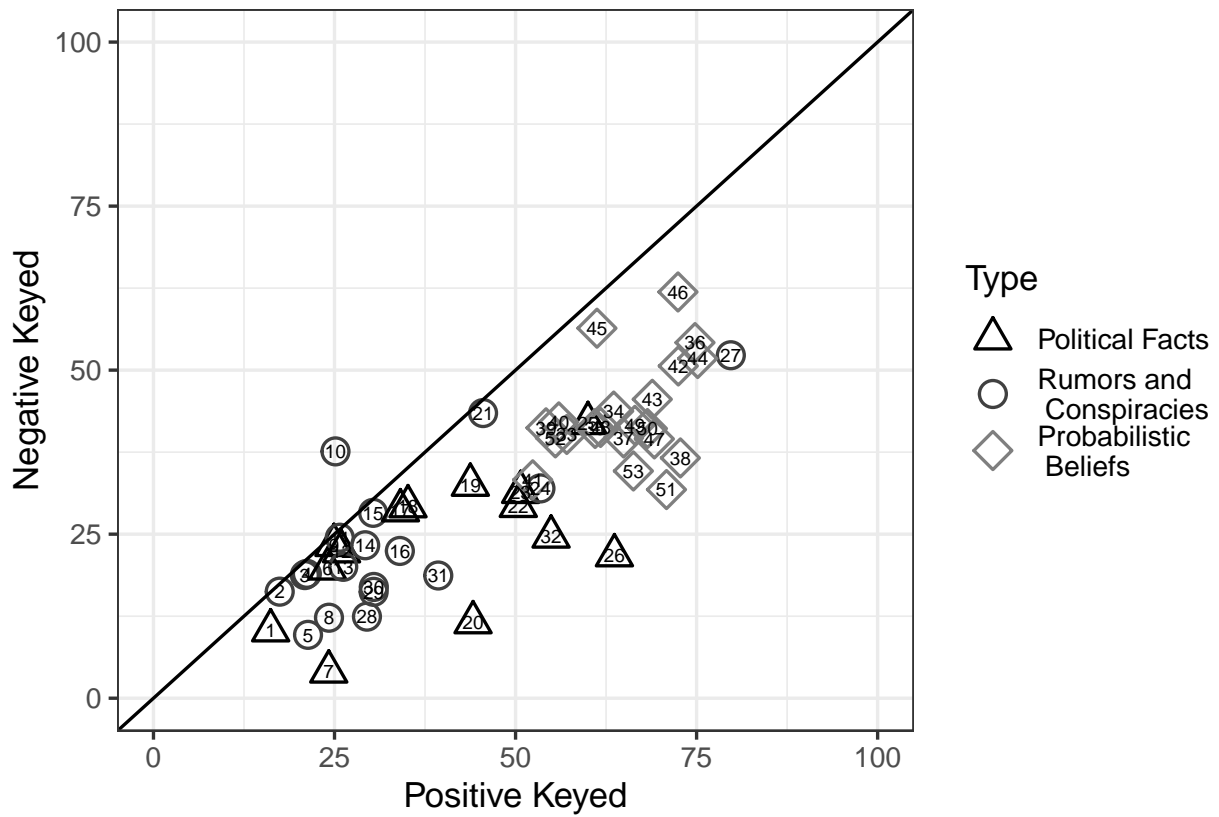
1	Pope endorse (US 2020)	15	Soros conspiracy (US 2020)	29	GMO Bioterrorism (China 2021)	43	Exports (China 2019)
2	US created coronavirus (US 2020)	16	Death panels (US 2020)	30	Soros conspiracy (China 2021)	44	Overseas Students (China 2019)
3	Ireland asylum (US 2020)	17	Immigrant pop Canada (US 2020)	31	9/11 conspiracy (China 2021)	45	Jobs 2016 (US 2016)
4	FBI agent (US 2020)	18	Immigrant pop US (US 2020)	32	GMO Increase Yield (China 2021)	46	Currency devaluation (US 2016)
5	Contrails conspiracy (US 2020)	19	Iraq WMD (US 2020)	33	2016 GDP Growth (China 2017)	47	GDP 2014 (US 2016)
6	Vitamin C (US 2020)	20	FBI Clinton charges (US 2020)	34	Currency devaluation (China 2017)	48	Jobs 2016 (US 2017-18)
7	Hand wash (US 2020)	21	Coronavirus weapon (US 2020)	35	Military Spending (China 2017)	49	Currency devaluation (US 2017-18)
8	Violence for votes (US 2020)	22	Clinton stumbled (US 2020)	36	GDP 2015 (China 2017)	50	GDP 2014 (US 2017-18)
9	RuPaul groped (US 2020)	23	Beyonce rally (US 2020)	37	Qualcomm Fine (China 2017)	51	Food stamps 2018 (US 2017-18)
10	9/11 conspiracy (US 2020)	24	Clinton deplorables (US 2020)	38	IMF Currency Basket (China 2017)	52	Currency devaluation (US 2020)
11	Clinton Foundation (US 2020)	25	FBI director (US 2020)	39	ZTE Fine (China 2019)	53	Food stamps 2018 (US 2020)
12	Pence vulgar (US 2020)	26	Accept election loss (US 2020)	40	Hong Kong Work Week (China 2019)		
13	Iraq conspiracy (US 2020)	27	Obama birth (China 2021)	41	Scholarships (China 2019)		
14	Obama birth certificate (US 2020)	28		42	Military Parade (China 2019)		

2.1 Replications of Previous Work

In two studies fielded in 2020 and 2021, we asked about conspiratorial beliefs and beliefs about politically-relevant facts in online surveys of respondents in the U.S. and China. The first was fielded using the online survey platform Lucid, which provided us a nationally-representative sample of 2,055 respondents in December 2020. To account for differential non-response, we created post-stratification weights to Census targets and 2020 presidential vote. We present details on the sample in Appendix Section B.2. The second was fielded using the online survey platform Qualtrics in China in March-April 2021. Qualtrics provided a quota-based sample matched to population targets. We present details on this Qualtrics survey in Appendix Section B.1.

2.1.1 Rumors and conspiracies

We categorize our questions into three groups though acknowledge up front the taxonomy imperfect. For rumors and conspiracies in the U.S., we replicate Allcott and Gentzkow (2017) who analyze consumption of fake news and its relation to beliefs about conspiracy theories, Berinsky (2017) who considers misinformation surrounding national health care policy, Clayton et al. (2021) who present results from survey experiments on population endorsement of democratic norms, Oliver and Wood (2014) who examine conspiracy theories and paranoia, and Jamieson and Albarracin (2020) who consider misinformation and media consumption about the COVID-19 pandemic.[4]

In China, we ask questions similar to Cui and Shoemaker (2018) who study attitudes toward conspiracies and facts about genetically modified food in China. We also ask questions related to some U.S. conspiracies from Oliver and Wood (2014) to the China sample.

The studies we replicate fielded different question designs from binary true/false response options to agree/disagree with between two and seven categories. We replicated the original response options in our surveys but for presentation here we compute for each item the percentage of responses indicating agreement, meaning "true," "agree," or "yes," with or without qualification ("definitely" as well as "probably").

The circled dots in Figure 1 represent results for the rumor and conspiracy items (see also Appendix Figure A1). The difference between percent agreeing to positive- and negative-keyed versions varies across questions from 43 points on the Obama born abroad question fielded to the China 2021 sample to -12.5 points on the Oliver and Wood (2014) question on government planning of the 9/11 terrorist attacks.[5] All but this one difference is positive. In general, the questions in the China sample had greater levels of acquiescence bias than the U.S. sample, which could be partially due to the fact that many of the conspiracies asked were related to the U.S. and may not be as familiar to the respondents.[6]

In the U.S., questions with differences greater than 5 points generally reflect partisan issues such as democratic transitions (which was partisan in December 2020 when the survey was fielded), the Affordable Care Act, the Iraq war, and the birthplace of Barack Obama. The exception is the Oliver and Wood conspiracy about airplane contrails. Smaller differences surrounded conspiracies and rumors about the FBI, Ireland asylum, George Soros, the Clinton Foundation, and the etiology of the coronavirus.

---

[4]Oliver and Wood (2014, 957) discuss the potential for acquiescence bias to influence their results.

[5]We suspect the latter result might be due to poor phrasing in negation on our part.

[6]Some additional evidence of this is there was a high non-response rate to the questions fielded in China.

On some questions, the divergent estimates from the original versus negated version of the question are of political importance. In the original question of Allcott and Gentzkow, our sample indicates more than half (53 percent) believed "At the third presidential debate, Donald Trump refused to say whether he would concede the election if he lost" versus less than one third (32 percent) in the negated version. In the original question of Clayton et al., our sample indicates that 80 percent of the public believes "Presidential candidates should accept the outcome of elections even if they narrowly lose" while only 52 percent disagree that "Presidential candidates need not accept the outcome of elections if they narrowly lose." The percent who agreed that "Sometimes regular people need to be a little violent to make sure votes are counted correctly" was double (24 versus 12 percent) the percent who disagreed that "Regular people DO NOT need to be a little violent to make sure votes are counted correctly."[7]

2.2 Politically-relevant facts

The triangles in Figure 1 represent results on what we categorize as politically-relevant facts (see also Appendix Figure A2). In addition to the five studies above, we also include questions from Nyhan and Reifler (2010) on weapons of mass destruction in Iraq and from Hopkins, Sides, and Citrin (2018) on beliefs about the size of the immigrant population in the United States. Here, positive-keyed is when the agreeable response matches the truth of the fact.

As with the conspiracies and rumors, differences in endorsement on politically-relevant facts varies considerably across items from 44 to less than 2 points. All differences are positive. In some cases, conclusions scholars might draw about politically-relevant beliefs meaningfully diverge by question wording. Under the original Allcott and Gentzkow wording, our sample indicates 64 percent believed that "On October 28th, the FBI director alerted members of Congress that the FBI had discovered new emails relevant to its investigation of Hillary Clinton's personal server" versus only 22 percent in the negated condition, and 44 versus 11 percent believed that "Two days before the election, the FBI director told Congress that a newer batch of emails linked to Hillary Clinton's private email server changed his conclusion that Clinton should face no charges over her handling of classified information."

Jamieson and Albarracin, p 2 conclude from their study on beliefs about the effectiveness of hand washing that, "Because hand washing and social distancing can prevent the spread of respiratory viruses including the flu, the finding that early in March, 87% believed that these practices were preventative signals a success of public health messaging." We find a 20 point difference in this belief between the original and negated version of the question.

2.3 Surveys Eliciting Probabilistic Beliefs

The instruments used by the studies we have so far replicated have each been closed-ended responses with a small number of categories. In this section we document acquiescence bias when beliefs are elicited on a continuous probability scale.

The squares in Figure 1 present results from 16 politically-relevant facts on which we elicited beliefs from survey respondents in China and the U.S. from four surveys spanning years 2016 through 2020 (see also Appendix Figure A3). In contrast to the closed-ended responses (e.g., true vs false) in the two tables above, we elicited subjects' continuous probabilistic beliefs. We asked

---

[7]Some of our readers have reacted that the Clayton et al. (2021) questions read more like opinions than beliefs. We fielded the items thinking about the construct as "beliefs about how democracies operate" but acknowledge the alternative classification is also reasonable.

how likely they believed each fact to be true on a zero to 100 scale with incentives for an accurate report of their probabilistic beliefs.[8] See Appendix Sections B.4, B.5, and B.6 for details on the samples in these studies.

As with the closed-ended instrument, we find variable but in some cases large differences in reported probabilities by question wording. In China, differences vary from 9 points on ZTE Fine to 27 points on China's currency becoming part of the IMF currency basket. In the U.S., differences vary from 6 points on beliefs about GDP growth in 2014 to 46 points on beliefs about the size of cuts to food stamps in the Trump Administration's 2018 budget proposal (this difference from the TESS 2017-18 survey; on the Lucid 2020 survey, the difference was 32 points). All differences are positive.

These results show that acquiescence bias can be of substantive importance for conclusions (a) about population beliefs about objective facts, (b) from instruments that elicit beliefs as probabilities rather than closed-ended categorical responses, and (c) does not seem to be resolved by the use of incentives.

To summarize the results of this section, we have shown that, while acquiescence bias does not always influence inferences about conspiratorial or factual beliefs, in many cases scholars would come to substantively different conclusions about population beliefs if they asked a positive- rather than a negative-keyed question. These biases exist across instruments of measurement, source of population samples, topics, and in both the U.S. and in China.

3 Correlates of conspiracy beliefs and acquiescence bias

In addition to the incidence of conspiratorial or factual beliefs in the population, scholars are interested in the correspondence between such beliefs and citizen characteristics. For example, Allcott and Gentzkow (2017) report that having beliefs aligned with the partisan or ideological implications of a news headline corresponds to greater endorsement of the veracity of that headline. We replicate the Allcott and Gentzkow (2017) finding first with responses from our subjects assigned the original version of the question. We then compare conclusions with analysis of negative-keyed questions.

Table 1 presents OLS regression results of the relationship between subject ideological alignment with the news headline (as coded by the original authors) and endorsement of the news headline, for both positive-keyed and negative-keyed versions. As before, we recode responses to negative-keyed versions so that "agreement" indicates agreement with the conspiratorial headline.

The negative-keyed questions (columns two and four) suggest substantively smaller correlation between ideological alignment and reported belief. The magnitude of decline is especially large among the subset of news headline described by Allcott and Gentzkow (2017) as "Big Fake," fake news stories that were mentioned in at least three mainstream media articles. Among these stories, the relationship to ideological alignment is 45% smaller with the negative-keyed question.

The cause of the divergent correlations in Table 1 is that acquiescence bias varies substantially between groups. In Table 2 we present regressions pooling positive- and negative-keyed questions and interacting an indicator of positive-keyed – recall key is assigned at random a the question level – with ideological alignment and with characteristics of the subject. We find substantively large interactions with ideological alignment. Questions aligned with respondent ideology exhibit

---

[8]The Hopkins, Sides, and Citrin questions asked subjects the population percentage foreign-born in the U.S. so also on a 0-100 scale but not a probability.

	Dependent variable: Agreement with the Conspiratorial Headline			
	Pos Keyed	Neg Keyed	Pos Keyed, Big Fake	Neg Keyed, Big Fake
	(1)	(2)	(3)	(4)
Aligned	0.094***	0.069***	0.150***	0.083***
	(0.008)	(0.009)	(0.011)	(0.012)
Constant	0.550***	0.410***	0.380***	0.380***
	(0.006)	(0.006)	(0.008)	(0.009)
Observations	8,411	7,020	4,262	3,445
R^2	0.015	0.008	0.040	0.013
Adjusted R^2	0.015	0.008	0.039	0.012

Note: *p<0.1; **p<0.05; ***p<0.01

Table 1: Correlation between aligned ideology and belief in news headlines, positive-keyed versus negative-keyed items. Left two regression show results for all news headlines, right two regression show results only for items labeled by Allcott and Gentzkow (2017) as "Big Fake."

more acquiescence bias especially for "Big Fake" statements.

In Table 3 we evaluate how acquiesence bias affects the correlation between ideology and belief in a conspiratorial fact for the questions we replicated from Oliver and Wood (2014). Because ideological alignment was not coded in that work, we interact indicators of very conservative, very liberal, and other covariates with positive-keyed. The interaction term between very conservative and positive-keyed is large and statistically significant in all cases, indicating larger acquiescence bias for this subgroup. In many instances, including conspiratorial headlines related to Iraq, 9-11, and Contrails, the researcher would come to a different conclusion about the correlation between ideology and conspiratorial thinking had they fielded the positive-keyed question instead of the negative-keyed question.

In Appendix Table A5, we evaluate acquiescence bias as in Tables 2 and 3 for Jamieson and Albarracin (2020). We find similar results. The magnitudes of correlation between endorsement of conspiracies and covariates and between misperception of facts and covariates depends upon whether the question was fielded as original or negated. In particular, there are consistently interactions of the positive-keyed version of the question and very conservative respondents, very liberal respondents, numeracy, and age.

In sum, for each of the studies that we replicate we find that magnitude of acquiescence bias varies by characteristics of the respondent. Acquiescence bias for these questions is more common among very conservative subjects, very liberal subjects, younger subjects, those with innumeracy, and those with lower education.

Given these correlations with individual characteristics, acquiescence bias would cause inaccurate estimates for scholars interested in both the incidence of and the correlation with characteristics of conspiratorial or factual beliefs. We next turn to suggesting fixes for both problems.

	Dependent variable: Agreement with the Conspiratorial Headline			
	All	Big Fake	All	Big Fake
	(1)	(2)	(3)	(4)
Aligned	0.069***	0.083***	0.069***	0.083***
	(0.009)	(0.013)	(0.009)	(0.012)
Pos Keyed	0.140***	–0.006	0.170***	0.140***
	(0.009)	(0.012)	(0.019)	(0.025)
Numeracy			–0.005	–0.022***
			(0.006)	(0.008)
Age			–0.0002	–0.001***
			(0.0003)	(0.0003)
Education			0.00002	0.00002
			(0.00002)	(0.00002)
Pos Keyed*Aligned	0.024**	0.068***	0.025**	0.069***
	(0.012)	(0.017)	(0.012)	(0.017)
Pos Keyed*Numeracy			–0.002	–0.007
			(0.008)	(0.011)
Pos keyed:Age			–0.001**	–0.003***
			(0.0003)	(0.0005)
Pos Keyed:Education			–0.0001**	–0.00003
			(0.00002)	(0.00003)
Constant	0.410***	0.380***	0.420***	0.450***
	(0.006)	(0.009)	(0.014)	(0.018)
Observations	15,431	7,707	15,419	7,701
R^2	0.046	0.029	0.048	0.060
Adjusted R^2	0.046	0.029	0.047	0.059

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2: The impact of Allcott and Gentzkow (2017) question wording on belief varies by ideological alignment. Left regression is for all statements, right regression is only for "Big Fake" statements.

	Iraq	9-11	Obama	Contrails	Soros
	Dependent variable: Agreement with Conspiracy				
	(1)	(2)	(3)	(4)	(5)
Pos Keyed	1.000***	1.400***	0.580***	0.860***	0.530***
	(0.150)	(0.160)	(0.180)	(0.150)	(0.160)
Very Conservative	−0.420***	−0.085	0.850***	0.068	0.630***
	(0.120)	(0.120)	(0.130)	(0.120)	(0.130)
Very Liberal	0.050	−0.170	−0.440***	−0.190*	−0.540***
	(0.110)	(0.120)	(0.130)	(0.110)	(0.120)
Numeracy	0.005	−0.010	−0.170***	−0.200***	−0.074
	(0.049)	(0.050)	(0.053)	(0.047)	(0.053)
Age	−0.004	0.018***	0.0005	−0.016***	0.004*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Education	−0.0001	−0.0001	−0.0001	0.0001	0.0002
	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0002)
Pos Keyed*Very Conservative	0.450***	0.330*	0.410**	0.360**	0.500***
	(0.170)	(0.170)	(0.190)	(0.170)	(0.180)
Pos Keyed*Very Liberal	0.340**	0.560***	0.084	0.150	0.270*
	(0.150)	(0.160)	(0.180)	(0.150)	(0.160)
Pos Keyed*Numeracy	−0.160**	−0.120*	0.029	−0.024	−0.091
	(0.067)	(0.070)	(0.077)	(0.066)	(0.070)
Pos keyed*Age	−0.017***	−0.043***	−0.011***	−0.011***	−0.013***
	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
Pos Keyed*Education	0.0001	−0.0001	−0.0001	−0.00004	−0.0004*
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Constant	−0.290***	−0.690***	−0.640***	−0.230**	−0.230**
	(0.110)	(0.110)	(0.120)	(0.110)	(0.120)
Observations	1,962	1,864	1,847	1,844	1,845
R^2	0.093	0.160	0.094	0.170	0.092
Adjusted R^2	0.088	0.150	0.088	0.170	0.086

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3: Acquiescence bias in correlations, Oliver and Wood (2014)

4 Accounting for Acquiescence Bias

Fielding positive- and negative-worded versions of 53 questions eliciting beliefs shows that acquiescence bias influences survey estimates of three quantities of interest to political scientists: rates of endorsement of conspiracies, rates of political misperceptions, and correlations between features of individuals and their beliefs. In this section, we present two simple statistical procedures to estimate (a) population beliefs controlling for acquiescence bias and (b) correlations purged of acquiescence bias.

Our methods correct sample data to provide an acquiescence-free estimate of population quantities. We show how random assignment of positive- and negative-keyed versions of each item allows for estimation of the magnitude of and a correction for bias.

Consider a conspiratorial or factual statement about the world, for example that Barack Obama was born in the United States. Let the random variable Y_i represent the response by subject i to a measurement instrument – e.g., a survey question – used to measure subject i's belief about the statement. Let Y_i^* be the subject's belief about the statement could it be elicited without acquiescence bias.

Our replication studies strongly suggest that the response Y_i is perturbed by acquiescence bias on many conspiratorial and factual statements of interest such that $E(Y_i) \neq Y_i^*$. To characterize the bias $E(Y_i - Y_i^*)$, define $Y_i(p)$ and $Y_i(n)$ the responses of subject i were she asked the positive-keyed (p) or negative-keyed (n) version of the measurement instrument. $Y_i(p)$ and $Y_i(n)$ should be coded so that increasing values indicate increasing endorsement of the belief, not increasing agreement with the question asked. For example, if question version p is "True or False: Barack Obama was born in the United States" and version n is "True or False: Barack Obama was not born in the United States," and if $Y_i(p) = 1$ represents response "True" and $Y_i(p) = 0$ represents the response "False," then $Y_i(n) = 1$ would represent "False" and $Y_i(n) = 0$ would represent "True." For both versions, the original response "True" is the agreeable response possibly subject to acquiescence bias.

With these definitions, the data-generating process of subject i measured response Y_i is

$$Y_i(p) = Y_i^* + \delta_i,$$
$$Y_i(n) = Y_i^* - \delta_i \tag{1}$$

with δ_i acquiescence bias for subject i on that question.

Let D_i represent the version of instrument fielded such that $D_i = 1$ if i responds to the positive-key version p and $D_i = -1$ if i responds to the negative-key version n. We then consolidate (1) into

$$Y_i(D_i) = Y_i^* + D_i \delta_i. \tag{2}$$

The observed response $Y_i(D_i)$ is the acquiescence-free belief Y_i^* plus (when $D_i = 1$) or minus (when $D_i = -1$) the subject's acquiescence-bias δ_i. Given data-generating equation (2), consider an ordinary least-squares (OLS) regression of Y on D with a representative sample from the target population

$$Y_i = \alpha + \beta D_i + \varepsilon_i$$

with ε_i an independent and identically distributed error term. When the expected value of ε_i is zero, the coefficient α estimates the population average belief $E(Y_i^*)$ purged of acquiescence bias.

14

The coefficient β estimates the population average acquiescence bias $\Delta = E(\delta_i)$ under the usual assumptions for ordinary least-squares regression. This regression might also use post-stratification survey weights to target a population-level average in the presence of individual heterogeneity (assuming random assignment of D).

4.1 Extending to covariates

In addition to population rates, scholars of political conspiracies and beliefs want to learn about correlations between characteristics of individuals and their beliefs. For example, Oliver and Wood (2014) and Bullock et al. (2015) examine the relationship between ideology and political beliefs and Pennycook and Rand (2019) argue that non-analytical reasoning causes endorsement of conspiracy theories. Our replication results presented in Tables 2 and 3, however, show that correlations can diverge between positive- and negative-keyed versions of a question.

Consider the data-generating process in Eq. 2. To model the relationship between beliefs and characteristics of the subject, substitute the quantity Y_i^* with the linear combination $x_i'\gamma$, x a k-vector of covariates and γ a k-vector of coefficients. Thus γ represents the correspondence between individual characteristics x_i and acquiescence-free beliefs Y_i^*. This substitution leads to the updated data-generating process of Y_i

$$Y_i(D_i) = x_i'\gamma + D_i\delta_i. \tag{3}$$

To see the problem of acquiescence bias for estimating correlates of political beliefs, consider a sample of size $\mathcal{N}$ from the population of interest. Define the $\mathcal{N}$-vectors Y, D, and δ for the sample values of Y_i, D_i, and δ_i. Define X the $\mathcal{N}$-by-k matrix of k covariates. The OLS estimate of γ, $\hat{\gamma}$ is

$$\hat{\gamma} = (X'X)^{-1}X'Y.$$

Substituting the known data-generating process of Y we have

$$\hat{\gamma} = (X'X)^{-1}X'(X\gamma + D\delta)$$
$$= \gamma + (X'X)^{-1}X'D\delta. \tag{4}$$

The OLS estimator of the k correlations of interest γ is biased by the second term in (4), the covariance between x_i and δ_i. Covariance between x_i and δ_i was exactly the problem diagnosed by the *The American Voter* discussed in Section 1; because studies had fielded only positive-keyed versions of questions measuring authoritarianism, correlation between acquiescence bias and education inflated estimates of the relationship between authoritarianism and education. Tables 2 and 3 in Section 3 indicate that there may also be a correlation between acquiescence bias and conservatism, which could inflate correlations between conservatism and belief in conspiracies. In other words, if D_i equals one for all subjects because only a positive-keyed item is fielded, and if the quantity $(X'X)^{-1}X'\delta$ is non-zero, $\hat{\gamma}$ is biased.

Equation 4 suggests two options to ameliorate bias in $\hat{\gamma}$. In each, the analyst assigns at random positive- and negative-keyed versions of each item. In the first solution, the analyst can regress Y_i on covariates of interest x_i, a variable D_i taking the values 1 and -1 indicating version positive and negative, and an interaction between D_i and x_i:

$$Y_i = \beta_0 + \beta_1 x_i + \beta_2 D_i + \beta_3 D_i x_i + \epsilon_i$$

The coefficient on x_i, β_1, estimates the correlation between x_i and Y_i^* free of acquiescence bias. The coefficient β_2 is an estimate of average acquiescence bias and the coefficient(s) β_3 estimate heterogeneity of acquiescence bias related to x_i. As before, one could estimate this regression with post-stratification survey weights.

Second, the analyst can implement weighted least-squares (WLS) without including the variable D_i directly into the regression. When the vector D is perfectly balanced – when the number of positive- and negative-keyed responses is exactly equal – its expected value is zero because one $D_i = 1$ offsets each $D_i = -1$. When the vector D is assigned at random, the covariance of D_i and δ_i is zero. When the covariances of D_i and δ_i is zero, the expected values of their product, $D_i\delta_i$, is zero. Therefore, the expected value of the bias terms in (4) is zero when D is balanced and assigned at random.

To achieve exact balance on D, define the matrix W with elements on the diagonal $\frac{\mathcal{N}}{2N}$ for the N subjects asked question version p and $\frac{\mathcal{N}}{2\tilde{N}}$ for the $\tilde{N}$ subjects asked question version n, off-diagonal cells zero, $N + \tilde{N} = \mathcal{N}$. Then

$$\hat{\gamma}_{\text{WLS}} = (X'WX)^{-1}X'WY \qquad (5)$$

is an unbiased estimate of γ. If the data set includes post-stratification weights, the values N and $\tilde{N}$ should be the sum of the weights in each sample rather than the count of observations and the diagonals of the weight matrix W should be the product of the terms above and the original survey weight. With survey weights, the regression method is likely to be easier for scholars to implement.[9]

We emphasize that these fixes depend upon the untested assumption that acquiescence bias does not vary by key. Because of this assumption, we recommend that researchers present analysis separately by version of the question in addition to implementing the proposed solutions.

5   Applying the method

We apply the weighted least squares approach to examine the correlation between ideology and belief in conspiracy theory in each of the studies. For each question, we calculated the weights for positive-keyed and negative-keyed responses as $\frac{\mathcal{N}}{2N}$ and $\frac{\mathcal{N}}{2\tilde{N}}$ and use weighted least squares. Because of the large number of items, we present coefficient plots below. While we plot and discuss the correlation between conspiratorial belief and the subject identifying as very conservative in the main body, we apply the method to coefficients for very liberal, numeracy, age, and education and present full regression tables in Appendix Section E.

We plot coefficients and 95 percent confidence intervals on the *very conservative* indicator for each question for (1) the sample assigned the positive-keyed version of the question, (2) the sample assigned the negative-keyed version of the question, and (3) the weighted least squares method combining the two (point with solid line). Figures 2, 3, and 4 present the questions we replicated from Allcott and Gentzkow (2017), Jamieson and Albarracin (2020), and Oliver and Wood (2014).

In most cases, there are substantively significant differences between the coefficient estimated in the positive- versus negative-keyed samples. For example, Figure 3 shows that with the positive-key version of the question, one would conclude that identifying as conservative is significantly

---

[9]A reader has noted a third method: to run two separate regressions on the positive- and negative-keyed samples then use the two sets of coefficient estimates to construct precision-weighted averages.

correlated with false statements about hand washing and COVID-19. But if the scholar fielded the negative-keyed version, the coefficient estimate is less than half the magnitude and not statistically distinguishable from zero.

Across the three figures, multiple questions exhibit patterns similar to the hand-washing item. Point estimate differences and differences in range of confidence intervals show that scholars would come to different conclusions with the two versions of the question.

6 Alternative: Multiple-choice response options

In this section, we evaluate an alternative instrument to elicit beliefs that attempts to remove the problem of acquiescence bias. We term this instrument "multiple-choice" because the instrument provides the respondent a set of concrete responses from which to choose rather than querying agreement with a statement in the question text (as advocated in Clifford, Kim, and Sullivan, 2019). The multiple-choice version is, presumably, not subject to acquiescence bias when no response option appears agreeable.

Although in some cases multiple-choice might be a better instrument, multiple-choice adds complications that might lead other scholars to prefer agree/disagree items. First, not all questions have an obvious multiple-choice instrumentation for a scholar's construct of interest, particularly if interest lies in degree of endorsement. Second, multiple-choice might be more cognitively demanding for the subject than agree/disagree items.

Third, multiple-choice instruments might depend more than agree/disagree on the set of response options. For example, the original Berinsky (2017) question asked "True of False: Changes to the health care system enacted by Congress and the Obama administration created 'death panels' which have the authority to determine whether or not a gravely ill or injured person should receive health care based on their 'level of productivity in society'?" We created a multiple-choice version which asked "Which component was part of the Affordable Care Act passed by Congress in 2010? (select one)" with response options "Prohibited insurers to deny coverage to those with preexisting conditions," "Established 'death panels' which have the authority to determine whether or not a gravely ill or injured person should receive health care based on their 'level of productivity in society,'" and "Nationalized healthcare, disallowing private insurance companies from continuing to provide coverage."

For subjects who might be answering this multiple choice version of the *death panels* question with cheerleading rather than accuracy motives (Bullock et al., 2015), it might be that the "nationalized healthcare" response provided greater expressive benefit than the "death panel" response. Indeed, we find that 19.7 percent of subjects who endorsed the rumor of death panels in the original true/false question wording later selected the nationalize response to the multiple-choice version.

With no absolute benchmark, it is difficult to know which method most effectively captures beliefs. Scholars wanting to get a best estimate might benefit by fielding multiple methods. That said, in our sample we find individual-level inconsistencies between conspiracy beliefs elicited with the multiple-choice and agree/disagree instruments.

To illuminate the challenge, we compare estimates of population endorsement of three conspiracies/rumors in Figure 5. We present in the darker bars the percent of respondents selecting the conspiratorial belief with a multiple-choice instrument and the lighter bars the WLS-adjusted
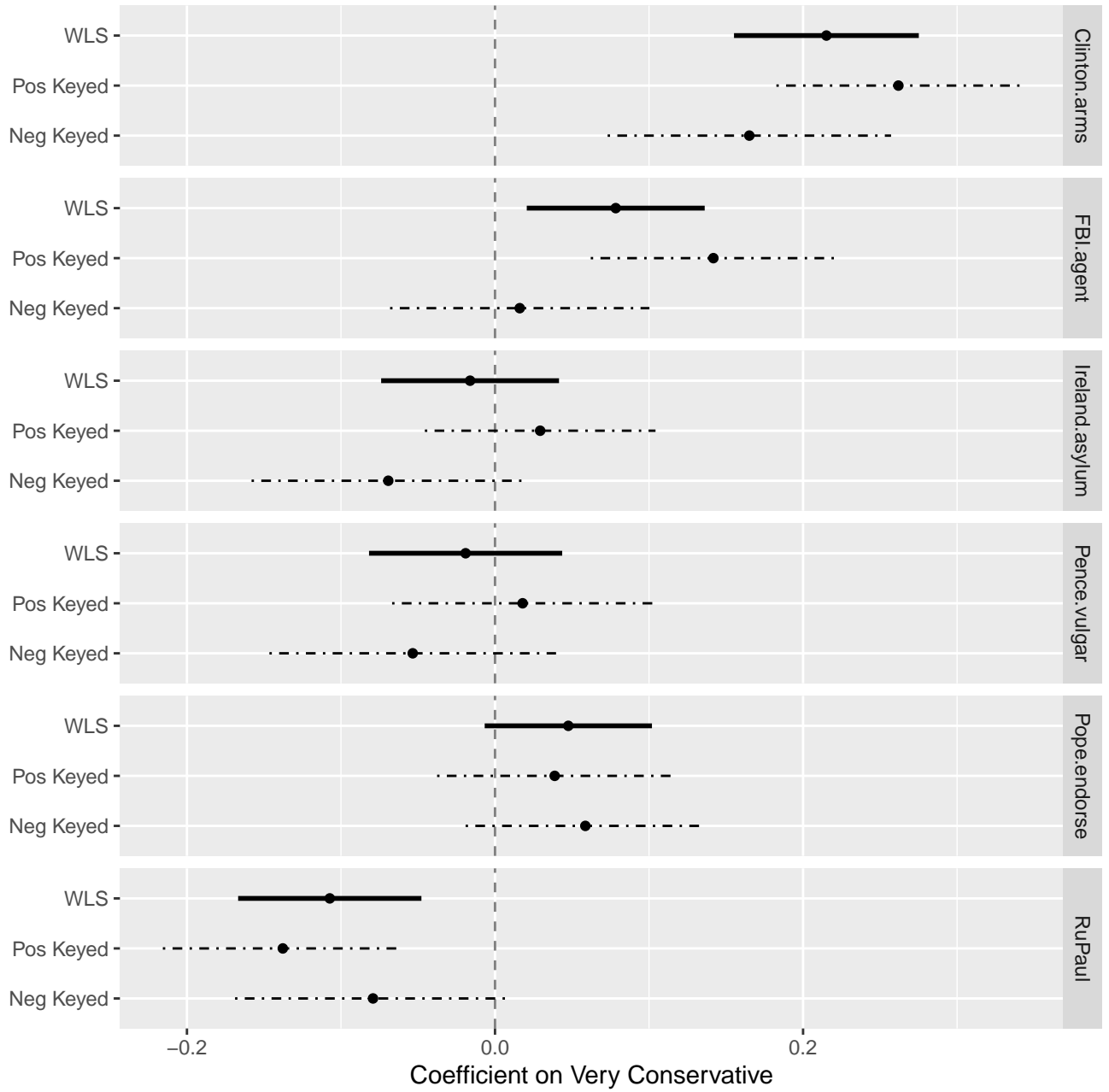
Figure 2: Coefficient indicating the correlation between very conservative and belief in the conspiracy theory, Allcott and Gentzkow (2017)
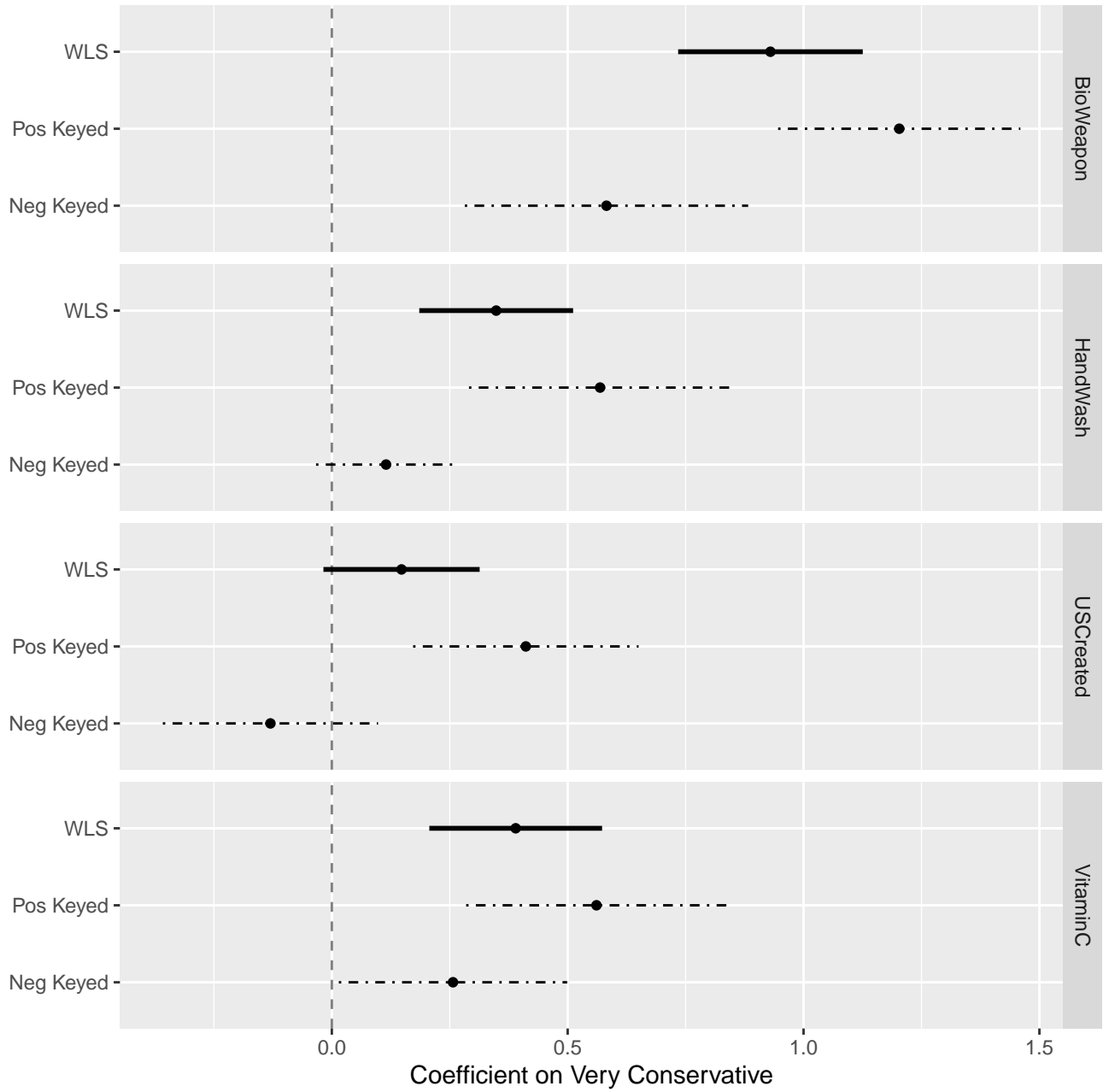
Figure 3: Coefficient indicating the correlation between very conservative and belief in the conspiracy theory, Jamieson and Albarracin (2020)
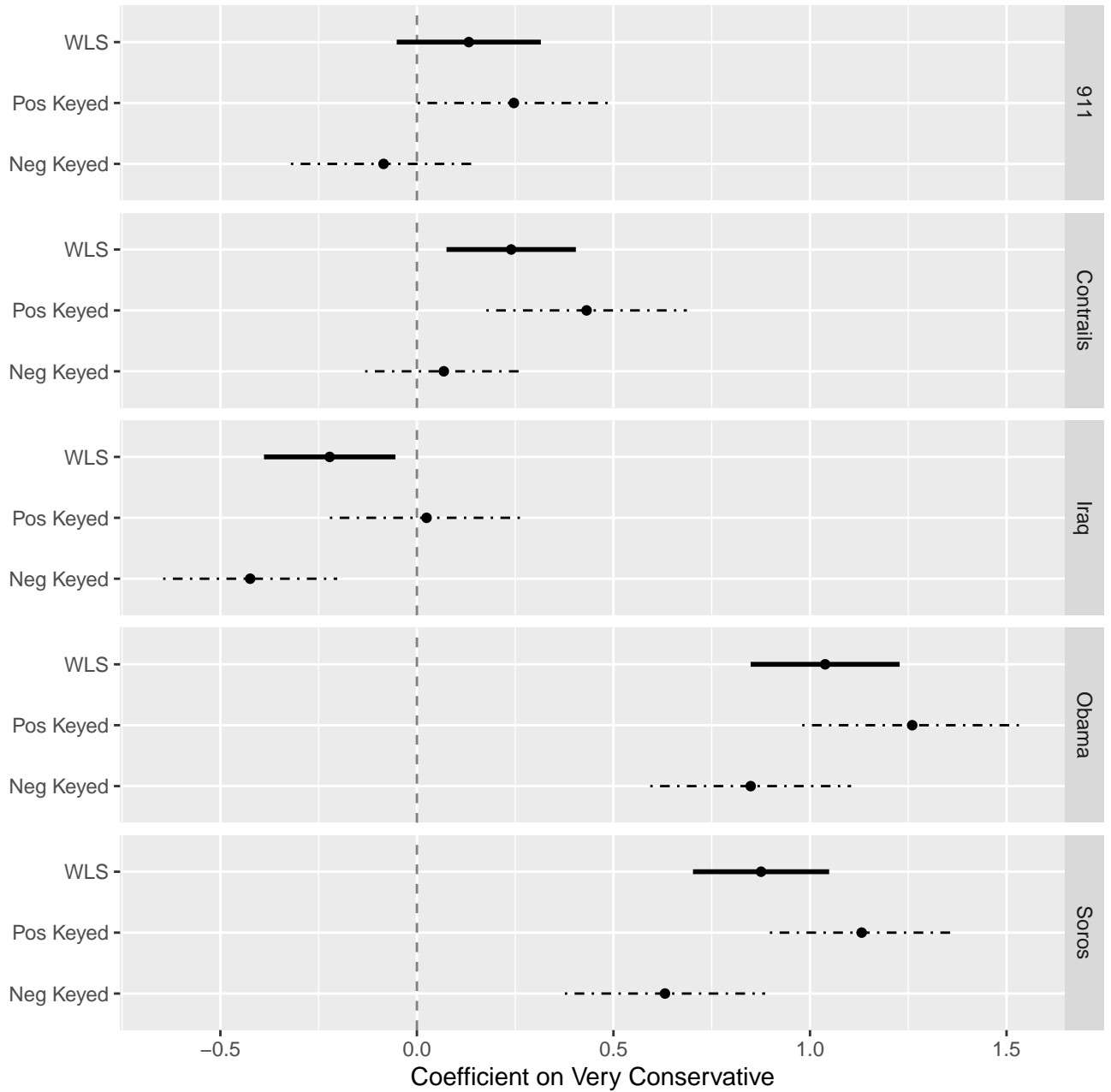
Figure 4: Coefficient indicating the correlation between very conservative and belief in the conspiracy theory, Oliver and Wood (2014)
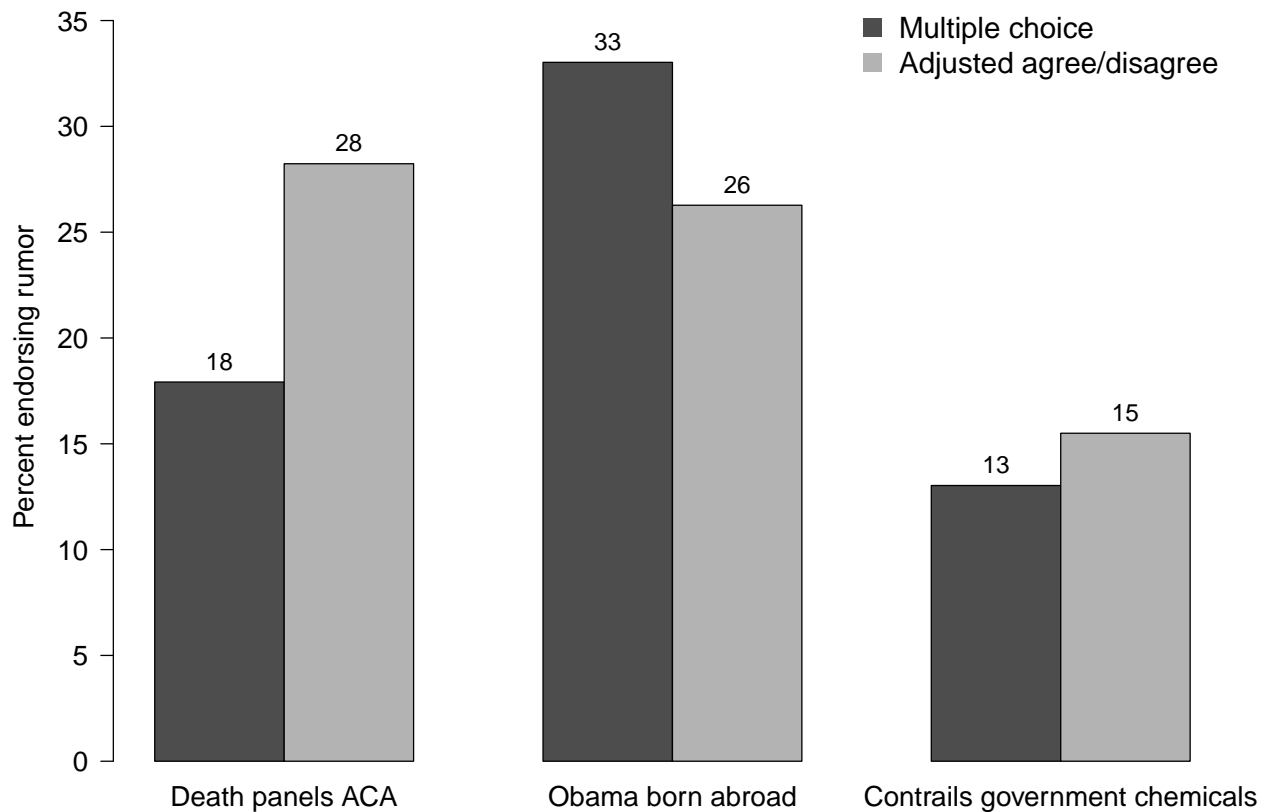
Figure 5: Endorsement of political rumors, multiple-choice versus adjusted agree/disagree instrument. Bar heights extend to weighted sample estimates of population percent endorsing rumor. Agree/disagree adjusted with estimator.

estimate from the agree/disagree version.[10] The results show that neither instrument consistently delivers lower rates of conspiratorial beliefs. The multiple-choice instrument suggests fewer people endorse believe in death panels but that more believe Barack Obama born abroad. Multiple-choice responses do not appear to provide an unalloyed benchmark of beliefs.

7   Discussion: More and less bias?

We did not design our studies to understand variation in the magnitude of acquiescence bias by features of the question wording, topic, or individual. This might be a promising path for future research. With the variety of topics in our studies, however, we offer here brief speculation on what factors appear related to the magnitude of bias.

First, we see some evidence of greater acquiescence bias on the most politically-salient items. Among American subjects, the largest differences between positive- and negative-key (Figure 2) appear on the items about accepting the result of a lost election, whether Donald Trump stated that he would concede a lost election, and whether the Affordable Care Act created death panels.

---

[10]We randomized the order of multiple-choice response options for each respondent.

One might have thought that subjects would have strongly-held positions on these matters and so would be *less* likely to err towards acquiescence. Our findings in Table 1 also suggest that stronger ideologues exhibit *more* acquiescence bias on the most politically-salient topics.

Second, for questions of politically-relevant facts, we see some evidence that the more complicated the topic or the greater the length of the question text, the larger the acquiescence bias. The largest differences on facts (Figure 2) occur for a somewhat-complicated question about a Trump budget proposal, actions surrounding the FBI Director's statements about Hillary Clinton's emails in October 2016, and GDP and job growth statistics. This is consistent with the theory that subjects select the agreeable response when they don't know the answer or don't care to exert the cognitive effort required.[11] This might be consistent with our observation of large amounts of acquiescence bias in the survey fielded about U.S. conspiracies in China where respondents might not have been familiar with the content.

Third, we speculate that the more far-fetched or unusual the conspiracy, the larger the acquiescence bias. The Oliver and Wood (2014) conspiracies about September 11th and airplane contrails exhibit relatively large acquiescence bias. This pattern might be evidence of subjects satisficing with the agreeable response or might alternatively indicate that subjects derive some satisfaction from agreeing with unusual claims (trolling). That said, the conspiracy about the motive for the invasion of Iraq exhibits little bias.

We also find that acquiescence bias appears more prevalent among ideologues. This strikes us as something that merits future investigation. What about ideologues would lead to a stronger tendency to acquiesce? One might have thought that stronger political beliefs would lead to less bias as ideologues draw on their ideological worldview to construct responses to survey questions.

8 Recommendations and conclusion

We believe our evidence suggests that acquiescence bias causes substantively important problems for research on political conspiracies and beliefs. We hope to have brought this challenge to the attention of scholars and to have provided knowledge and methods to improve future research.

While we do not believe there is a one-size-fits-all recommendation because each topic of interest will have its own idiosyncrasies, our investigation brings forward a set of suggested practices.

First, we recommend fielding both positive- and negative-keyed instruments to elicit beliefs. Examining differences in estimates and presenting them to readers should be the first practice. We would then, in general, recommend combining results from the two instruments with one of our proposed estimators.

Second, we would encourage scholars to consider multiple-choice or other instruments without agreeable responses. This approach, however, presents the first-order challenge of selecting the appropriate set of response options. Best practice response options is an opportunity for future research.

This second suggestion is consistent with a more general recommendation. If scholars are interested to learn about sensitive or complicated political beliefs, measurement challenges should be of central concern. Best practice in survey research is to pilot, test, and evaluate different measurement instruments. Too often scholars put one question on a survey without testing and validating, and then assume the responses to that particular item accurately reflect the beliefs of

---

[11]That said, we also see large acquiescence bias on the hand-washing question and whether Clinton stumbled at an event in 2016, which both seem to us relatively straightforward questions.

their subjects. We encourage scholars to consider how the choice of instrument might materially influence the inferences they draw about political conspiracies and beliefs.

References

Allcott, Hunt, and Matthew Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives* 31(2): 211–236.

Ansolabehere, Stephen, Jonathan Rodden, and James M. Snyder. 2008. "The Strength of Issues: Using Multiple Measures to Gauge Preference Stability, Ideological Constraint, and Issue Voting." *American Political Science Review* 102(2): 215–232.

Berinsky, Adam J. 2017. "Rumors and Health Care Reform: Experiments in Political Misinformation." *British Journal of Political Science* 47(2): 241–262.

Billiet, Jaak B., and McKee J. McClendon. 2000. "Modeling Acquiescence in Measurement Models for Two Balanced Sets of Items." *Structural Equation Modeling* 7(4): 608–628.

Budak, Ceren, Sharad Goel, and Justin M. Rao. 2016. "Fair and Balanced? Quantifying Media Bias Through Crowdsourced Content Analysis." *Public Opinion Quarterly* 80: 250–271.

Bullock, John G., Alan S. Gerber, Seth J. Hill, and Gregory A. Huber. 2015. "Partisan Bias in Factual Beliefs About Politics." *Quarterly Journal of Political Science* 10(4): 519–578.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. New York: Wiley.

Carlson, Taylor N. 2019. "Through the Grapevine: Informational Consequences of Interpersonal Political Communication." *American Political Science Review* 113(2): 325–339.

Clayton, Katherine, Nicholas T. Davis, Brendan Nyhan, Ethan Porter, Timothy J. Ryan, and Thomas J. Wood. 2021. "Elite rhetoric can undermine democratic norms." *Proceedings of the National Academy of Sciences* 118(23).

Clifford, Scott, Yongkwang Kim, and Brian W. Sullivan. 2019. "An Improved Question Format For Measuring Conspiracy Beliefs." *Public Opinion Quarterly* 83(4): 690–722.

Cui, Kai, and Sharon P Shoemaker. 2018. "Public perception of genetically-modified (GM) food: a nationwide Chinese consumer study." *npj Science of Food* 2(1): 1–8.

Flynn, DJ, Brendan Nyhan, and Jason Reifler. 2017. "The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics." *Political Psychology* 38: 127–150.

Garrett, R. Kelly, and Robert M. Bond. 2021. "Conservatives' susceptibility to political misperceptions." *Science Advances* 7(23): eabf1234.

Gentzkow, Matthew, and Jesse M. Shapiro. 2010. "What Drives Media Slant? Evidence from U.S. Daily Newspapers." *Econometrica* 78(1): 35–71.

Gohse, Tia. 2016. "Half of Americans Believe in 9/11 Conspiracy Theories." https://www.livescience.com/56479-americans-believe-conspiracy-theories.html.

Groseclose, Tim, and Jeffrey Milyo. 2005. "A Measure of Media Bias." *Quarterly Journal of Economics* 120(4): 1191–1237.

Guess, Andrew M, Brendan Nyhan, and Jason Reifler. 2020. "Exposure to untrustworthy websites in the 2016 US election." *Nature human behaviour* 4(5): 472–480.

Hill, Seth J. 2017. "Learning Together Slowly: Bayesian Learning About Political Facts." *Journal of Politics* 79(4): 1403–1418.

Hill, Seth J., and Gregory A. Huber. 2019. "On The Meaning of Survey Reports of Roll Call 'Votes'." *American Journal of Political Science* 63(3): 611–625.

Hopkins, Daniel J., John Sides, and Jack Citrin. 2018. "The Muted Consequences of Correct Information about Immigration." *Journal of Politics* 81(1): 315–320.

Huang, Haifeng. 2017. "A War of (Mis)Information: The Political Effects of Rumors and Rumor Rebuttals in an Authoritarian Country." *British Journal of Political Science* 47(2): 283–311.

Jamieson, Kathleen Hall, and Dolores Albarracin. 2020. "The Relation between Media Consumption and Misinformation at the Outset of the SARS-CoV-2 Pandemic in the US." *Harvard Kennedy School Misinformation Review* 1.

Jerit, Jennifer, and Yangzi Zhao. 2020. "Political misinformation." *Annual Review of Political Science* 23(1): 77–94.

Krosnick, Jon A. 1991. "Response strategies for coping with the cognitive demands of attitude measures in surveys." *Applied Cognitive Psychology* 5(3): 213–236.

Krosnick, Jon A. 1999. "Survey Research." *Annual Review of Psychology* 50: 537–567.

Lazer, David MJ, Matthew A Baum, Yochai Benkler, Adam J Berinsky, Kelly M Greenhill, Filippo Menczer, Miriam J Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild et al. 2018. "The science of fake news." *Science* 359(6380): 1094–1096.

Lumley, Thomas. 2011. *survey: analysis of complex survey samples*. R package version 3.24-1.

National Public Radio. 2020. "Even If It's 'Bonkers,' Poll Finds Many Believe QAnon And Other Conspiracy Theories." https://www.npr.org/2020/12/30/951095644/even-if-its-bonkers-poll-finds-many-believe-qanon-and-other-conspiracy-theories.

Nyhan, Brendan, and Jason Reifler. 2010. "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior* 32(2): 303–330.

Oliver, J. Eric, and Thomas J. Wood. 2014. "Conspiracy Theories and the Paranoid Style(s) of Mass Opinion." *American Journal of Political Science* 58(4): 952–966.

Pennycook, Gordon, and David G Rand. 2019. "Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning." *Cognition* 188: 39–50.

Prior, Markus, and Arthur Lupia. 2008. "Money, Time, and Political Knowledge: Distinguishing Quick Recall and Political Learning Skills." *American Journal of Political Science* 52(1): 169–183.

R Development Core Team. 2015. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Schuman, Howard, and Stanley Presser. 1981. *Questions and Answers in Attidude Surveys*. Academic Press.

Vosoughi, Soroush, Deb Roy, and Sinan Aral. 2018. "The spread of true and false news online." *Science* 359(6380): 1146–1151.

Watson, Dorothy. 1992. "Correcting for acquiescent response bias in the absence of a balanced scale: An application to class consciousness." *Sociological Methods & Research* 21(1): 52–88.

Westwood, Sean, Justin Grimmer, Matthew Tyler, and Clayton Nall. 2021. "American Support for Political Violence is Low.".

Appendix

A Positive- and negative-key question wordings

Tables A1, A2, A3, and A4 present question wordings for all items fielded for the paper. We document details of each individual survey in Section B below.

Table A1: China 2021 question wordings, Chinese

Item	Positive keyed	Negative keyed
GMO Bioterrorism, China	起源于美国的转基因技术是用于对中国进行生物恐怖袭击的科技。	起源于美国的转基因技术不是用于对中国进行生物恐怖袭击的科技。
GMO Increase Yield, China	与传统的杂交技术相比，精密的转基因技术将会增加和保持产量，改善食品的质量并且延长食品的保质期。	与传统的杂交技术相比，精密的转基因技术不会增加和保持产量，也不会改善食品的质量或者延长食品的保质期。
Obama birth, China	美国前总统奥巴马并非真的从美国出生，也没有一个确凿的夏威夷出生证明。	美国前总统奥巴马是从美国出生，并且拥有一个确凿的夏威夷出生证明。
Soros conspiracy, China	金融巨鳄乔治-索罗斯暗地建立了自己的秘密组织，企图破坏美国政府的稳定，操控媒体，从而控制世界。	金融巨鳄乔治-索罗斯没有建立自己的秘密组织，没有企图破坏美国政府的稳定，操控媒体，或控制世界。
9/11 consipracy, China	某些美国的政府官员谋划了"911事件"（2001年9月11日的美国本土袭击），因为他们希望美国借此对中东发动战争。	美国的政府官员没有谋划了"911事件"（2001年9月11日的美国本土袭击），也没有希望美国借此对中东发动战争。

Table A2: China 2021 question wordings, English

Item	Positive keyed	Negative keyed
GMO Bioterrorism, China	GM technologies originated in the USA for the purpose of being directed as bioterrorism against China.	GM technologies originated in the USA not for the purpose of being directed as bioterrorism against China.
GMO Increase Yield, China	Compared with traditional hybrid technologies, precision GM technology will increase and maintain yield, improve food quality and extend food shelf life.	Compared with traditional hybrid technologies, precision GM technology will not increase and maintain yield, improve food quality or extend food shelf life.

Obama birth, China	President Barack Obama was not really born in the United States and does not have an authentic Hawaiian birth certificate.	President Barack Obama was born in the United States and has an authentic Hawaiian birth certificate.
Soros conspiracy, China	Billionaire George Soros is behind a hidden plot to destabilize the American government, take control of the media, and put the world under his control.	Billionaire George Soros is not behind a hidden plot to destabilize the American government, take control of the media, and put the world under his control.
9/11 consipracy, China	Certain U.S. government officials planned the attacks of September 11, 2001, because they wanted the United States to go to war in the Middle East.	U.S. government officials did not plan the attacks of September 11, 2001, because they wanted the United States to go to war in the Middle East.

Table A3: U.S. Lucid 2020 study question wordings

Item	**Positive keyed**	**Negative keyed**
9/11 conspiracy	Certain U.S. government officials planned the attacks of September 11, 2001, because they wanted the United States to go to war in the Middle East.	U.S. government officials tried to prevent the attacks of September 11, 2001. They DID NOT want the United States to go to war in the Middle East.
Accept election loss	Presidential candidates should accept the outcome of elections even if they narrowly lose.	Presidential candidates need not accept the outcome of elections if they narrowly lose.
Beyonce rally	The musicians Beyonce and Jay Z appeared at a rally in support of Hillary Clinton.	The musicians Beyonce and Jay Z DID NOT appear at a rallies in support of Hillary Clinton.
China currency	In August 2015, the Chinese currency, the Yuan, was worth more against the U.S. dollar than it had been in the time period from September 2012 to July 2015.	In August 2015, the Chinese currency, the Yuan, was worth less against the U.S. dollar than it had been in the time period from September 2012 to July 2015.
Clinton Foundation	The Clinton Foundation bought $137 million in illegal arms.	The Clinton Foundation DID NOT buy millions of dollars worth of illegal arms.
Clinton deplorables	Hillary Clinton said that you could put half of Trump's supporters into what I call the basket of deplorables.	Hillary Clinton DID NOT say that you could put half of Trump's supporters into what I call the basket of deplorables.

Item	Positive keyed	Negative keyed
Clinton stumbled	At the 9/11 memorial ceremony, Hillary Clinton stumbled and had to be helped into a van.	At the 9/11 memorial ceremony, Hillary Clinton was healthy and walked herself back to a van.
Contrails conspiracy	Vapor trails left by aircraft are actually chemical agents deliberately sprayed in a clandestine program directed by government officials.	Vapor trails left by aircraft are produced by aircraft engine exhaust or changes in air pressure and have nothing to do with government officials.
Coronavirus weapon	The Chinese government created the coronavirus as a biological weapon. Do you believe this is...	The Chinese government DID NOT create the coronavirus as a biological weapon. Do you believe this is...
Death panels	True of False: Changes to the health care system enacted by Congress and the Obama administration created death panels which have the authority to determine whether or not a gravely ill or injured person should receive health care based on their level of productivity in society?	True of False: Changes to the health care system enacted by Congress and the Obama administration DID NOT create death panels which have the authority to determine whether or not a gravely ill or injured person should receive health care based on their level of productivity in society?
FBI Clinton charges	Two days before the election, the FBI director told Congress that a newer batch of emails linked to Hillary Clinton's private email server changed his conclusion that Clinton should face no charges over her handling of classified information.	Two days before the election, the FBI director told Congress that a newer batch of emails linked to Hillary Clinton's private email server DID NOT change his conclusion that Clinton should face no charges over her handling of classified information.
FBI agent	An FBI agent connected to Hillary Clinton's email disclosures murdered his wife and shot himself.	An FBI agent connected to Hillary Clinton's email disclosures DID NOT murder his wife or shoot himself.
FBI director	On October 28th, the FBI director alerted members of Congress that the FBI had discovered new emails relevant to its investigation of Hillary Clinton's personal server.	In late October, the FBI director kept to himself and DID NOT alert members of Congress that the FBI had discovered new emails relevant to its investigation of Hillary Clinton's personal server.
Hand wash	Regular hand washing and avoiding those showing symptoms DOES NOT help prevent infection with coronavirus. Do you believe this is...	Regular hand washing and avoiding those showing symptoms helps prevent infection with coronavirus. Do you believe this is...

Item	Positive keyed	Negative keyed
Immigrant pop Canada	Out of every 100 people living in Canada, how many do you think were born outside of Canada?	Out of every 100 people living in Canada, how many do you think were born in Canada?
Immigrant pop US	Out of every 100 people living in the United States, how many do you think were born outside of the country?	Out of every 100 people living in the United States, how many do you think were born in the United States?
Iraq WMD	Immediately before the U.S. invasion in 2003, Iraq had an active weapons of mass destruction program and large stockpiles of WMD.	Immediately before the U.S. invasion in 2003, Iraq DID NOT have either an active weapons of mass destruction program nor large stockpiles of WMD.
Iraq conspiracy	The U.S. invasion of Iraq was not part of a campaign to fight terrorism, but was driven by oil companies and Jews in the U.S. and Israel.	The U.S. invasion of Iraq was part of a campaign to fight terrorism, not driven by oil companies and Jews in the U.S. and Israel.
Ireland asylum	In May 2016, Ireland announced that it was officially accepting Americans requesting political asylum from a Donald Trump presidency.	Ireland DID NOT announce that it was officially accepting Americans requesting political asylum from a Donald Trump presidency.
Obama birth certificate	Barack Obama was not really born in the United States and DOES NOT have an authentic Hawaiian birth certificate.	Barack Obama was born in the United States and has an authentic Hawaiian birth certificate.
Pence vulgar	Mike Pence said that Michelle Obama is the most vulgar First Lady we've ever had.	Mike Pence DID NOT say that Michelle Obama is the most vulgar First Lady we've ever had.
Pope endorse	Pope Francis endorsed Donald Trump.	Pope Francis DID NOT endorse Donald Trump.
RuPaul groped	Celebrity RuPaul said that Donald Trump mistook him for a woman and groped him at a party in 1995.	Celebrity RuPaul DID NOT say that Donald Trump mistook him for a woman and groped him at a party in 1995.
Soros conspiracy	Billionaire George Soros is behind a hidden plot to destabilize the American government, take control of the media, and put the world under his control.	Billionaire George Soros is not trying to destabilize the American government, take control of the media, or put the world under his control.
Trump concede	At the third presidential debate, Donald Trump refused to say whether he would concede the election if he lost.	At the third presidential debate, Donald Trump said he would concede the election if he lost.

Item	Positive keyed	Negative keyed
Trump food stamps	The 2018 budget of the Trump administration proposed to cut spending on food stamps (the Supplemental Nutrition Assistance Program) by more than $60 billion over the fiscal years 2018 to 2022.	The 2018 budget of the Trump administration proposed to cut spending on food stamps (the Supplemental Nutrition Assistance Program) by less than $60 billion over the fiscal years 2018 to 2022.
US created coronavirus	The US government created the coronavirus. Do you believe this is...	The US government DID NOT create the coronavirus. Do you believe this is...
Violence for votes	Sometimes regular people need to be a little violent to make sure votes are counted correctly.	Regular people DO NOT need to be a little violent to make sure votes are counted correctly.
Vitamin C	Taking vitamin C can prevent a person from being infected with the coronavirus. Do you believe this is...	Taking vitamin C DOES NOT prevent a person from being infected with the coronavirus. Do you believe this is...

Table A4: U.S. MTurk 2016, TESS 2017, China 2017, and China 2019 study question wordings

U.S. MTurk and TESS	1	In the third quarter of 2014, gross domestic product (GDP) of the United States grew at the fastest quarterly rate since [2003/1998].
	2	The U.S. Economy added [fewer/more] than 45,000 net jobs in May, 2016.
	3	In August 2015, the Chinese currency, the Yuan, was worth [less/more] against the U.S. dollar than it had been in the time period from September 2012 to July 2015.
	4	On January 8, 2012, the length of the day from sunrise to sunset in the city of Doha, Qatar was [less/more] than 11 hours.
China 2017	1	GDP growth in 2015 was [above/below] 6.5%.
	2	Defense spending increased by [more/less] than 10% in 2015.
	3	The renminbi entered the IMF SDR currency basket at a rate [greater/less] than 10%.
	4	In August 2015, the Chinese currency, the Yuan, was worth [less/more] against the U.S. dollar than it had been in the time period from September 2012 to July 2015.
	5	In 2015, the National Development and Reform Commission of China fines Qualcomm for monopoly behavior with a fine [more/less] than 4 billion yuan.
	6	China's GDP growth rate in 2016 was [higher/lower] than the GDP growth rate in 2015.
China 2019	1	In 2017, ZTE agreed to pay a fine of [more than/less than] US$1 billion to the United States.
	2	A survey in 2015 showed that Hong Kong people work [more than/less than] 60 per week on average.
	3	U.S. soybeans exported to China [more than/less than] a quarter of the total U.S. production.
	4	In 2018, there were [more than/less than] 70 fighters participating in the South China Sea parade.
	5	In 2017, the number of Chinese students studying abroad were [more than/less than] 600,000.
	6	In 2016, the Chinese government provided scholarships to [more than/less than] 100,000 international students..

Note: We asked about a fifth fact of the U.S. MTurk sample on monthly changes in retail sales in June, 2015. We believe the statement itself was ambiguous to participants, as on average subjects learned from all four headlines in the wrong (away from truth) direction. We exclude this statement from analysis below.

B   Details of studies

B.1   April 2021 China.

We fielded a survey of 1,058 respondents aged 18 and older, from China, and who passed an attention screener from the online survey platform Qualtrics, beginning March 30, 2020. Respondents were recruited by quota sampling on age, gender, education, and geography to reflect urban residents in China. As part of this survey, respondents were asked about their beliefs in 5 conspiratorial and factual statements relevant to GMOs and U.S. politics. We drew these 5 statements from prominent studies in the U.S. and China. We present both versions of the 5 questions in Chinese and English respectively in Tables A1 and A2.

B.2   December 2020 U.S.

We fielded a survey to 2,055 subjects aged 18 and older, U.S. citizens, and who passed an attention screener from the online survey platform Lucid between December 21 and December 23, 2020. We advertised prominently on the consent screen that our study did not deceive. Prior to creating post-stratification weights, respondents were 55 percent female, 76 percent identified their race as white, 11 percent reported Hispanic ethnicity, 20 percent reported abstaining in the 2020 election, 46 percent voted Biden, and 30 percent voted Trump. We raked post-stratification weights to target marginals to account for differential non-response using the R Development Core Team (2015) package survey (Lumley, 2011). Weights varied in size from 0.2 to 4.

We followed the survey design of Allcott and Gentzkow (2017) in, after consenting to participate, asking subjects to confirm they would provide thoughtful answers. Subjects then proceeded to report their beliefs about each of 29 political conspiratorial or factual statements of fact relevant to politics. We drew these 29 statements from prominent studies in the realm of political conspiracies and facts. We created a negative-keyed version of each question and randomized delivery at the level of subject-question. We randomized question order for the twelve Allcott and Gentzkow (2017). We broke up the factual and conspiratorial questions with batteries that elicited policy attitudes and that measured numeracy. The survey closed with a set of demographic questions. We present both versions of the 29 questions in Table A3.

B.3   March to July 2019 China

We recruited 1,025 survey respondents from China, 18 and older using the online survey platform Qualtrics between March 26 and July 21, 2019. Participants were paid through Qualtrics, with the opportunity to make an additional eight yuan depending on their performance. The sample was about 53% male, 47% female, with a mean age of 42. Like the 2017 sample, we selected six statements of fact about economics and politics released from official sources. We report these statements of facts below.

B.4   December 2017 U.S.

We fielded a study with American subjects December 2017 through January 2018 with NORC, a survey firm at the University of Chicago. Partial funding for this study was provided through Time-Sharing Experiments for the Social Sciences. NORC maintains an online panel that is nationally representative, and delivered to us responses from 1,992 18+ U.S. citizens. Subjects were compensated with standard rewards for participation through NORC, plus an additional $2.00 along with the opportunity to earn $0.10 per round via the crossover scoring mechanism.

The design of the NORC study follows that of the MTurk study but for the following changes. NORC's survey team made edits to our presentation of the incentive rule and elicited probabilities using a graphical slider rather than a text input box. Subjects completed one section of headlines without source, two sections of headlines with source, and a final section where they could choose from which source they would observe a headline. In this section, only two rounds of beliefs were elicited – prior beliefs before seeing the source and beliefs after choosing and observing the source. Each section had only three rounds rather than five.

We fielded three of the five facts used in the MTurk study along with one new fact, and rewords the true/false version of one of the continuing facts. See Table A4 for all facts.

B.5 February to April 2017 China

We recruited 1,109 survey respondents from China, 18 and older using the online survey platform Qualtrics between February 3 and April 21, 2017. Participants were paid through Qualtrics, with the opportunity to make an additional eight yuan depending on their performance. The sample was 56% male, 44% female, with a mean age of 40.

We selected six statements of fact about economics and politics released from official sources. Like the U.S. case, we selected facts that were informative about domestic and international politics. We were interested in using a mix of official, commercial, and international sources and also wanted to select one overlapping fact with the U.S. sample. We used the facts in Table A4 with headlines from four sources: BBC, Xinhua, Global Times, and Nanfang Dushibao. BBC is a Western source, Xinhua is one of the two main official sources of the Chinese Communist Party, Global Times is a official paper that is known to have a nationalistic and anti-Western slant, and Nanfang Dushibao is a commercial Chinese paper in the South and known to be more independent of the government.

B.6 September 2016 U.S.

We recruited 794 subjects aged 18 and older and U.S. citizens from Amazon.com's Mechanical Turk (MTurk) worker platform between September 8 and 12, 2016. Participants were paid a $0.60 flat fee and offered the opportunity to earn bonuses of up to $4.00 depending upon their performance in the experiment, which was advertised to and did take about 15 minutes. The study did not deceive, which was advertised prominently on the consent screen. Respondents were 53 percent female and had an average age of 35, and 48 percent had a four year college degree or more. The sample wasn't overly political – 65 percent reported voting in the 2012 presidential election – but tilted Democratic and liberal, with 54 percent Democrat (including leaners) and 28 percent Republican, and only 19 percent conservative or very conservative.

Upon consenting to participate, subjects first took an IQ-like quiz. They had two minutes to answer up to 15 logic and reasoning questions. They were paid $0.10 for each point of their total score on the quiz, which was the number answered correctly less the number answered incorrectly, skipped questions not counted. The average quiz score was 2.1, with a minimum of -11 and a maximum of 11. Subjects were told that money would not be deducted from the show-up fee for scores less than zero. After the IQ-like quiz, subjects were taught about the main section of the experiment. They were told that they would participate in a section consisting of 25 rounds. For each round won, they would be paid a $0.10 bonus, $0.00 otherwise. In each round, they would be asked to evaluate a difficult factual statement with a number from 0 to 100 that described

how likely they believed the statement to be true.[12] The instructions presented the response as a probability in terms designed to be accessible to those not trained in statistics and then explained how participants would win each round, which was a function of their probabilistic belief through the crossover design. The experiment highlighted at multiple points that the subject's chances of winning would be highest if they accurately reported their probabilistic belief.

After presenting the overview of the section and the mechanism of payment, subjects were instructed that they would evaluate the same factual statement in multiple rounds, and that in some rounds they would receive a signal from the computer about whether or not the statement was true. They were told that the signal from the computer would indicate that the correct answer was true or false, and that this signal would be correct three out of four times on average. They were also told that in other rounds, they may receive a news headline related to the statement of fact. They were told that they might want to change their beliefs in response to signals or headlines.

After the instructions for the section, the subjects played three practice rounds evaluating the factual statement "It rained (more than 0.00 inches of precipitation) in Santa Fe, New Mexico on July 7, 2004." Mimicking the section they would play, in the first round they evaluated the statement without any signal from the computer. In the second round they received signals from the computer and again evaluated the statement. In the third round, they received a headline from a newspaper related to the fact and again evaluated the statement. After the third round, the instructions explained how they would be paid as a function of their response.

Once the practice section was complete, participants then proceeded to the main section for which they were paid based upon their performance. For each of five statements of fact, beliefs were elicited for five rounds. Beliefs were elicited in the first round for each statement prior to the delivery of any signal or headline, measuring the subject's initial belief. In each round subsequent to the first, their previous response was presented for their reference.[13]

The first and fifth section were sections where subjects received signals from the computer as to the truth of the statement. In rounds two through five of these two sections, they received one new (independent) signal in each round from the computer about the statement and reported their (potentially-updated) belief. In each round with a signal, the subject was reminded that the signal would be correct three out of four times. With this design, we observe how subjective beliefs about the statement change over time in response to the noisy signals received.

In the second, third, and fourth sections, subjects received news headlines published on the day of or shortly after the statistic related to the statement of fact in rounds two, three, four, and five. That is, for each statement of fact we collected four news headlines, and these four headlines were delivered in random order to each subject. In the second section, the four headlines were delivered without information about the news source that produced the headline. This allows us to observe learning about each headline when subjects are not yet aware that we might deliver the source of headlines in subsequent rounds. In the third and fourth sections, headlines were presented along with the news source that published it. We thus observe how subjects change their beliefs about the truth of the statement in response to each headline and in response to the same headline along with the news source that published it, with randomized order across subjects.

---

[12] The prompt in each round was "Please tell us how likely you believe this statement is true: [Statement presented]. How likely you believe that the statement is true (for example, 1 if you believe it almost certainly false, 99 if you believe it almost certainly true, 50 if totally unsure): [textbox entry]."

[13] In all rounds, subjects had 20 seconds to evaluate the statement, to limit the option of searching for the truth on the web. After 20 seconds, responses were recorded and they were automatically forwarded to the next round.

All subjects had beliefs elicited about five different statements of fact for five rounds each. All subjects evaluated one statement of fact about an abstract, non-economic, non-political topic, the length of the day in Doha, Qatar on January 8, 2012. Subjects only evaluated this fact in one of the two computer signal sections (the first or fifth section, at random) as we did not have headlines related to this fact. This fact serves as a benchmark of learning about non-economic, non-political facts. Subjects additionally evaluated four other facts. Half of the sample evaluated one fact not presented in this essay for use in a separate research project. This half evaluated three of the four facts for this project in addition to the Doha fact. The other half of the sample evaluated all four of the facts for this project in addition to the Doha fact. Which facts were delivered with computer signals versus headlines versus headlines with news sources were all randomized at the subject level. Finally, we randomized for each statement whether the subjects evaluated a true or a false version of the statement to protect against any global bias toward evaluating statements as true or false. We find little difference on this dimension, and so recode all signals and response in the direction of true.

We selected four statements of fact about economic statistics released from official sources that the news media deemed sufficiently salient to mention in a news story headline. In Section G, we document how we selected facts and news sources for the U.S. studies, with a similar procedure for the China study. Our goal was to select facts that were difficult but similar to economic judgments the public would have to make, particularly in evaluating the economic outcomes under incumbent governments. We sought four facts related to economic releases with an objective value (government-produced statistics or equity market values) that each had headlines relevant to the fact from the same four news sources. We also wanted economic statistics relevant to political knowledge and judgments. After extensive searching over multiple topics and multiple years, we identified the facts (see Table A4 below) with such headlines for four U.S. news sources: the New York Times, USA Today, CNN, and Fox News. While these are not an exact representative sample from the full set of headlines, our belief having read hundreds of headlines related to economic and political news is that these are roughly representative of how these news sources covered such events during this time period.

Finally, after completing the five sections, participants answered a series of survey questions about their demographics, political attitudes, and political behaviors. This includes standard demographics and political questions such as partisanship and ideology. On the final screen, a code was presented to the subjects for them to submit on Mechanical Turk in order to collect any bonuses from the IQ-like quiz responses and the five sections.

C Results Plots

D Acquiescence bias in correlations for additional studies
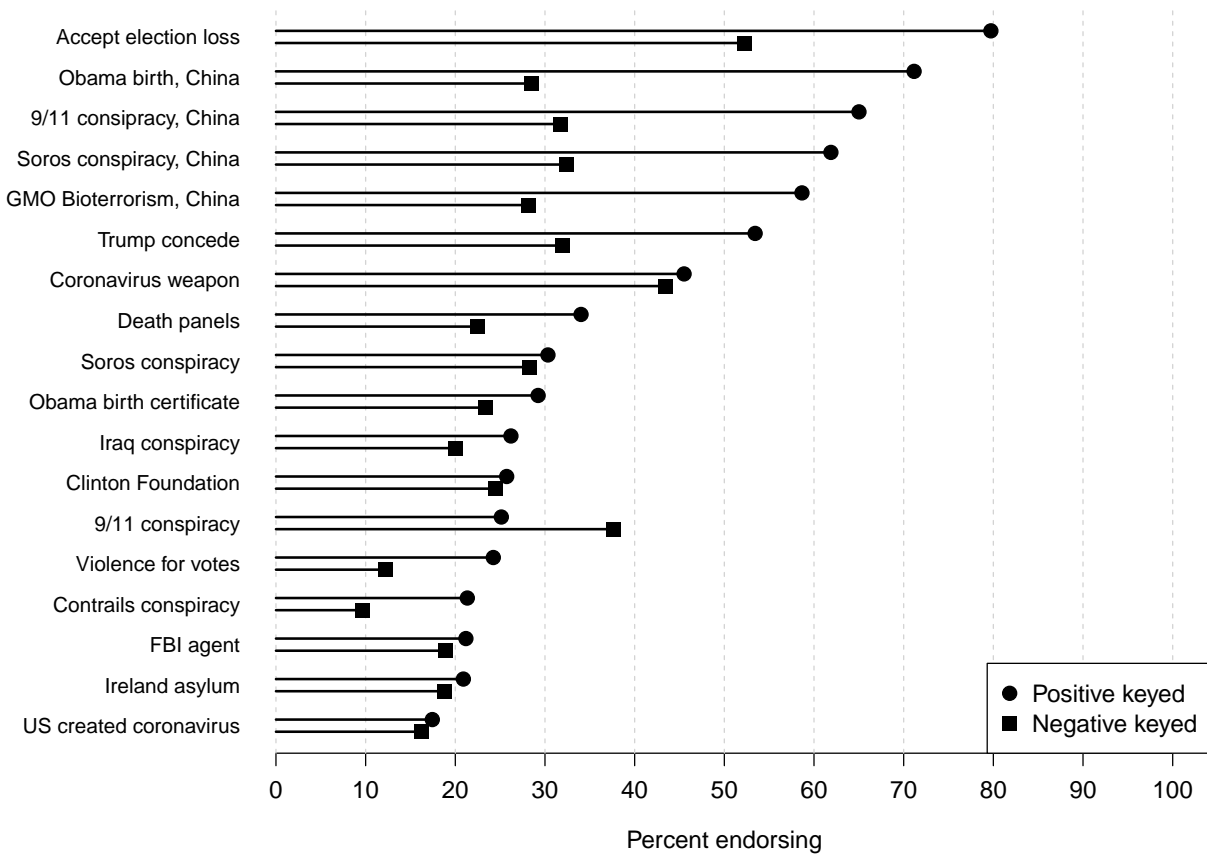
Figure A1: Effect of question wording on reported endorsement of rumors and conspiracies
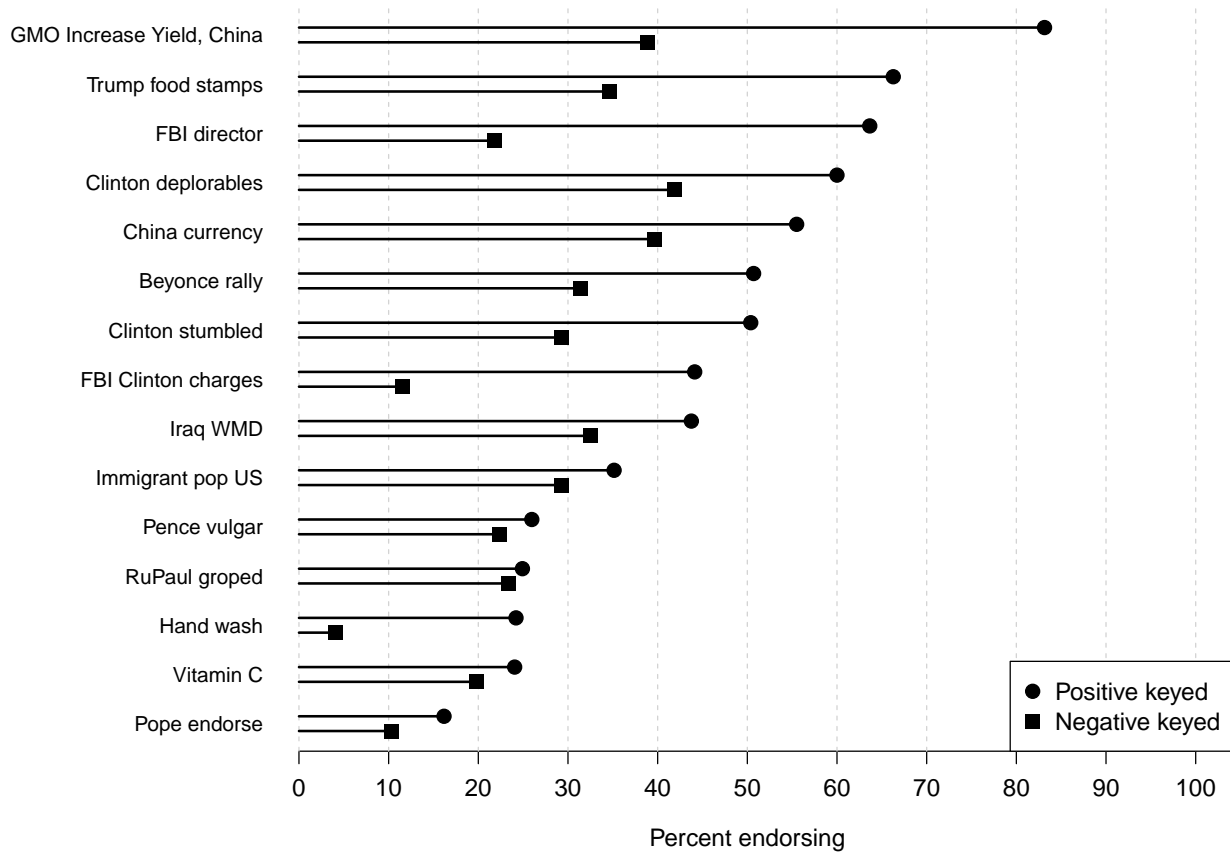
Figure A2: Effect of question wording on reported belief of politically-relevant facts, Replications
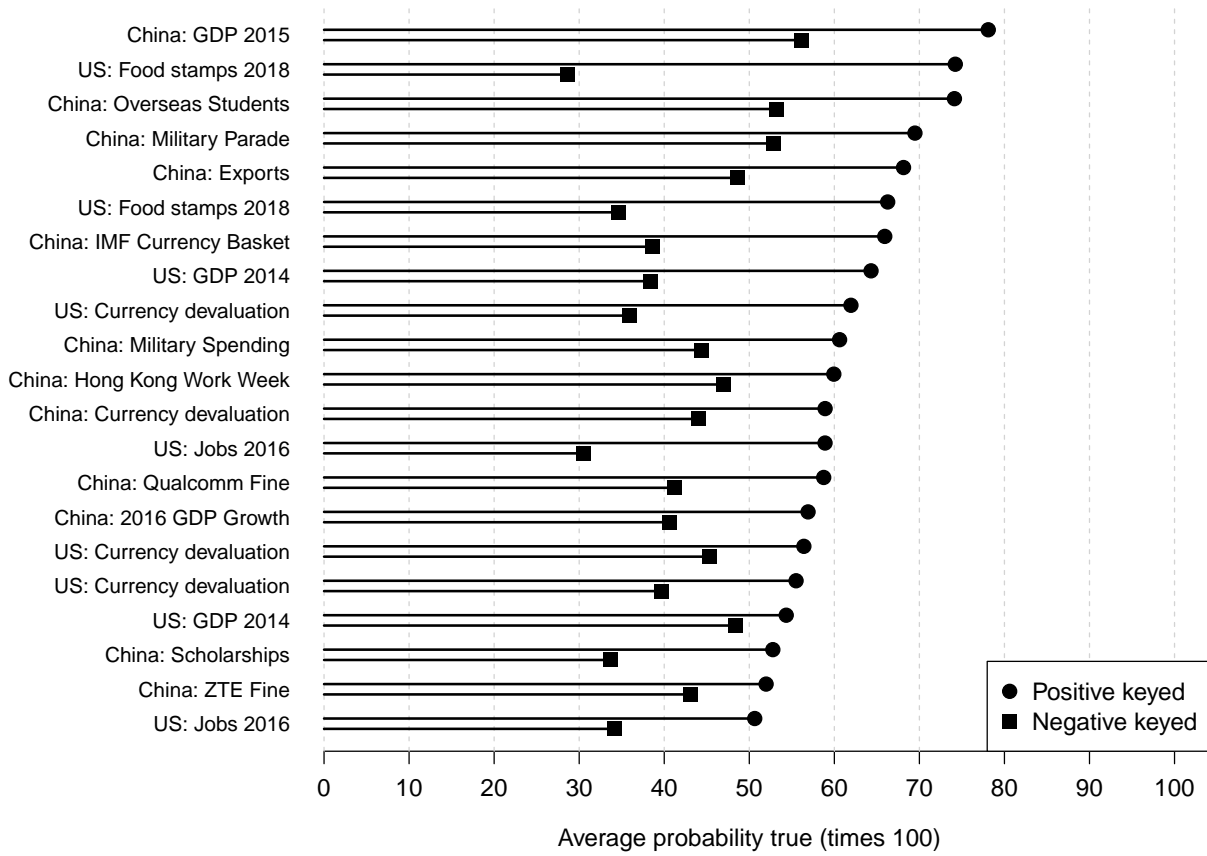
Figure A3: Effect of question wording on reported probabilistic belief in politically-relevant facts, Authors

Table A5: Acquiescence bias in correlations, Jamieson and Albarracin (2020)

	Dependent variable: Agreement with statement			
	Vitamin C	Bio Weapon	Hand Wash	US Created
	(1)	(2)	(3)	(4)
Pos Keyed	0.280*	0.200	0.930***	0.210
	(0.170)	(0.180)	(0.140)	(0.150)
Very Conservative	0.260**	0.580***	–0.570***	–0.130
	(0.120)	(0.150)	(0.110)	(0.120)
Very Liberal	–0.460***	–0.780***	–0.140	–0.330***
	(0.120)	(0.130)	(0.100)	(0.110)
Numeracy	–0.220***	–0.270***	0.220***	–0.230***
	(0.054)	(0.057)	(0.044)	(0.048)
Age	–0.009***	0.003	0.011***	–0.017***
	(0.002)	(0.003)	(0.002)	(0.002)
Education	0.0001	–0.0002	0.0004***	–0.0001
	(0.0001)	(0.0002)	(0.0001)	(0.0001)
Pos Keyed*Very Conservative	0.300	0.620***	0.450***	0.540***
	(0.190)	(0.200)	(0.160)	(0.170)
Pos Keyed*Very Liberal	0.820***	0.520***	0.320**	0.290*
	(0.170)	(0.180)	(0.150)	(0.150)
Pos Keyed*Numeracy	0.130*	–0.007	–0.220***	0.076
	(0.074)	(0.079)	(0.063)	(0.067)
Pos Keyed*Age	–0.009***	–0.008**	–0.007***	–0.006**
	(0.003)	(0.004)	(0.003)	(0.003)
Pos Keyed*Education	0.0002	–0.0001	–0.0004*	0.00001
	(0.0002)	(0.0003)	(0.0002)	(0.0002)
Constant	2.600***	2.900***	1.600***	2.800***
	(0.120)	(0.120)	(0.099)	(0.110)
Observations	2,050	2,052	2,053	2,054
R^2	0.064	0.091	0.110	0.110
Adjusted R^2	0.059	0.086	0.100	0.100

E   Full Weighted Least Squares Results

E.1   Results from Oliver and Wood (2014) Questions

Table A6: Covariate Relationships for 911 Conspiracy from Replication of Oliver and olivwood (2014)

	Belief in 911 Conspiracy		
	Positively Keyed Only	Weighted Least Squares	Interaction with D
	(1)	(2)	(3)
Very Conservative	0.250**	0.130	0.081
	(0.120)	(0.093)	(0.087)
D			0.700***
			(0.081)
Very Liberal	0.390***	0.150*	0.110
	(0.120)	(0.088)	(0.081)
Numeracy	–0.130***	–0.090**	–0.072**
	(0.049)	(0.037)	(0.035)
Age	–0.025***	–0.004***	–0.004**
	(0.002)	(0.002)	(0.002)
Education	–0.0002	–0.0002*	–0.0002
	(0.0002)	(0.0001)	(0.0001)
Very Conservative:D			0.170*
			(0.087)
Very Liberal:D			0.280***
			(0.081)
Numeracy:D			–0.062*
			(0.035)
Age:D			–0.022***
			(0.002)
Education:D			–0.00004
			(0.0001)
Constant	0.700***	0.024	0.004
	(0.120)	(0.087)	(0.081)
Observations	916	1,864	1,864
R^2	0.160	0.012	0.160
Adjusted R^2	0.160	0.010	0.150

Note: *p<0.1; **p<0.05; ***p<0.01

Table A7: Covariate Relationships for Soros Conspiracy from Replication of Oliver and olivwood (2014)

	Belief in Soros Conspiracy		
	Positively Keyed Only	Weighted Least Squares	Interaction with D
	(1)	(2)	(3)
Very Conservative	1.100***	0.880***	0.880***
	(0.120)	(0.088)	(0.088)
D			0.270***
			(0.081)
Very Liberal	−0.260**	−0.390***	−0.400***
	(0.110)	(0.082)	(0.082)
Numeracy	−0.160***	−0.120***	−0.120***
	(0.047)	(0.035)	(0.035)
Age	−0.009***	−0.002	−0.002
	(0.002)	(0.002)	(0.002)
Education	−0.0002	0.00001	−0.00001
	(0.0001)	(0.0001)	(0.0001)
Very Conservative:D			0.250***
			(0.088)
Very Liberal:D			0.140*
			(0.082)
Numeracy:D			−0.045
			(0.035)
Age:D			−0.006***
			(0.002)
Education:D			−0.0002*
			(0.0001)
Constant	0.300***	0.029	0.032
	(0.110)	(0.082)	(0.081)
Observations	974	1,845	1,845
R^2	0.110	0.075	0.092
Adjusted R^2	0.110	0.072	0.086

Note: *p<0.1; **p<0.05; ***p<0.01

Table A8: Covariate Relationships for Contrails Conspiracy from Replication of Oliver and oliv-wood (2014)

	Belief in Contrails Conspiracy		
	Positively Keyed Only	Weighted Least Squares	Interaction with D
	(1)	(2)	(3)
Very Conservative	0.430***	0.240***	0.250***
	(0.130)	(0.084)	(0.082)
D			0.430***
			(0.077)
Very Liberal	−0.039	−0.110	−0.110
	(0.120)	(0.078)	(0.077)
Numeracy	−0.230***	−0.220***	−0.220***
	(0.050)	(0.033)	(0.033)
Age	−0.027***	−0.023***	−0.022***
	(0.002)	(0.002)	(0.002)
Education	0.0001	0.0002	0.0001
	(0.0002)	(0.0001)	(0.0001)
Very Conservative:D			0.180**
			(0.082)
Very Liberal:D			0.075
			(0.077)
Numeracy:D			−0.012
			(0.033)
Age:D			−0.006***
			(0.002)
Education:D			−0.00002
			(0.0001)
Constant	0.640***	0.250***	0.210***
	(0.120)	(0.078)	(0.077)
Observations	938	1,844	1,844
R^2	0.160	0.140	0.170
Adjusted R^2	0.150	0.130	0.170

Note: *p<0.1; **p<0.05; ***p<0.01

Table A9: Covariate Relationships for Iraq Conspiracy from Replication of Oliver and olivwood (2014)

	Belief in Iraq Conspiracy		
	Positively Keyed Only	Weighted Least Squares	Interaction with D
	(1)	(2)	(3)
Very Conservative	0.025	–0.220***	–0.200**
	(0.120)	(0.085)	(0.084)
D			0.510***
			(0.077)
Very Liberal	0.390***	0.250***	0.220***
	(0.110)	(0.078)	(0.077)
Numeracy	–0.150***	–0.082**	–0.075**
	(0.047)	(0.034)	(0.034)
Age	–0.020***	–0.012***	–0.012***
	(0.002)	(0.002)	(0.002)
Education	0.0001	–0.00001	–0.00001
	(0.0001)	(0.0001)	(0.0001)
Very Conservative:D			0.220***
			(0.084)
Very Liberal:D			0.170**
			(0.077)
Numeracy:D			–0.080**
			(0.034)
Age:D			–0.008***
			(0.002)
Education:D			0.0001
			(0.0001)
Constant	0.720***	0.240***	0.220***
	(0.110)	(0.079)	(0.077)
Observations	1,011	1,962	1,962
R^2	0.120	0.054	0.093
Adjusted R^2	0.120	0.052	0.088

Note: *p<0.1; **p<0.05; ***p<0.01

Table A10: Covariate Relationships for Obama Conspiracy from Replication of Oliver and oliv-wood (2014)

	Belief in Obama Conspiracy		
	Positively Keyed Only	Weighted Least Squares	Interaction with D
	(1)	(2)	(3)
Very Conservative	1.300***	1.000***	1.100***
	(0.140)	(0.097)	(0.096)
D			0.290***
			(0.089)
Very Liberal	−0.350***	−0.390***	−0.400***
	(0.130)	(0.090)	(0.090)
Numeracy	−0.140**	−0.150***	−0.160***
	(0.056)	(0.038)	(0.038)
Age	−0.010***	−0.005***	−0.005***
	(0.003)	(0.002)	(0.002)
Education	−0.0002	−0.0002	−0.0002
	(0.0002)	(0.0001)	(0.0001)
Very Conservative:D			0.200**
			(0.096)
Very Liberal:D			0.042
			(0.090)
Numeracy:D			0.015
			(0.038)
Age:D			−0.005***
			(0.002)
Education:D			−0.0001
			(0.0001)
Constant	−0.068	−0.360***	−0.360***
	(0.130)	(0.090)	(0.089)
Observations	905	1,847	1,847
R^2	0.110	0.085	0.094
Adjusted R^2	0.100	0.082	0.088

Note: *p<0.1; **p<0.05; ***p<0.01

E.2   Results from Jamieson and Albarracin (2020) Questions

Table A11: Covariate Relationships for VitaminC Conspiracy from Replication of Jamieson and Albarracin (2020)

| | Belief in VitaminC Conspiracy | | |
| | Positively Keyed Only | Weighted Least Squares | Interaction with D |
	(1)	(2)	(3)
Very Conservative	0.560***	0.390***	0.410***
	(0.140)	(0.093)	(0.094)
D			0.140*
			(0.083)
Very Liberal	0.360***	−0.028	−0.049
	(0.120)	(0.086)	(0.085)
Numeracy	−0.095*	−0.160***	−0.160***
	(0.050)	(0.037)	(0.037)
Age	−0.017***	−0.013***	−0.013***
	(0.002)	(0.002)	(0.002)
Education	0.0002	0.0001	0.0001
	(0.0002)	(0.0001)	(0.0001)
Very Conservative:D			0.150
			(0.094)
Very Liberal:D			0.410***
			(0.085)
Numeracy:D			0.064*
			(0.037)
Age:D			−0.004***
			(0.002)
Education:D			0.0001
			(0.0001)
Constant	2.900***	2.700***	2.700***
	(0.120)	(0.084)	(0.083)
Observations	1,017	2,050	2,050
R^2	0.082	0.046	0.064
Adjusted R^2	0.077	0.043	0.059

Note: *p<0.1; **p<0.05; ***p<0.01

Table A12: Covariate Relationships for HandWash Conspiracy from Replication of Jamieson and Albarracin (2020)

	Belief in HandWash Conspiracy		
	Positively Keyed Only	Weighted Least Squares	Interaction with D
	(1)	(2)	(3)
Very Conservative	0.570***	0.350***	0.340***
	(0.140)	(0.083)	(0.079)
D			0.460***
			(0.071)
Very Liberal	0.140	−0.009	−0.016
	(0.130)	(0.076)	(0.073)
Numeracy	−0.220***	−0.110***	−0.110***
	(0.055)	(0.033)	(0.031)
Age	−0.011***	−0.008***	−0.008***
	(0.003)	(0.001)	(0.001)
Education	−0.0004***	−0.0003***	−0.0003***
	(0.0002)	(0.0001)	(0.0001)
Very Conservative:D			0.230***
			(0.079)
Very Liberal:D			0.160**
			(0.073)
Numeracy:D			−0.110***
			(0.031)
Age:D			−0.004***
			(0.001)
Education:D			−0.0002*
			(0.0001)
Constant	2.600***	2.100***	2.100***
	(0.130)	(0.075)	(0.071)
Observations	996	2,053	2,053
R^2	0.060	0.031	0.110
Adjusted R^2	0.055	0.028	0.100

Note: *p<0.1; **p<0.05; ***p<0.01

Table A13: Covariate Relationships for BioWeapon Conspiracy from Replication of Jamieson and Albarracin (2020)

	Belief in BioWeapon Conspiracy		
	Positively Keyed Only	Weighted Least Squares	Interaction with D
	(1)	(2)	(3)
Very Conservative	1.200***	0.930***	0.890***
	(0.130)	(0.100)	(0.100)
D			0.100
			(0.089)
Very Liberal	−0.260**	−0.500***	−0.520***
	(0.130)	(0.091)	(0.091)
Numeracy	−0.270***	−0.270***	−0.270***
	(0.055)	(0.040)	(0.039)
Age	−0.005**	−0.002	−0.001
	(0.002)	(0.002)	(0.002)
Education	−0.0003	−0.0002*	−0.0002*
	(0.0002)	(0.0001)	(0.0001)
Very Conservative:D			0.310***
			(0.100)
Very Liberal:D			0.260***
			(0.091)
Numeracy:D			−0.003
			(0.039)
Age:D			−0.004**
			(0.002)
Education:D			−0.00004
			(0.0001)
Constant	3.100***	3.000***	3.000***
	(0.130)	(0.089)	(0.089)
Observations	1,044	2,052	2,052
R^2	0.100	0.081	0.091
Adjusted R^2	0.100	0.079	0.086

Note: *p<0.1; **p<0.05; ***p<0.01

Table A14: Covariate Relationships for USCreated Conspiracy from Replication of Jamieson and Albarracin (2020)

	Belief in USCreated Conspiracy		
	Positively Keyed Only	Weighted Least Squares	Interaction with D
	(1)	(2)	(3)
Very Conservative	0.410***	0.150*	0.140*
	(0.120)	(0.084)	(0.084)
D			0.100
			(0.076)
Very Liberal	−0.038	−0.190**	−0.180**
	(0.120)	(0.078)	(0.078)
Numeracy	−0.150***	−0.180***	−0.190***
	(0.049)	(0.034)	(0.033)
Age	−0.023***	−0.020***	−0.020***
	(0.002)	(0.001)	(0.001)
Education	−0.0001	−0.0001	−0.0001
	(0.0002)	(0.0001)	(0.0001)
Very Conservative:D			0.270***
			(0.084)
Very Liberal:D			0.140*
			(0.078)
Numeracy:D			0.038
			(0.033)
Age:D			−0.003**
			(0.001)
Education:D			0.00000
			(0.0001)
Constant	3.000***	2.900***	2.900***
	(0.110)	(0.076)	(0.076)
Observations	1,034	2,054	2,054
R^2	0.120	0.100	0.110
Adjusted R^2	0.110	0.099	0.100

Note: *p<0.1; **p<0.05; ***p<0.01

E.3   Results from Allcott and Getnzkow (2017) "Big Fake" Questions

Table A15: Covariate Relationships for Pope.endorse Conspiracy from Replication of Allcott and Gentzkow (2017)

	Belief in Pope.endorse Conspiracy		
	Positively Keyed Only	Weighted Least Squares	Interaction with D
	(1)	(2)	(3)
Very Conservative	0.039	0.048*	0.049*
	(0.039)	(0.028)	(0.028)
D			0.016
			(0.029)
Very Liberal	−0.023	−0.064**	−0.056**
	(0.037)	(0.025)	(0.026)
Numeracy	−0.049***	−0.044***	−0.045***
	(0.017)	(0.012)	(0.012)
Age	−0.005***	−0.005***	−0.005***
	(0.001)	(0.001)	(0.001)
Education	0.0001	0.0001*	0.0001*
	(0.0001)	(0.00003)	(0.00003)
Very Conservative:D			−0.010
			(0.028)
Very Liberal:D			0.033
			(0.026)
Numeracy:D			−0.004
			(0.012)
Age:D			0.0001
			(0.001)
Education:D			−0.00000
			(0.00003)
Constant	0.600***	0.590***	0.590***
	(0.039)	(0.028)	(0.029)
Observations	709	1,282	1,282
R^2	0.071	0.077	0.082
Adjusted R^2	0.064	0.073	0.074

Note: *p<0.1; **p<0.05; ***p<0.01

Table A16: Covariate Relationships for RuPaul Conspiracy from Replication of Allcott and Gentzkow (2017)

| | Belief in RuPaul Conspiracy | | |
| | Positively Keyed Only | Weighted Least Squares | Interaction with D |
	(1)	(2)	(3)
Very Conservative	–0.140***	–0.110***	–0.110***
	(0.040)	(0.030)	(0.030)
D			0.120***
			(0.030)
Very Liberal	0.110***	0.041	0.045
	(0.038)	(0.028)	(0.027)
Numeracy	0.001	–0.006	–0.007
	(0.018)	(0.013)	(0.013)
Age	–0.005***	–0.002***	–0.002***
	(0.001)	(0.001)	(0.001)
Education	–0.00005	–0.00002	–0.00003
	(0.0001)	(0.00004)	(0.00004)
Very Conservative:D			–0.029
			(0.030)
Very Liberal:D			0.069**
			(0.027)
Numeracy:D			0.007
			(0.013)
Age:D			–0.003***
			(0.001)
Education:D			–0.00002
			(0.00004)
Constant	0.730***	0.610***	0.610***
	(0.041)	(0.031)	(0.030)
Observations	703	1,281	1,281
R^2	0.100	0.031	0.062
Adjusted R^2	0.094	0.027	0.054

Note: *p<0.1; **p<0.05; ***p<0.01

Table A17: Covariate Relationships for Clinton.arms Conspiracy from Replication of Allcott and Gentzkow (2017)

	Belief in Clinton.arms Conspiracy		
	Positively Keyed Only	Weighted Least Squares	Interaction with D
	(1)	(2)	(3)
Very Conservative	0.260***	0.210***	0.210***
	(0.040)	(0.031)	(0.031)
D			0.036
			(0.032)
Very Liberal	−0.080**	−0.110***	−0.110***
	(0.037)	(0.028)	(0.029)
Numeracy	−0.055***	−0.042***	−0.044***
	(0.017)	(0.013)	(0.013)
Age	−0.004***	−0.003***	−0.003***
	(0.001)	(0.001)	(0.001)
Education	−0.00003	−0.00001	−0.00000
	(0.00004)	(0.00004)	(0.00004)
Very Conservative:D			0.048
			(0.031)
Very Liberal:D			0.030
			(0.029)
Numeracy:D			−0.012
			(0.013)
Age:D			−0.001
			(0.001)
Education:D			−0.00003
			(0.00004)
Constant	0.640***	0.600***	0.600***
	(0.041)	(0.031)	(0.032)
Observations	719	1,282	1,282
R^2	0.110	0.076	0.085
Adjusted R^2	0.100	0.072	0.077

Note: *p<0.1; **p<0.05; ***p<0.01

Table A18: Covariate Relationships for Ireland.asylum Conspiracy from Replication of Allcott and Gentzkow (2017)

	Belief in Ireland.asylum Conspiracy		
	Positively Keyed Only	Weighted Least Squares	Interaction with D
	(1)	(2)	(3)
Very Conservative	0.029	–0.016	–0.020
	(0.038)	(0.029)	(0.030)
D			0.100***
			(0.030)
Very Liberal	0.120***	0.058**	0.054**
	(0.036)	(0.027)	(0.027)
Numeracy	–0.021	–0.024*	–0.026**
	(0.017)	(0.013)	(0.013)
Age	–0.004***	–0.002***	–0.002***
	(0.001)	(0.001)	(0.001)
Education	0.00004	0.00003	0.00004
	(0.00004)	(0.00004)	(0.00004)
Very Conservative:D			0.049*
			(0.030)
Very Liberal:D			0.070***
			(0.027)
Numeracy:D			0.005
			(0.013)
Age:D			–0.002***
			(0.001)
Education:D			–0.00000
			(0.00004)
Constant	0.660***	0.550***	0.550***
	(0.041)	(0.030)	(0.030)
Observations	714	1,284	1,284
R^2	0.069	0.020	0.047
Adjusted R^2	0.063	0.016	0.039

Note: *p<0.1; **p<0.05; ***p<0.01

Table A19: Covariate Relationships for Pence.vulgar Conspiracy from Replication of Allcott and Gentzkow (2017)

	Belief in Pence.vulgar Conspiracy		
	Positively Keyed Only	Weighted Least Squares	Interaction with D
	(1)	(2)	(3)
Very Conservative	0.018	–0.019	–0.018
	(0.043)	(0.032)	(0.032)
D			0.068**
			(0.032)
Very Liberal	0.110***	0.066**	0.063**
	(0.040)	(0.029)	(0.029)
Numeracy	–0.015	–0.003	–0.004
	(0.018)	(0.014)	(0.014)
Age	–0.005***	–0.004***	–0.004***
	(0.001)	(0.001)	(0.001)
Education	–0.00000	0.00001	0.00002
	(0.00005)	(0.00004)	(0.00004)
Very Conservative:D			0.036
			(0.032)
Very Liberal:D			0.049*
			(0.029)
Numeracy:D			–0.011
			(0.014)
Age:D			–0.001**
			(0.001)
Education:D			–0.00002
			(0.00004)
Constant	0.700***	0.630***	0.640***
	(0.045)	(0.032)	(0.032)
Observations	702	1,283	1,283
R^2	0.072	0.039	0.051
Adjusted R^2	0.065	0.036	0.043

Note: *p<0.1; **p<0.05; ***p<0.01

Table A20: Covariate Relationships for FBI.agent Conspiracy from Replication of Allcott and Gentzkow (2017)

	Belief in FBI.agent Conspiracy		
	Positively Keyed Only	Weighted Least Squares	Interaction with D
	(1)	(2)	(3)
Very Conservative	0.140***	0.078***	0.079***
	(0.041)	(0.029)	(0.030)
D			0.051*
			(0.030)
Very Liberal	−0.035	−0.062**	−0.067**
	(0.036)	(0.027)	(0.028)
Numeracy	−0.040**	−0.037***	−0.037***
	(0.017)	(0.013)	(0.013)
Age	−0.003***	−0.002***	−0.002***
	(0.001)	(0.001)	(0.001)
Education	−0.0001	−0.0001*	−0.0001
	(0.00004)	(0.00004)	(0.00004)
Very Conservative:D			0.063**
			(0.030)
Very Liberal:D			0.031
			(0.028)
Numeracy:D			−0.002
			(0.013)
Age:D			−0.001**
			(0.001)
Education:D			0.00000
			(0.00004)
Constant	0.580***	0.540***	0.530***
	(0.041)	(0.030)	(0.030)
Observations	714	1,289	1,289
R^2	0.046	0.024	0.034
Adjusted R^2	0.039	0.021	0.025

Note: *p<0.1; **p<0.05; ***p<0.01

F   Acquiescence bias in learning about political facts

We focus in the main body of our essay on rates and correlates of beliefs in political conspiracies and politically-relevant facts. Recent scholarship, however, has been interested not only in population beliefs, but in how delivery of information relevant to those beliefs changes – or does not change – such beliefs (e.g., Hill, 2017; Hill and Huber, 2019; Nyhan and Reifler, 2010). For example, Y_{i1} could be subject beliefs that a statement of fact about economic growth is true when first queried about a topic. But after delivered government reports, news headlines or articles (Nyhan and Reifler, 2010), or political speeches (Carlson, 2019), they may update beliefs to $Y_{i2} \neq Y_{i1}$. One may be tempted to subtract the sample estimate of Y_{i2} from Y_{i1} to estimate the informational value of the stimulus delivered (perhaps in an experimental setting to gain an estimate of the causal effect). This is particularly attractive because it would difference off any fixed acquiescence bias to the instrument, δ_i. However, if there is also *acquiescence bias in learning*, the difference $(Y_{i2}-Y_{i1})$ is a biased estimate of true learning. That is, it may be that subjects overlearn towards true and underlearn towards false.

A model similar to that above provides structure to this problem and leads to an unbiased estimator for learning in the presence of acquiescence bias. Begin by extending the data-generating Equation 2 to two separate elicitations at time 1 and time 2. Subject i's response at the two periods are $Y_{i1}(D_i)$ and $Y_{i2}(D_i)$.

Suppose participants are delivered some stimulus between time 1 and time 2. We define learning from that stimulus by

$$Y_{i2}(D_i) = Y_{i1}(D_i) + \lambda_i.$$

Here λ_i is the change in beliefs after delivery of the stimulus. For example, if the stimulus increases agreement with the conspiracy or fact $\lambda_i > 0$.

Absent acquiescence bias in learning, one could subtract Y_{i1} from Y_{i2} as an estimate of λ_i. If, however, we are concerned that acquiescence bias influences how subjects respond to a stimulus, the difference would be biased by this influence. Define the data-generating process for Y_{i2} in the presence of bias in learning

$$Y_{i2}(D_i) = \underbrace{Y_i^* + D_i\delta_i}_{\text{Original belief + bias}} + \underbrace{\lambda_i + D_i\kappa_i}_{\text{Learning + bias}} \tag{A1}$$

where κ_i is acquiescence bias in learning with similar operation to acquiescence bias δ_i.

To estimate average learning from the stimulus, $E(\lambda)$, consider a least-squares regression specification

$$(Y_2 - Y_1) = \alpha + \beta D + \varepsilon \tag{A2}$$

where Y_t and D are $\mathcal{N}$-vectors for the sample of Y_i and D_i and ε a $\mathcal{N}$-vector of idiosyncratic mean-zero errors.

The parameter estimates from Eq. A2 map back to the parameters of Eq. A1. The intercept α estimates average learning from the stimulus $E(\lambda)$, and the coefficient β the average bias in learning $E(\kappa)$.

F.1   Application to US and China learning samples, learning

In this section, we estimate how acquiescence bias in learning varies across political statements of fact and samples. We compare how respondents asked the positive-keyed version learn after

Table A21: Estimates of learning, true versus false statements

Statement	Naive False	Naive True	Learning	Learning Bias
US: Jobs 2016	23.0	10.7	16.9	-6.2
US: Jobs 2016	15.9	5.1	10.5	-5.4
US: Food stamps 2018	3.8	-4.4	-0.3	-4.1
US: Currency devaluation	9.6	7.5	8.5	-1.1
China: GDP 2015	-4.3	-5.6	-5.0	-0.7
US: Currency devaluation	15.8	15.0	15.4	-0.4
China: Hong Kong Work Week	-6.1	-4.6	-5.4	0.7
China: Overseas Students	-0.8	1.4	0.3	1.1
China: 2016 GDP Growth	-2.0	0.4	-0.8	1.2
China: Scholarships	-1.6	0.9	-0.4	1.3
US: GDP 2014	2.8	5.8	4.3	1.5
China: Military Parade	-0.9	4.4	1.8	2.7
China: Exports	-3.8	1.8	-1.0	2.8
China: ZTE Fine	-2.6	3.9	0.7	3.3
China: Currency devaluation	-1.4	6.5	2.5	4.0
China: Military Spending	-6.1	2.4	-1.8	4.3
China: Qualcomm Fine	-3.5	7.5	2.0	5.5
China: IMF Currency Basket	-4.5	8.3	1.9	6.4
US: GDP 2014	-8.5	13.4	2.4	10.9

being presented with news headlines about the statement compared to respondents presented the negative-keyed version. We provide details of the headlines delivered to subjects in Section G below. In Table A21, we present estimates applying Eq. A2 of acquiescence bias in learning and learning for each statement of fact our subjects evaluated.

Like our estimates on overall beliefs in political conspiracies, we see differences in estimates of learning when using the negatively keyed version of the question as compared to the positively keyed version of the question. As presented in the Learning Bias column of Table A21, on average, we overestimate learning for positively keyed questions, and underestimate learning for negatively keyed questions. However, this finding is less consistent across questions than acquiescence bias in overall beliefs, and in some cases we do see negative learning bias, where the negatively keyed version of the question has higher estimated learning.

G   Selection of facts, headlines, and news sources

To select news headlines for our MTurk, NORC, and China studies, we selectied six major U.S. news outlets: Fox News, MSNBC, CNN, the New York Times, National Public Radio (NPR), and USA Today. We selected these sources because they all have written articles that readers can access online and extensive previous research has explored their potential bias (e.g., Budak, Goel, and Rao, 2016; Groseclose and Milyo, 2005; Gentzkow and Shapiro, 2010). USA Today

and CNN are generally considered moderate news outlets without political slant, while Fox News is more conservative and the New York Times is more liberal (Budak, Goel, and Rao, 2016). After selecting the news outlets, we deliberately excluded opinion or editorial sections as we started our search for news headlines.

Next, we created a list of objective economic facts with political implications about which to find news headlines. We searched for facts that would be (1) relevant to voter decision-making, especially if Americans make decisions based on the state of the economy, and (2) could be verified in an objective data report. We thus focused primarily on economic indicators such as GDP growth, manufacturing data, retail sales, the unemployment rate, and Consumer Price Index (CPI). We then sought topics that were specifically focused on U.S.-China relations to allow for meaningful comparisons across the countries. We focused on trade deficits between the countries, currency exchange rates, and the U.S.-China cap-and-trade deal.

After brainstorming this list of topics, we located the data release dates. For example, the Bureau of Labor Statistics releases a monthly report called "The Employment Situation" that contains information on the unemployment rate, the number of jobs added to the economy, and other economic indicators. Similarly, the Bureau of Economic Analysis releases quarterly reports, as well as revised estimates, of the Gross Domestic Product. We then searched for articles about the relevant topics on the date of the data report releases across each of our six news outlets. We started by using the advanced search tools on Google News, using keywords, restricting the date to the date of the data release, and the source to the news outlet of interest. The search terms used for each topic are listed in Table A22 below. After collecting headlines from Google News, we went to each outlet's website and used the same keyword searches and date restrictions, where applicable, to locate more headlines. Some websites, such as CNN, did not allow us to restrict search results by date, so we used the search terms and sorted the search results by date and scrolled through the list until we reached the data release date. We then cross-referenced these headlines with Lexis Nexis searches. Lexis Nexis worked well for media outlets, such as the New York Times, that have a physical print publication, but it was not comprehensive for news sources that do not have a print publication. Finally, we inspected the content of each article to make sure that it was about the intended topic.

Most statistic releases did not have corresponding headlines from every news source. After compiling a list of headlines for each release, we tabulated the news outlets most commonly having headlines for the releases. For the U.S., we found that the New York Times, USA Today, CNN, and Fox News most often covered the statistics we searched. We decided to use these four sources, and then selected the subset of releases that had headlines from all four of these news outlets. We then selected the four facts above from this list that covered different economic facts, including one U.S.-China fact, and that had headlines that we deemed modestly related to the fact of interest – often, the news article might mention the release but the headline would be on another topic, e.g. how the equity markets had fared. In sum, our goal was to identify four facts that covered important economic indicators that each had enough informative headlines from four consistent news sources to measure how subjects learned and if they attributed any bias to the news sources. We present the facts in Table A4.

Table A22: Topics considered and search terms used to identify facts and headlines

Topic	Search Terms
GDP Annualized Quarter-over-Quarter	GDP, gross domestic product, economic growth, growth rate, growth, annualized
China devalues the yuan	Yuan, China, currency, devalue
US-China cap and trade	US, China, cap and trade, emissions, carbon, CO_2
Nonfarm Payrolls	nonfarm, payrolls, non farm
ISM Manufacturing	ISM Manufacturing
Retail Sales	Retail sales
GDP Quarterly Estimates	GDP, gross domestic product, economic growth, growth rate, growth
Affordable Care Act Enrollments	HHS, enrollment, sign up, signup, Obamacare, ACA, Affordable Care Act, enroll, healthcare, health care, insurance, health insurance, marketplace, healthcare.gov
The Employment Situation	unemployment, jobs, unemployment rate, job growth, employment, employed, unemployment benefits, labor force
Consumer Price Index	consumer price index, CPI, consumer prices, consumer price, inflation
Trade Deficit	trade, goods and services, goods, services, trade deficit, treasury, foreign trade balance, China, international trade