

Supplementary Materials

Study Purpose and Objective: directly compare average treatment effects obtained in experiments implemented with different samples. That is, compare the treatment effect obtained in an experiment implemented with a population-based sample with the treatment effect obtained from the same experiment (using the same stimuli, questions, procedures) implemented with convenience samples.

STUDY 1 SUPPLEMENTARY MATERIALS

Study 1 DREAM Act Treatment Effects for All Conditions (Relative to Control)

	Polarized Pro	Agreement Pro	Polarized Con	Agreement Con
Exit Poll	-0.06 (0.04)	-0.01 (0.04)	-0.00 (0.04)	-0.02 (0.04)
Student	0.04 (0.06)	0.04 (0.07)	0.12 (0.07)	0.17 (0.07)
Staff	-0.08 (0.09)	0.04 (0.09)	-0.06 (0.10)	0.14 (0.09)
MTurk	-0.02 (0.04)	0.04 (0.04)	0.06 (0.03)	0.08 (0.04)
TESS	-0.01 (0.05)	-0.03 (0.05)	0.14 (0.05)	0.18 (0.04)

STUDY 2 SUPPLEMENTARY MATERIALS

Materials: Study Descriptions and Full Results

Full experimental materials (complete stimuli, all questions) and information regarding response rates and recruitment for each study can be found at:

<http://www.tessexperiments.org/previousstudies.html>

[Experiment 1]

Brandt, Mark. 2009. "Onset and Offset Controllability in Perceptions and Reactions to Home Mortgage Foreclosures."

Abstract:

The circumstances and rhetoric surrounding home foreclosures provide an ideal and timely backdrop for an extension of research on attributional judgments. While people face foreclosure for many reasons, the current debate surrounding the mortgage crisis has highlighted reasons that are either onset or offset controllable; that is, the initial cause, or the subsequent solution may be seen as controllable. In the current study, I examine how people use attributional evidence from multiple time points to determine affective reactions and helping intentions for people undergoing foreclosure, as well as ideological differences in these attributional processes. Participants read about people who were undergoing foreclosure for onset and offset controllable or uncontrollable reasons and then answer questions about their perceptions of these targets.

Participants were randomly assigned to *one of four* conditions where they read about individuals who were in control (or not) of the onset or offset of the mortgage foreclosure situation.:

- Onset Controllable: Some people have a large monthly mortgage payment because they wanted to purchase a larger house than they needed.
- Onset Uncontrollable: Some people have a large monthly mortgage payment because they were misled by a mortgage loan officer that their payments would remain low, when in fact, their payments ended up being very high.
- Offset Controllable: Now they are facing foreclosure because they do not want to continue paying the mortgage, even though they are able to afford the payments.
- Offset Uncontrollable: Now they are facing foreclosure because the primary income earner in the household lost their job due to their company closing and they can no longer afford payments.

4 conditions

First main DV: How much control did these people have over getting into their current financial situation? 1 (no control), 4 (neutral), 7 (complete control)

Raw Unweighted Means by Sample and Condition

Sample	Group 1	Group 2	Group 3	Group 4
TESS	5.90 (0.08) SD=1.41 N=312	5.27 (0.09) SD=1.54 N=313	5.00 (0.09) SD=1.62 N=299	4.47 (0.10) 1.74 N=294
Mturk	5.93 (0.05) SD=1.27 N=786	4.84 (0.05) SD=1.43 N=786	4.74 (0.06) SD=1.52 N=763	3.88 (0.06) SD=1.59 N=716

Group 2 v 1:

TESS: $p < 0.000$

Mturk: $p < 0.000$

Group 3 v 1:

TESS: $p < 0.000$

Mturk: $p < 0.000$

[Experiment 2]

Caprariello, Peter. 2011. "To Do, to Have, or to Share? Valuing Experiences and Material Possessions by Involving Others."

All subjects told:

We are interested in ways you spend your *discretionary money*. Discretionary money refers to money that is spent on anything that is NOT essential to basic activity (that is, essentials refer to things like tuition and textbooks, groceries, transportation, rent, gas for a car, health care, etc.). We'd like you to answer the questions that follow for money that you spent on something discretionary.

Please think of the last time you spent at least \$10 (but no more than \$10,000) of your discretionary money in order

[Randomized to 1 of 4 conditions]

- **to do something with at least one other person.** The primary focus of this expense should have been on an activity – doing something with at least one other person – and not on buying something that could be kept. Maybe you bought tickets to see a movie with some people, maybe you paid to visit an art museum with friends, maybe you and some other people went to a spa together ... any of these would be legitimate examples of spending money to do something with others.
- **to do something by yourself.** The primary focus of this expense should have been on an activity – doing something by yourself – and not on buying something that could be kept. Maybe you bought a ticket to see a movie by yourself, maybe you paid to enter an art museum, maybe you went to a spa by yourself ... any of these would be legitimate examples of spending money to do something by yourself.
- **to acquire a material possession to use with at least one other person.** Maybe you bought a sound system to use with others, maybe you acquired new clothes or jewelry for dressing up to go out with others, or maybe you bought a game to play with others ... any of these would be legitimate examples of spending money to acquire material goods to use with others.

[List experiment]

Below you will read three things that sometimes people oppose or are against. After you read all three, just tell us HOW MANY of them you OPPOSE. We don't want to know which ones, just HOW MANY.

- (1) the federal government increasing assistance to the poor
- (2) professional athletes making millions of dollars per year
- (3) large corporations polluting the environment

[Control gets 3 items, treatment groups get different 4th item]

- (4) granting citizenship to a legal immigrant who is Muslim
- (4) granting citizenship to a legal immigrant who is Christian
- (4) cutting off all immigration to the United States

Raw Unweighted Means by Sample and Condition

Sample	Group 1 (control)	Group 2	Group 3	Group 4
TESS	1.90 (0.03) SD=0.74 N=799	2.12 (0.03) SD=0.98 N=807	2.14 (0.04) SD=0.98 N=777	2.23 (0.03) SD=0.94 N=816
Mturk	1.79 (0.02) SD=0.66 N=725	1.88 (0.03) SD=0.76 N=748	2.01 (0.03) SD=0.83 N=783	2.33 (0.03) SD=0.80 N=786

Group 2 v 1

TESS: $p < 0.000$

Mturk: $p < 0.019$

Group 3 v 1

TESS: $p < 0.000$

Mturk: $p < 0.000$

[Experiment 4]

Flavin, Patrick J. 2011. "Public Attitudes about Political Equality."

Abstract:

Most studies of political equality rest on the normative assumption that citizens, in general, widely support it as a fundamental principle for American democracy. But, do Americans actually support political equality? When phrased vaguely as "political equality," we might expect high levels of public support. But does support remain high when survey respondents are presented with a precise definition of political equality? Moreover, does support for political equality vary depending on the definition presented? Despite the large literature on public attitudes about economic

inequality and equality of opportunity in the United States (e.g., Hochschild 1981; McClosky and Zaller 1984; Feldman 1988; Feldman and Zaller 1992; Bartels 2008), research focusing on public opinion about political equality in particular is nearly nonexistent. This lack of understanding about citizens' beliefs regarding political equality is surprising given the normative importance placed on the concept in so many political science studies. To further our understanding, this TESS project examines citizens' support for various definitions of "political equality" that are experimentally manipulated.

5 conditions:

- Some people think that the United States should place a greater emphasis on promoting political equality. How about you, do you strongly support, somewhat support, neither support nor oppose, somewhat oppose, or strongly oppose promoting political equality?
- Some people think that the United States should place a greater emphasis on promoting political equality. By political equality we mean making sure every citizen has the right to vote and participate in politics to make their opinions known to government. How about you, do you strongly support, somewhat support, neither support nor oppose, somewhat oppose, or strongly oppose promoting political equality?
- Some people think that the United States should place a greater emphasis on promoting political equality. By political equality we mean making sure every citizen has access to an education and relevant information that will allow them to stay informed about political affairs. How about you, do you strongly support, somewhat support, neither support nor oppose, somewhat oppose, or strongly oppose promoting political equality?
- Some people think that the United States should place a greater emphasis on promoting political equality. By political equality we mean making sure elected officials listen and respond to the opinions of all citizens equally – whether they are rich or poor, black or white, male or female – when making important policy decisions. How about you, do you strongly support, somewhat support, neither support nor oppose, somewhat oppose, or strongly oppose promoting political equality?
- Some people think that the United States should place a greater emphasis on promoting political equality. By political equality we mean making sure citizens have equal political influence by limiting the amount of money an individual or group can give to a candidate during a political campaign. How about you, do you strongly support, somewhat support, neither support nor oppose, somewhat oppose, or strongly oppose promoting political equality?

Strongly support (1) to Strongly oppose (5)

Raw Unweighted Means by Sample and Condition

Sample	Group 1	Group 2	Group 3	Group 4	Group 5
TESS	2.61 (0.05) SD=0.93 N=385	1.89 (0.05) SD=0.98 N=385	2.23 (0.05) SD=1.04 N=405	1.78 (0.05) SD=0.96 N=405	2.06 (0.05) SD=1.07 N=426
Mturk	2.02	1.62	1.73	1.54	1.85

	(0.04) SD=0.96 N=490	(0.04) SD=0.90 N=513	(0.04) SD=0.89 N=538	(0.037) SD=0.84 N=523	(0.04) SD=0.99 N=531
--	----------------------------	----------------------------	----------------------------	-----------------------------	----------------------------

Group 2 v 1:

TESS: $p < 0.000$

Mturk: $p < 0.000$

Group 3 v 1:

TESS: $p < 0.000$

Mturk: $p < 0.000$

[Experiment 5]

Gash, Alison and Michael Murakami. 2009. "Understanding How Policy Venue Influences Public Opinion."

Design:

Participants randomly assigned to one of four groups. In each group, participants are asked to state their opinion on a policy that prevents companies from considering gender in their hiring practices. Individuals in group one are asked to imagine that the policy was issued by a court in their state; those in group two are asked to imagine that their state legislature produced the policy; and those in group three are asked to imagine that the policy was produced through an initiative process in their state. Participants in group four are simply asked whether they agree that companies should be barred from considering gender in hiring. No information about venue is given to group four participants.

Do you agree or disagree with the court's [or legislature's, voters', or control] decision? 1 (Strongly agree), 2 (Somewhat agree), 3 (Somewhat disagree), 4 (Strongly disagree)

Raw Unweighted Means by Sample and Condition

Sample	Group 2 (Court)	Group 3 (Legislature)	Group 4 (Voters)	Group 1 (No prompt)
TESS	1.813 (0.051) SD=0.809 N=252	1.90 (0.06) SD=0.90 N=249	1.83 (0.05) SD=0.88 N=263	2.348 (0.053) SD=0.83 N=244
Mturk	1.734 (0.029) SD=0.81 N=778	1.81 (0.03) SD=0.84 N=780	1.82 (0.03) SD=0.86 N=743	2.293 (0.033) SD=0.903 N=741

Group 2 v 1:

TESS: $p < 0.000$

Mturk: $p < 0.000$

Group 3 v 1:

TESS: $p < 0.000$
 Mturk: $p < 0.000$

[Experiment 6]

Mello, Michelle. 2010. "Patient Responses to Medical Error Disclosure: Does Compensation Matter?"

Abstract:

Disclosing medical errors to patients and families is both a regulatory requirement and an ethical imperative. However, physicians' fear of malpractice liability is a major barrier to disclosure. Disclosure may increase medicolegal risk by alerting patients that they have been injured by medical management. To manage this risk, several healthcare institutions have implemented programs through which they make rapid offers of compensation (sometimes at modest levels) following disclosure of a medical injury, but little is known about these programs' effectiveness in deterring malpractice claims. Using a series of medical error vignettes, this study investigated the effects of early compensation offers on patients' propensity to sue following disclosure of a harmful error, compared to apology alone, including the importance of the amount of compensation offered.

- 16 conditions

How likely would you be to seek legal advice about suing Dr. S/T? Very unlikely (1), somewhat unlikely (2), Somewhat likely (3), Very likely (4)

Raw Unweighted Means by Sample and Condition

Sample	Group 1	Group 2	Group 3	Group 4
TESS	3.17 (0.08) SD=0.97 N=132	3.21 (0.09) SD=0.99 N=126	3.10 (0.09) SD=1.01 N=136	3.27 (0.08) SD=0.91 N=131
Mturk	3.19 (0.06) SD=0.82 N=187	3.24 (0.07) SD=0.97 N=197	3.31 (0.06) SD=0.83 N=180	3.08 (0.07) SD=1.00 N=199

Sample	Group 5 (for tests this is called Group 2)	Group 6	Group 7	Group 8
TESS	2.74 (0.09) SD=1.03 N=139	2.70 (0.09) SD=1.04 N=125	2.87 (0.09) SD=1.07 N=136	2.74 (0.10) SD=1.06 N=106
Mturk	2.79 (0.07) SD=0.94 N=182	2.74 (0.08) SD=1.00 N=176	2.68 (0.07) SD=1.00 N=173	2.63 (0.08) SD=1.08 N=203

Sample	Group 9 (for tests this is called Group 3)	Group 10	Group 11	Group 12
TESS	2.79 (0.09) SD=0.96 N=124	2.82 (0.09) SD=1.09 N=132	3.00 (0.09) SD=1.05 N=125	2.85 (0.09) SD=1.01 N=131
Mturk	2.99 (0.07) SD=0.91 N=164	2.79 (0.07) SD=1.00 N=192	2.91 (0.07) SD=1.00 N=172	3.00 (0.07) SD=0.98 N=217

Sample	Group 13	Group 14	Group 15	Group 16
TESS	2.27 (0.08) SD=0.93 N=130	2.33 (0.09) SD=1.05 N=132	2.51 (0.10) SD=1.12 N=120	2.48 (0.09) SD=1.08 N=133
Mturk	2.43 (0.07) SD=0.94 N=198	2.45 (0.08) SD=1.03 N=184	2.51 (0.07) SD=1.01 N=195	2.54 (0.07) SD=1.05 N=213

NOTE: For experimental tests of main dimensions, we compare Group 1 to Group 5, and Group 1 to Group 9. (Despite their labeling here, Groups 5 and 9 will be labeled as Groups 2 and 3 in Figures).

Group 2 v 1:

TESS: $p < 0.001$

Mturk: $p < 0.000$

Group 3 v 1:

TESS: $p < 0.002$

Mturk: $p < 0.032$

[Experiment 7]

Jacobsen, Rebecca. 2011. "Informing the Public or Information Overload? The influence of school accountability data format on public satisfaction."

4 conditions:

- Below are report card data for Oak High School. The performance of the students at Oak High School has been measured and the school's performance index scores are listed for each area. The performance index is a weighted average of the tests given to students at all grade levels. This results in a scale from 0 to 200 points. Considering the provided data, please answer the accompanying questions.

Oak High School

Educational Goal	Performance Index (0-200)
Academics	164
Arts	190
Citizenship and Community Responsibility	190

- Below are report card data for Oak High School. The performance of the students at Oak High School has been measured and the resulting letter grades have been earned for each area. Letter grades include A, B, C, D and F. Considering the provided data, please answer the accompanying questions.

Oak High School	
Educational Goal	Letter Grade
Academics	B
Arts	A
Citizenship and Community Responsibility	A

- Below are report card data for Oak High School. The performance of the students at Oak High School has been measured and the percent of students meeting or exceeding the standards are listed for each area. The percent can range from 0 percent of students to 100 percent of students meeting the standards. Considering the provided data, please answer the accompanying questions.

Oak High School	
Educational Goal	Percent Meeting or Exceeding Standards
Academics	82%
Arts	95%
Citizenship and Community Responsibility	95%

- Below are report card data for Oak High School. The performance of the students at Oak High School has been measured and when compared to established performance standards, the following achievement levels have been earned in each area. The categories of performance standards include Advanced, Proficient, Basic, Below Basic and Failing. Considering the provided data, please answer the accompanying questions.

Oak High School	
Educational Goal	Achievement Level
Academics	Proficient
Arts	Advanced
Citizenship and Community Responsibility	Advanced

Satisfaction means many things. Overall, how SATISFIED are you with Oak School based on these data? 1 (Very dissatisfied), 7 (very satisfied)

Note: this is only one school. You could look at Elm and Cedar Schools.

Raw Unweighted Means by Sample and Condition

Sample	Group 1	Group 2	Group 3	Group 4
TESS	4.49 (0.09) SD=1.50 N=277	5.26 (0.08) SD=1.34 N=265	5.05 (0.09) SD=1.44 N=284	5.00 (0.09) SD=1.48 N=267
Mturk (not bundled)	4.86 (0.08) SD=1.24 N=229	5.50 (0.07) SD=1.15 N=256	5.17 (0.06) SD=1.03 N=266	5.42 (0.07) SD=1.12 N=257
Mturk Bundled	4.79 (0.08) SD=1.32 N=253	5.52 (0.07) SD=1.18 N=254	5.09 (0.07) SD=1.19 N=256	5.50 (0.07) SD=1.10 N=279

Note: the bundled group had 2 different ordering. Here, they are put into a singled bundled category.

Group 2 v 1:

TESS: $p < 0.000$

Mturk: $p < 0.000$

Mturk bundled: $p < 0.000$

Group 3 v 1:

TESS: $p < 0.000$

Mturk: $p < 0.003$

Mturk bundled: $p < 0.009$

[Experiment 8]

Piazza, James A. 2011. "Terrorism Suspect Identity and Public Support for Controversial Detention and Interrogation Practices."

Abstract:

This study proposes to examine the effects that the religious identity of a terrorism suspect has on public support for the application of controversial interrogation techniques and detention practices that have become part of United States counterterrorism policy since the September 11th 2001 terrorist attacks. It tests the hypothesis that the public is more permissive of physically abusive interrogation of Muslim terrorist suspects and is more accepting of indefinite detention and transfer of accused terrorists to military commissions for suspects identified as Muslims or as associated with extremist Islamic movements. This study executes an original national survey and finds that respondents are generally more supportive of subjecting terror suspects with stereotypical Muslim names or that are associated with a radical Islamic terrorist group to harsher treatment than non-Muslim suspects associated with domestic, right-wing terrorist movements.

The proposed study executes a survey experiment involving four treatment vignettes and one control vignette and 17 survey questions administered to 1,050 respondents. Respondents are randomly assigned to one of the five treatments which depict a short AP newswire blurb describing

an arrest of two terrorist suspects in suburban Chicago. The treatments are identical to one another except they vary the names of the suspects (stereotypical Arabic/Muslim vs. Anglo-American) and the names of the terrorist movement the suspects are alleged to be members of (radical Islamists vs. right-wing American extremist). The control vignette omits any identification of the suspect names or groups.

- 5 conditions

Please tell us if you would support or oppose each of these items as a method of getting information from a suspect in the case described earlier: Applying electric shocks to the suspect.

1 (strongly support), 2 (support), 3 (neither support nor oppose), 4 (oppose), 5 (strongly oppose)

Raw Unweighted Means by Sample and Condition

Sample	Group 1	Group 2	Group 3	Group 4	Group 5
TESS	3.44 (0.08) SD=1.21 N=208	3.45 (0.08) SD=1.30 N=235	3.57 (0.08) SD=1.25 N=226	3.51 (0.09) SD=1.27 N=221	3.56 (0.08) SD=1.25 N=220
Mturk (not bundled)	3.74 (0.09) SD=1.26 N=201	3.68 (0.08) SD=1.23 N=211	3.56 (0.10) SD=1.36 N=203	3.79 (0.08) SD=1.21 N=212	3.72 (0.08) SD=1.23 N=215
Mturk bundled*	3.84 (0.08) SD=1.20 N=216	3.81 (0.08) SD=1.22 N=231	3.78 (0.09) SD=1.22 N=199	3.89 (0.09) SD=1.27 N=202	3.91 (0.09) SD=1.22 N=194

*This study was bundled in 2 different orderings. Here they are combined into a single bundled group.

Group 2 v 1:

TESS: $p < 0.938$

Mturk: $p < 0.660$

Mturk bundled: $p < 0.775$

Group 3 v 1:

TESS: $p < 0.277$

Mturk: $p < 0.169$

Mturk bundled: $p < 0.649$

[Experiment 9]

Shafer, Emily Fitzgibbons. 2010. "Why Hillary Rodham Became Hillary Clinton: Consequences of Non-Traditional Last Name Choice in Marriage."

Abstract:

This survey experiment tests whether a woman's choice of last name in marriage has an impact on how she is viewed by others. I propose a vignette-based survey experiment in which I vary a woman's last name in two situations – one in which she is arriving late to work because of her marital commitment and one in which she is arriving late at home because of her work commitment. This will allow me to test whether a woman's last name choice has an impact on how she is viewed as an employee and as a wife even when additional information about how the woman is performing in these roles is given.

I propose to address these unanswered questions through a survey experimental design in which respondents will read a short vignette (one of six conditions) and answer a series of questions regarding the woman as an employee and wife. I manipulate a woman's last name by giving her a last name identical to her husband's, different than her husband's

6 conditions:

- Carol Sherman is married to Bill Cook. Carol has been late to her office job on several occasions because she is currently caring for Bill's mother who has a serious, but non-life threatening illness. Her fellow employees are starting to feel burdened as they are picking up her slack.
- Carol Sherman-Cook is married to Bill Cook. Carol has been late to her office job on several occasions because she is currently caring for Bill's mother who has a serious, but non-life threatening illness. Her fellow employees are starting to feel burdened as they are picking up her slack.
- Carol Cook is married to Bill Cook. Carol has been late to her office job on several occasions because she is currently caring for Bill's mother who has a serious, but non-life threatening illness. Her fellow employees are starting to feel burdened as they are picking up her slack.
- Carol Sherman is married to Bill Cook. Carol has been spending a lot of extra hours at her office job hoping for a promotion. Bill is starting to feel burdened by her absence, as he is picking up her slack in housework.
- Carol Sherman-Cook is married to Bill Cook. Carol has been spending a lot of extra hours at her office job hoping for a promotion. Bill is starting to feel burdened by her absence, as he is picking up her slack in housework.
- Carol Cook is married to Bill Cook. Carol has been spending a lot of extra hours at her office job hoping for a promotion. Bill is starting to feel burdened by her absence, as he is picking up her slack in housework.

Please rate how *committed* you think Carol is as an *employee*. 1=Extremely committed, 5=not at all committed

Raw Unweighted Means by Sample and Condition

Sample	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
TESS	2.97 (0.04) SD=0.89 N=429	2.91 (0.04) SD=0.87 N=441	2.89 (0.04) SD=0.85 N=426	1.85 (0.04) SD=0.76 N=424	1.87 (0.04) SD=0.76 N=446	1.82 (0.04) SD=0.73 N=419
Mturk	3.02 (0.04) SD=0.76	2.97 (0.04) SD=0.73	3.02 (0.04) SD=0.81	1.60 (0.03) SD=0.57	1.53 (0.03) SD=0.56	1.59 (0.03) SD=0.55

	N=435	N=405	N=439	N=434	N=478	N=405
--	-------	-------	-------	-------	-------	-------

Group 2 v 1:

TESS: $p < 0.349$

Mturk: $p < 0.375$

Group 3 v 1:

TESS: $p < 0.204$

Mturk: $p < 0.968$

[Experiment 10]

Thompson, Suzanne C. 2011. "Terrorist Threat: Overreactions, Underreactions, and Realistic Reactions"

Design:

Participants are then randomly assigned to one of three conditions:

- [High Fear] Terrorists have many ways to attack the U.S. mainland. Snipers can target crowded areas to incite fear and suicide bombers can disrupt transportation and communication. The destruction of power grids and water sources can have widespread effects. If a large scale attack occurs, it may be a while before basic services are restored. You should take precautions now. If you are flying, be prepared for extra security measures and screening at the airport. At home, prepare your family for a possible attack by making a home kit that contains items you would need if basic services are interrupted (e.g., water, food, batteries, flashlight, battery-based radio, extra eye-glasses and medications). In addition, you and your family should have a plan for emergencies. For example, make sure every member of your family has a cell-phone or phone card and knows the number to call in an emergency. Take the time to sit down with family members to plan where you will go from work or school if an emergency occurs. Whether on a plane flight, in your neighborhood, or at work, all Americans should continue to be vigilant, take notice of their surroundings, and report suspicious items or activities to local authorities immediately. If you see something, say something.
- [Plain low fear] The Department of Homeland Security is protecting you against terrorist activities in a number of ways and there are additional precautions that you can take to protect yourself and your family. If you are flying, be prepared for extra security measures and screening at the airport. At home, prepare your family for a possible attack by making a home kit that contains items you would need if basic services are interrupted (e.g., water, food, batteries, flashlight, battery-based radio, extra eye-glasses and medications). In addition, you and your family should have a plan for emergencies. For example, make sure every member of your family has a cell-phone or phone card and knows the number to call in an emergency. Take the time to sit down with family members to plan where you will go from work or school if an emergency occurs. Whether on a plane flight, in your neighborhood, or at work, all Americans should continue to be vigilant, take notice of their surroundings, and report suspicious items or activities to local authorities immediately. If you see something, say something.

- [Low fear positive image] The Department of Homeland Security is protecting you against terrorist activities in a number of ways and there are additional precautions that you can take to protect yourself and your family. If you are flying, be prepared for extra security measures and screening at the airport. More experienced travelers appreciate the security of knowing that this extra level of security is being used. At home, you can protect your family in several ways. John Mercer, for example, has followed the Homeland Security suggestions for how to be prepared by making a home kit that contains items his family would need if basic services are interrupted (e.g., water, food, batteries, flashlight, battery-based radio, extra eye-glasses and medications). In addition, he made sure that his family had a plan for emergencies, for example, each member has a cell-phone or phone card and knows the number to call in an emergency. The family took the time to decide how each family member would get home from work or school if an emergency occurs. According to John, it was the right thing to do for his family and easy to accomplish. Whether on a plane flight, in your neighborhood, or at work, all Americans should continue to be vigilant, take notice of their surroundings, and report suspicious items or activities to local authorities immediately. If you see something, say something.

To what extent do these statement describe your thoughts about terrorist attacks? I am concerned about terrorist attacks. 1 (not at all), 7 (very much)

Raw Unweighted Means by Sample and Condition

Sample	Group 1	Group 2	Group 3
TESS	4.15 (0.12) SD=1.70 N=202	4.05 (0.12) SD=1.71 N=198	4.27 (0.12) SD=1.69 N=187
Mturk (not bundled)	3.52 (0.10) SD=1.73 N=322	3.47 (0.09) SD=1.73 N=348	3.42 (0.09) SD=1.66 N=360
Mturk bundled*	3.58 (0.09) SD=1.70 N=337	3.46 (0.09) SD=1.72 N=338	3.55 (0.10) SD=1.82 N=338

*This study was bundled in 2 different orderings. They combined into one bundled group here.

Group 2 v 1:

TESS: $p < 0.566$

Mturk: $p < 0.673$

Mturk bundled: $p < 0.350$

Group 3 v 1:

TESS: $p < 0.49$

Mturk: $p < 0.405$

Mturk bundled: $p < 0.818$

Three manipulation check questions:

a. Which of the following items was **not** on the list of needed items mentioned in the message?

Batteries

Medications

Can opener

All of these were included **[SP]**

b. Which of the following was **not** on the list of what you should report to the authorities?

Suspicious items

Suspicious looking people

Suspicious activities

All of these were included **[SP]**

c. Which of the following was **not** on the list of what you should do?

Be vigilant

Speak out if you see something

Be sure you have batteries

All of these were included **[SP]**

Manipulation Check	TESS	MTurk (not bundled)	Mturk (bundled)
Question 1	0.621 (0.020) SD=0.486 N=591	0.672** (0.015) SD=0.470 N=1046	0.651 (0.015) SD=0.477 N=1022
Question 2	0.261 (0.018) SD=0.439 N=591	0.290 (0.014) SD=0.454 N=1038	0.311** (0.015) SD=0.463 N=1012
Question 3	0.611 (0.020) SD=0.488 N=591	0.684*** (0.014) SD=0.465 N=1030	0.664** (0.015) SD=0.472 N=1010

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Two-tailed tests.

Note: Significance is reported relative to TESS mean. There are no significant difference between Mturk bundled and not bundled for any of the questions.

Question 1 (Mean Correct)

TESS: $M = 0.6209814$ (0.019973). $SD = 0.4855537$. $N = 591$

Mturk Solo: $M = 0.6720841$ (0.0145223). $SD = 0.4696786$. $N = 1046$

Mturk Bundled: $M = 0.6506849$ (0.0149204). $SD = 0.4769871$. $N = 1022$

- No significant difference between mturk solo and bundled
- No significant difference between mturk bundled and TESS
- Significant difference between mturk solo and TESS ($p < 0.04$, two-tailed)

Question 2 (Mean Correct)

TESS: $M=0.260573$ (0.0180712). $SD=0.4393204$. $N=591$

Mturk Solo=: $M=0.2899807$ (0.0140906). $SD=0.4539719$. $N=1038$

Mturk Bundled: $M=0.3112648$ (0.0145618). $SD= 0.4632398$. $N=1012$

- No significant difference between mturk solo and bundled
- Significant difference between mturk bundled and TESS ($p < 0.03$, two-tailed)
- No significant difference between mturk solo and TESS

Question 3 (Mean Correct)

TESS: $M=0.6108291$ (0.0200726). $SD=0.4879752$. $N=591$

Mturk Solo: $M=0.684466$ (0.0144874). $SD=0.4649539$. $N=1030$

Mturk Bundled: $M=0.6643564$ (0.014866). $SD=0.4724489$. $N=1010$

- No significant difference between mturk solo and bundled
- Significant difference between mturk bundled and TESS ($p < 0.03$, two-tailed)
- Significant difference between mturk solo and TESS ($p < 0.01$, two-tailed)

[Experiment 11]

Turaga, Rama Mohana. 2010. "Environmental Values, Beliefs, and Behavior."

3 conditions

- [ST] Scientific data on mercury concentration in water bodies show that currently 40% of lake area in the country has mercury fish tissue concentrations exceeding environmental guidelines
- [AR] Scientific studies indicate that approximately 30% of all mercury released to the environment in the U.S. can be attributed to households such as yours.
- [CO] [Just the question below]

Given this information and the information on the previous screen, how unwilling or willing are you to engage in the following activities? Write a letter to your local newspaper supporting stronger mercury control policies. 1 (Very unwilling), 5 (Very willing)

Raw Unweighted Means by Sample and Condition

Sample	Group 2 ST	Group 3 AR	Group 1 CO
TESS	2.96 (0.07) SD=1.07 N=257	2.86 (0.07) SD=1.11 N=267	2.80 (0.07) SD=1.11 N=240
Mturk	2.87 (0.07)	2.73 (0.07)	2.60 (0.07)

	SD=1.34 N=358	SD=1.24 N=335	SD=1.21 N=324
--	------------------	------------------	------------------

Group 2 v 1:

TESS: $p < 0.108$

Mturk: $p < 0.005$

Group 3 v 1:

TESS: $p < 0.534$

Mturk: $p < 0.147$

[Experiment 12]

Wallace, Geoffrey P.R. 2011. "The Reputational Consequences of International Law and Compliance."

Abstract:

Recent research on international law suggests states use international legal agreements as a commitment device to increase the credibility of their promises by raising the reputational consequences of violations. Testing this reputational mechanism for compliance has been complicated by selection effects both in a state's decision to join an agreement in the first place, as well as its likely anticipation of the consequences of subsequent noncompliance. This project uses an experiment embedded in a national survey to estimate the effect of a legalized commitment on a country's reputation, which is measured by the ability of the country to garner support from foreign actors for future cooperative agreements.

12 conditions [See <http://www.tessexperiments.org/previousstudies.html> for details]

The United States is [If XTESS097 =7-12: also] currently thinking about signing an agreement with this country that would involve cooperation over military issues. Would you support or oppose the United States signing a military agreement with this country? 1 (strongly support), 5 (Strongly oppose)

Raw Unweighted Means by Sample and Condition

Sample	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
TESS	3.48 (0.07) SD=1.05 N=233	2.75 (0.06) SD=0.98 N=234	3.29 (0.07) SD=1.04 N=241	2.76 (0.07) SD=1.04 N=239	3.12 (0.07) SD=1.03 N=250	2.90 (0.07) SD=1.03 N=238
Mturk	3.60 (0.07) SD=1.08 N=252	2.43 (0.07) SD=1.04 N=202	3.40 (0.07) SD=1.06 N=209	2.55 (0.06) SD=0.95 N=215	3.09 (0.07) SD=1.08 N=232	2.82 (0.07) SD=1.05 N=218

Sample	Group 7	Group 8	Group 9	Group 10	Group 11	Group 12
TESS	3.17	2.93	3.14	3.07	3.05	3.02

	(0.07) SD=1.08 N=239	(0.07) SD=1.02 N=237	(0.07) SD=1.06 N=238	(0.07) SD=1.02 N=229	(0.06) SD=0.99 N=239	(0.06) SD=1.03 N=255
Mturk	3.27 (0.08) SD=1.13 N=213	2.76 (0.06) SD=0.90 N=197	3.15 (0.07) SD=1.01 N=210	2.86 (0.06) SD=0.93 N=226	3.00 (0.07) SD=1.10 N=226	2.71 (0.07) SD=0.98 N=184

Group 2 v 1:
TESS: p<0.000
Mturk: p<0.000

Group 3 v 1:
TESS: p<0.05
Mturk: p<0.048

[Experiment 13]

Berinsky, Adam J. 2011. "Unmasking Expressive Responses to Political Rumor Questions Using a List Experiment."

4 conditions.

Group 1:

We are now going to give you a list of 4 statements. Please tell us HOW MANY of them are true for you. We don't want to know which ones, just HOW MANY.

- I enjoy listening to music
- I think it should be legal for two men to marry
- I support allowing prayer in public schools
- I like getting phone calls from telemarketers

Group 2 also gets:

- I believe Barack Obama is a Muslim.

Group 3:

We now want to know if the following statement is true for you:

"I believe Barack Obama is a Muslim."

True 1
False 2
Not sure 3

Group 4:

We now want to know if the following statement is true for you:

"I believe Barack Obama is a Muslim."

True 1
False 2

Raw Unweighted Means by Sample and Condition

Sample	Group 1	Group 2	Group 3	Group 4
TESS	2.06 (0.02) SD=0.65 N=1794	2.26 (0.02) SD=0.84 N=1757	%True= 19.96 %False=52.72 %Not Sure=27.32 N=1413	%True=27.43 %False= 71.84 N=1350
Mturk	2.17 (0.02) SD=0.56 N=854	2.30 (0.02) SD=0.67 N=886	% True= 8.57 % False= 73.33 % Not sure = 18.10 N=420	%True= 14.29 %False= 85.71 N=441

Group 2 v 1:

TESS: $p < 0.000$

Mturk: $p < 0.000$

[Experiment 14]

Kleykamp, Meredith. 2011. "Social Desirability Bias (list experiment)."

- 3 conditions

Group 1 asked: Below is a list of different kinds of people who might move to a home near you. How many of these types of people would you be pleased to have move in next door to you? We don't want to know which ones, just how many in total.

- A professional musician
- A recent veteran of the wars in Iraq/Afghanistan
- Someone convicted of driving under the influence of alcohol
- Someone who teaches kindergarten

Group 2 does NOT include "A recent veteran.."

Group 3: Below is a list of different kinds of people who might move to a home near you. Please indicate whether or not you would be pleased to have the following kinds of people move in next door to you. [Yes, No]

Raw Unweighted Means by Sample and Condition

Sample	Group 1	Group 2	Group 3 (different)*
TESS	3.77	2.92	87.56%

	(0.03) SD=1.10 N=1362	(0.02) SD=0.80 N=1369	N=1218
Mturk	3.78 (0.03) SD=0.89 N=997	2.97 (0.02) SD=0.65 N=1048	86.71% N=1001

*Group 3 is the percentage responding "Yes" to: "Please indicate whether or not you would be pleased to have the following kinds of people move in next door to you: A recent veteran of the wars in Iraq/Afghanistan."

Group 2 v 1:

TESS: $p < 0.000$

Mturk: $p < 0.000$

[Experiment 15]

Parmer, John. 2011. "Smallpox Vaccine Recommendations: Is Trust a Shot in the Arm?"

Design:

The proposed research study is a post-test only randomized experiment to explore the role of trust and confidence in influencing the public's response to a smallpox outbreak. Participants will begin by reading a short preface statement describing the threat of a smallpox outbreak from a bioterrorist attack and an introduction to a fictional smallpox outbreak scenario that will follow. Following the preface statement, and prior to exposure to experimental conditions, participants will respond to a single item measuring their current knowledge of the Department of Homeland Security (DHS).

Study participants will then be randomized into one of two groups and read short messages describing a smallpox outbreak scenario, the role of DHS in such an event, and vaccine information and recommendations for members of a community believed to have been exposed to the virus. Message content will be structured to contain information that enhances one of the two pathways to cooperation as described by the Trust, Confidence, and Cooperation (TCC) Model. One message condition will aim to enhance trust by highlighting the shared values dimension of trust. A second message condition will aim to enhance confidence in the DHS to effectively manage the smallpox outbreak by highlighting past performance during recent public health crises (i.e. anthrax and SARS) as well as approaches that were successfully employed in the past to control smallpox outbreaks.

- 2 conditions

If you had to make a decision now, would you get the recommended smallpox vaccine? 1=Yes, 0=No

Raw Unweighted Means by Sample and Condition

Sample	Group 1	Group 2
--------	---------	---------

TESS	0.838 (0.02) SD=0.37 N=260	0.846 (0.02) SD=0.36 N=259
Mturk (not bundled)	0.881 (0.01) SD=0.32 N=536	0.879 (0.01) SD=0.33 N=522
Mturk bundled*	0.884 (0.01) SD=0.32 N=481	0.852 (0.02) SD=0.36 N=548

*This study was bundled in 2 different orderings, here they are combined into a single bundled group.

Group 2 v 1:

TESS: $p < 0.825$

Mturk: $p < 0.949$

Mturk bundled: $p < 0.140$

[Experiment 16]

Converse, Benjamin. 2008. With God on Our Side.

Abstract:

People often reason egocentrically about others' beliefs, using their own beliefs as an inductive guide. We designed the current study to test for enhanced egocentrism in judgments of God's beliefs compared with judgments of the Average American's beliefs in a representative sample, and to extend our understanding of the causal direction of the proposed Self-God relationship. Specifically, participants indicated their own attitudes toward abortion and same-sex marriage, as well as their estimates of the average American's and God's opinions about each of these issues. We manipulated whether participants first indicated their own attitudes or God's attitudes.

- 4 conditions (varies order of DV)

Group 1 asks for people's personal opinion on abortion, while Group 2 first asks for God's opinion on abortion, and then people's personal opinion on abortion.

Please indicate your personal opinion about abortion.

[completely pro-choice (1) – completely pro-life (2)]

Raw Unweighted Means by Sample and Condition

Sample	Group 1 (SGA)	Group 2 (GSA)	Group 3 (GAS)	Group 4 (SAG)
TESS	3.590 (0.145)	4.360 (0.149)	4.198 (0.151)	3.889 (0.149)

	SD=2.346 N=261	SD=2.345 N=247	SD=2.376 N=247	SD=2.411 N=261
Mturk	2.638 (0.099) SD=2.114 N=453	2.798 (0.103) SD=2.169 N=440	2.921 (0.105) SD=2.236 N=454	2.678 (0.103) SD=2.173 N=447

Group 2 v 1:

TESS: $p < 0.001$

MTurk: $p < 0.265$

[Experiment 17] REFERENCED AS [RACE (A)]

Denny, Kathleen. 2012. Examining the Raced Fatherhood Premium.

- 12 conditions

Stimuli:

Imagine you are the hiring manager of Innovative Marketing Solutions, Inc., a mid-size marketing firm. You are in the process of hiring a new employee to increase your staff and increase your chances of acquiring more clients. On the next screen is a brief description of the position along with a memo sent to you by the human resources (“HR”) department summarizing its interview with a recent applicant. Please review the job description and human resources memo and answer the questions on the next few screens.

[Descriptions of job candidates varied the name (that implied different races), and whether the candidate had children and was/was not very involved with them]

How hardworking do you expect [Greg/Jamal/Victor/Samuel] to be, relative to other employees in similar positions at the company? 1 Not at all hardworking, 5 extremely hardworking

Raw Unweighted Means by Sample and Condition

Sample	Group 1	Group 2	Group 3	Group 4
TESS	3.724 (0.058) SD=0.711 N=152	3.943 (0.063) SD=0.754 N=141	3.912 (0.063) SD=0.732 N=137	3.740 (0.056) SD=0.694 N=154
Mturk	3.986 (0.050) SD=0.587 N=140	4.037 (0.057) SD=0.724 N=161	3.957 (0.047) SD=0.561 N=140	3.946 (0.051) SD=0.626 N=148

Sample	Group 5	Group 6	Group 7	Group 8
TESS	3.628 (0.056)	3.662 (0.063)	3.628 (0.058)	3.637 (0.059)

	SD=0.653 N=137	SD=0.733 N=136	SD=0.744 N=164	SD=0.713 N=146
Mturk	3.875 (0.054) SD=0.689 N=160	3.93 (0.064) SD=0.766 N=143	3.824 (0.057) SD=0.716 N=159	3.931 (0.058) SD=0.704 N=145

Sample	Group 9	Group 10	Group 11	Group 12
TESS	3.835 (0.059) SD=0.676 N=133	3.770 (0.064) SD=0.785 N=152	3.832 (0.057) SD=0.670 N=137	3.786 (0.061) SD=0.702 N=131
Mturk	3.967 (0.050) SD=0.614 N=152	4.087 (0.055) SD=0.677 N=149	4.039 (0.053) SD=0.656 N=154	4.021 (0.050) SD=0.593 N=140

In tests, we will compare Group 1 to Group 2 and Group 1 to Group 5. In summary figures and other aggregated data, Group 5 will become Group 3.

Group 2 v 1:

TESS: $p < 0.011$

Mturk: $p < 0.502$

[Experiment 18] REFERENCED AS [RACE (B)]

Pedulla, David. 2011. The Mechanisms of Labor Market Discrimination.

Abstract:

Extant field-experimental research demonstrates that racial discrimination against black men persists in the U.S. labor market. Among the mechanisms proposed to explain this persistent discrimination are stereotypes about black men as hyper-masculine, threatening, violent, and criminal. However, extant research has not explicitly tested the degree to which these stereotypes impact the evaluations of job applicants and how counter-stereotypical information may reduce discrimination against black men. Do individuals in stigmatized groups, in this case black men, receive more favorable evaluations when they present reviewers with information counter to the stereotypes about their group? Or, are individuals who are part of stigmatized groups penalized when they present counter-normative, counter-stereotypical information to job evaluators? This survey experiment begins to address these questions by having respondents evaluate a job applicant where the applicant's resume is experimentally manipulated along two dimensions. The race (black vs. white) of the applicant is manipulated along one axis by using racialized names. On the second axis, I manipulate the sex (male vs. female), gender presentation (masculine vs. feminine), or sexual orientation (straight vs. gay) of the applicant by varying their participation in college activities. After reviewing the resume to which they were randomly assigned, respondents were asked to evaluate the applicant along a host of dimensions – reliability, trustworthiness, warmth, etc. – and to make hiring and salary recommendations for the applicant.

- 10 conditions

Would you recommend hiring this applicant for the position? 1=yes, 0=No

Raw Unweighted Means by Sample and Condition

Sample	Group 1	Group 2	Group 3	Group 4	Group 5
TESS	0.838 (0.031) SD=0.370 N=142	0.856 (0.031) SD=0.352 N=132	0.841 (0.031) SD=0.367 N=138	0.866 (0.029) SD=0.342 N=142	0.868 (0.028) SD=0.340 N=144
Mturk	0.92 (0.021) SD=0.272 N=175	0.930 (0.020) SD=0.256 N=171	0.937 (0.019) SD=0.244 N=174	0.872 (0.026) SD=0.335 N=172	0.901 (0.022) SD=0.299 N=192

Sample	Group 6	Group 7	Group 8	Group 9	Group 10
TESS	0.912 (0.024) SD=0.285 N=136	0.904 (0.025) SD=0.295 N=136	0.887 (0.027) SD=0.317 N=142	0.784 (0.034) SD=0.413 N=148	0.818 (0.033) SD=0.388 N=137
Mturk	0.922 (0.020) SD=0.269 N=179	0.870 (0.026) SD=0.337 N=169	0.913 (0.020) SD=0.282 N=196	0.869 (0.026) SD=0.338 N=168	0.923 (0.019) SD=0.267 N=195

Main tests are Group 1 v 2 and Group 1 v 9. Below, and in figures/tests in paper, Group 9 will be referenced as Group 3 for purposes of simplification.

Group 2 v 1:

TESS: $p < 0.681$

Mturk: $p < 0.730$

[Experiment 19]

Jackson, Natalie. 2010. "An Experiment in the Measurement of Social and Economic Ideology."

Abstract:

It has been argued that political ideology consists of more than one dimension when the concept is used to explain policy preferences. These arguments are based on analyses of policy preferences that utilize dimension-reduction techniques to find at least two dimensions of the liberal-conservative scale at work—most frequently social and economic dimensions. However, no one has demonstrated whether individuals think of their ideological identifications in two dimensions. Do respondents' provide different self-placements for economic issues as compared to social issues? This paper uses data from a national survey experiment that directly measures the social and economic dimensions of ideology to determine whether respondents think of their ideological views and report their self-placement differently on social and economic issues; and whether the

two self-placement measures of social and economic ideology are more accurate predictors of policy preferences than the single measure.

Control: On a scale of political ideology, individuals can be arranged from strongly liberal to strongly conservative. Which of the following categories best describes your views?

Treatment (economic and social issue ideology randomized)

- On a scale of political ideology, individuals can be arranged from strongly liberal to strongly conservative. When thinking about your views on *economic* issues, which of the following categories best describes your views? “Economic issues” are questions of how to distribute resources among people within a society.
- On a scale of political ideology, individuals can be arranged from strongly liberal to strongly conservative. When thinking about your views on *social* issues, which of the following categories best describes your views? “Social issues” are problems that affect many or all members of society, and often involve cultural or moral values.
- Now, considering your responses to the previous two questions, which of the following categories best describes your views overall?

Options 1-7:

Strongly liberal, liberal, slightly liberal, middle of the road, slightly conservative, conservative, strongly conservative (don't think in these terms removed).

Raw Unweighted Means by Sample and Condition

Sample	Control, Ideology (%)	Treatment, Ideology (%)	Treatment, Economic Issues Ideology (%)	Treatment, Social Issues Ideology (%)
TESS	4.236 (0.074) SD=1.638 N=496	4.488 (0.073) SD=1.605 N=486	4.618 (0.074) SD=1.606 N=476	4.439 (0.076) SD=1.687 N=474
Mturk	3.237 (0.069) SD=1.588 N=528	3.323 (0.076) SD=1.647 N=471	3.711 (0.082) SD=1.798 N=477	3.006 (0.078) SD=1.715 N=481

Group 2 v 1:

TESS: $p < 0.015$

Mturk: $p < 0.402$

[Experiment 20]

Bergan, Daniel. 2012. *The Flexible Correction Model and Party Labels*.

Abstract:

The Flexible Correction Model holds that when made aware of potential sources of bias, people use naive theories to correct for that bias. We tested whether people instructed to correct for the influence of party labels attempt to correct for those biases, and if these attempts at correction are moderated by theories about the influence of party labels. Subjects were exposed to a short reading about a proposed health reform in which party labels attached to the reform were randomly assigned. Subjects were also randomly assigned to bias correction instructions or no instructions.

The experiment consists of 6 conditions: 3 (party label: Democrats, Republicans, or none) X 2 (bias correction instructions, no instructions). Subjects will be first exposed to a news article about a novel policy proposal claimed to be cut and pasted from The New York Times' online edition. Three different party labels will be attached to this policy proposal (Democratic, Republican, or "policymaker" in Congress (i.e. the last condition has no party label)). Then half of the participants will be exposed to the instruction "Please try not to let irrelevant factors influence your response or bias your judgments" before answering questions, but the other half people don't receive this instruction. Respondents will then answer questions about their attitude toward the novel policy, their perceived bias of party label, etc.

Please indicate to what extent you agree with the following statements about the Independent Payment Advisory Board, the proposed commission that would limit federal spending on health care by paying only for procedures that have demonstrated medical value.

I support establishing the Independent Payment Advisory Board (1=disagree strongly, 7=agree strongly)

Raw Unweighted Means by Sample and Condition

Sample	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
TESS	2.684 (0.118) SD=1.649 N=196	2.67 (0.121) SD=1.713 N=200	2.85 (0.118) SD=1.619 N=187	2.812 (0.113) SD=1.622 N=207	2.691 (0.112) SD=1.594 N=204	2.398 (0.119) SD=1.622 N=186
MTurk	2.892 (0.092) SD=1.553 N=287	3.307 (0.096) SD=1.666 N=300	3.112 (0.094) SD=1.618 N=295	2.94 (0.09) SD=1.615 N=319	3.427 (0.099) SD=1.674 N=288	3.017 (0.091) SD=1.572 N=301

To test main dimensions of interest, we compare Group 1 to Group 2, and then Group 1 to Group 4. Group 4 will later be referred to as Group 3 in figures for simplification.

Group 1 v Group 2:

TESS: $p < 0.936$

MTurk: $p < 0.002$

Manipulation Check

Can you recall, based on the article you read, who supports the Independent Payment Advisory Board, the proposed commission that would limit federal spending on health care by paying only for procedures that have demonstrated medical value?

- a. Republican Policymakers
- b. Democratic Policymakers
- c. Both Republican and Democratic Policymakers
- d. Neither Democrats nor Republicans were mentioned in the article
- e. I don't recall

TESS Correct	MTurk Correct
0.525 (0.014) SD=0.500 N=1206	0.692*** (0.011) SD=0.462 N=1913

***denote two-tailed significance $p < 0.001$

Attention by Sample

Table S1: Manipulation Check by Sample

Manipulation Check	TESS	MTurk (not bundled)	MTurk (bundled)
Question 1	0.621 (0.020) SD=0.486 N=591	0.672** (0.015) SD=0.470 N=1046	0.651 (0.015) SD=0.477 N=1022
Question 2	0.261 (0.018) SD=0.439 N=591	0.290 (0.014) SD=0.454 N=1038	0.311** (0.015) SD=0.463 N=1012
Question 3	0.611 (0.020) SD=0.488 N=591	0.684*** (0.014) SD=0.465 N=1030	0.664** (0.015) SD=0.472 N=1010
Question 4	0.525 (0.014) SD=0.500 N=1206	NA	0.692*** (0.011) SD=0.462 N=1913

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Two-tailed tests.

Note: Significance is reported relative to TESS mean. There are no significant difference between MTurk bundled and not bundled for any of the questions. Higher means reflect more correct responses to manipulation check questions. These results come from manipulation checks from two different studies (See Experiment 10 and Experiment 20 above for details).

We tested differences in attention between samples by comparing the percentage of correct responses to three manipulation-check questions in two of our studies – questions designed to assess if people are paying attention. (Unfortunately, the other studies examined in this paper do not have manipulation checks to analyze between samples.) The results are shown in Table S1 and reveal that the MTurk samples were actually significantly more likely to answer the questions correctly than the nationally representative sample—a finding consistent with existing research (Weinberg, Freese, and McElhattan 2014).

The Effects of Bundling Studies Together in MTurk

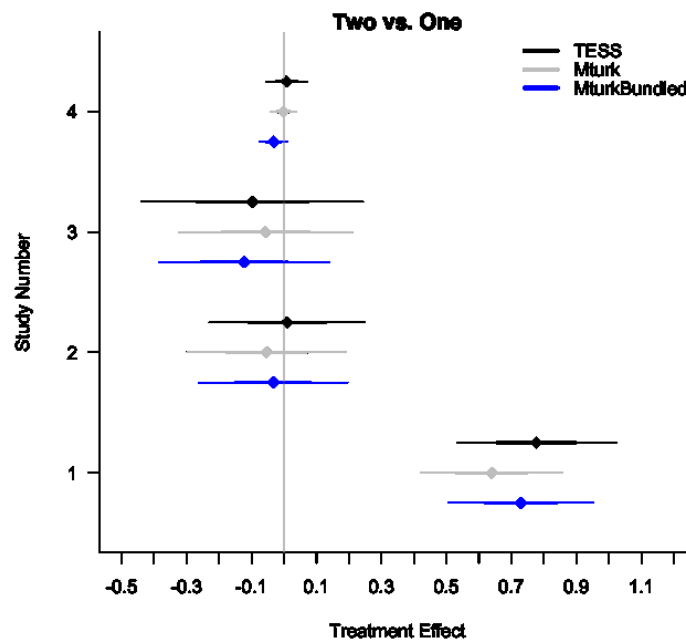


Figure S1: Bundling MTurk Studies. Experiment numbers here do not directly correspond to those in the other figures. They were renumbered for plotting purposes. Experiment 1 here corresponds to Experiment 8 in the other figures, 2 to 9, 3 to 11, and 4 to 16. Because the focus here is on comparing bundled and non-bundled MTurk studies, here we present unweighted treatment effects for each sample.

We conducted initial experiments to determine if fielding studies independently on MTurk yielded different results from bundling multiple studies into a single survey experiment to further

reduce costs. To do this, we executed a four of substantively distinct experiments in MTurk both separately and bundled together (testing also experimental arms with different orderings). The results are shown above and provide no evidence of any systematic effect of bundling. Building on this insight, the remaining MTurk experiments were implemented in a bundled fashion, with 16 studies fielded as 4 survey experiments.