

An Analysis of Surveys to Determine the Effectiveness of Bias in
Question Wording as a Method of Skewing Reported Opinions.

by
Benjamin Sharkansky

A THESIS

submitted to
Oregon State University
Honors College

in partial fulfillment of
the requirements for the
degree of

Honors Baccalaureate of Science in Mathematics and Economics
(Honors Scholar)

Presented May 24, 2019
Commencement June 2019

AN ABSTRACT OF THE THESIS OF

Benjamin Sharkansky for the degree of Honors Baccalaureate of Science in Mathematics and Economics presented on May 24, 2019. Title: An Analysis of Surveys to Determine the Effectiveness of Bias in Question Wording as a Method of Skewing Reported Opinions.

Abstract approved: _____

Clear and unbiased question wording is important to obtaining consistent and accurate survey results. However, if the goal is not to obtain accurate results but instead to obtain a specific result, then altering question wording may be effective. To examine this possibility, we created a survey with 9 different versions of a question, one being an unbiased control and the rest being intended to skew the results a particular way. The results show that while the question wording does indeed matter, the exact impact is harder to manipulate. While the results fail to demonstrate that survey results can be consistently skewed a particular way, they do demonstrate that it is feasible to skew results using these methods.

Key Words: statistics, statistical analysis, data, surveys, web surveys, public opinion, polling, question wording, MTurk, bias

Corresponding e-mail address: sharkanb@oregonstate.edu

©Copyright by Benjamin Sharkansky
May 24, 2019

An Analysis of Surveys to Determine the Effectiveness of Bias in
Question Wording as a Method of Skewing Reported Opinions.

by
Benjamin Sharkansky

A THESIS

submitted to

Oregon State University

Honors College

in partial fulfillment of
the requirements for the
degree of

Honors Baccalaureate of Science in Mathematics and Economics
(Honors Scholar)

Presented May 24, 2019
Commencement June 2019

Honors Baccalaureate of Science in Mathematics and Economics project of Benjamin Sharkansky presented on May 24, 2019.

APPROVED:

Katherine McLaughlin, Mentor, representing the Oregon State University Statistics Department

Sarah Emerson, Committee Member, representing the Oregon State University Statistics Department

Juan Restrepo, Committee Member, representing the Oregon State University Mathematics Department

Toni Doolen, Dean, Oregon State University Honors College

I understand that my project will become part of the permanent collection of Oregon State University, Honors College. My signature below authorizes release of my project to any reader upon request.

Benjamin Sharkansky, Author

Introduction:

We live in an increasingly data-driven world. Generally, the goal of this data-gathering is obtaining as much accurate information about a subject as possible. However, there are also instances where this may not be true. Take political polling. While most of it is done to, again, accurately estimate support for a particular candidate or policy, also consider the potential influence that such information has on politicians and voters. If a politician knows that a majority of their constituents support a policy, they may be more inclined to back that policy, or be more able to rally their supporters. On the other hand, the possibility of losing an election could motivate voters to turn out, so weaker visible support for a policy may be preferred for a stronger election day performance. As a result, there is an incentive to create polling data which expresses clear support for or against a particular policy. However, consider that this probably would not happen with an issue that is known to have widespread support, or no support at all, due to the reversal of support being a distinct anomaly. As a result, focus falls onto issues which are more evenly split in terms of support. Furthermore, in terms of legitimizing skewed results, it is better to have actual results gathered instead of unsupported, falsified data. This leaves the method of skewing data down to the questions used to gather them. By creating a survey for a controversial issue and narrowing the answers down to specific options, we can measure whether questions can reliably skew which opinions people express.

Background Research:

Starting with some establishing literature, Schuman and Presser refer to question wording experiments during the 1940s, which found that slight shifts in wording could cause major shifts in responses, and seek to expand upon them with improved methodology (Schuman and Presser 152-153). One experiment which they repeat compares how many people answer yes or no to the following questions, “Do you think the United States should allow public speeches against democracy” versus, “Do you think the United States should forbid public speeches against democracy” (Schuman and Presser 155). Note that a ‘yes’ response to the first question is equivalent to a ‘no’ response to the second. In the 1940s experiment, the results were that 25% of people supported allowing speech against democracy while 46% were against forbidding it, and the 1974 repeat found those levels of support to be 56% and 72% respectively (Schuman and Presser 155). While this is a big result, observe that the questions are vague. What constitutes speech against democracy? Does this include simple criticism of America's style of democracy, or does it only apply to violent revolutionary rhetoric? Allowing people to decide what they think the question means is a relevant factor, but not specifically addressed. Aware of this, Schuman and Presser do further tests on a variety of issues, experimenting with question length, formality, the addition of no opinion options, screening for familiarity, and the interaction effect of these things with education (Schuman and Presser 157-162). With the results just mentioned, they found that people with more education were less often swayed by the wording, but interactions with other terms were less significant (Schuman and Presser 167-168). In all, Schuman and Presser

tried a variety of general approaches with a variety of issues and got results which merited further examination.

Moving on to other kinds of questioning and wording and how they affect survey results, Kalton et al. surveyed issues which respondents were aware of, but not particularly familiar with, specifically British transportation law (Kalton et al. 150). They discuss the effects of including perceived alternatives in questions, finding that the perception of alternatives was sufficient to massively alter the survey results (Kalton et al. 151). This study demonstrated that adding additional information, or at least seeming to, altered survey results almost as much as Schuman and Presser did, especially when it came to offering a substantive alternative to a particular policy (Kalton et al. 159). Adding a neutral or no opinion option did as well. As for more controversial issues, Kenneth Rasinski examines question wording effects on responses regarding government spending on a wide variety of political issues (Rasinski 390). He found that responses differed by as few as four percentage points, to as many as forty, with 'welfare' being a particularly effective term for generating opposition (Rasinski 391). All his tests showed statistical significance, indicating that people can be swayed on less mundane issues as well (Rasinski 393). Moving on from wording specifically, Phillips and Clancy address the social desirability of certain answers and how that can influence responses. They found that people who report a certain trait as being desirable are more likely to report having that trait (Phillips and Clancy 927). How exactly this would impact other questions is unclear, but when discussing surveys about controversial issues, it merits consideration.

To examine the effects of question wording in the context of skewing responses to controversial topics, we created surveys via Oregon State Qualtrics, and posted them online via Amazon MTurk. One significant difference between these surveys, compared to all the surveys just mentioned, is that they were administered in a web environment, specifically Amazon MTurk. They also provided small financial incentives for completion. The following research papers address quirks of these formats and how they can impact results. Mainly in the context of online student evaluations, Nulty highlights issues in web surveys, particularly lower response rates, and gives suggestions for mitigating them (Nulty 302). These include sending numerous reminders, assuring anonymity, as well as providing incentives for taking the survey (Nulty 305). However, Nulty's scenario has a fixed number of surveys distributed and the response rate is the proportion returned. This does not apply to Amazon MTurk where a desired number of respondents is predetermined and the survey stays up until enough people have taken it. Furthermore, anyone with an Amazon worker account who meets a survey's requested demographics is free to take the survey until the number of respondents is met, creating a self-selecting, non-probability sample of MTurk users. As a result, a narrower focus is needed to determine potential issues with MTurk.

Berinsky et al. specifically focus on MTurk and how samples from it compare to those gathered using other methods, especially face to face surveys, in the context of political science (Berinsky et al. 353). They discuss the validity of MTurk samples, particularly how well they generalize to the U.S population compared to convenience samples which are usually taken in political science studies (Berinsky et al. 352). Replicating old political science surveys, they found that U.S MTurk respondents were almost always more racially diverse, engaged, and liberal than the other respondents, though they didn't always identify as Democrats (Berinsky et al. 358-360). Other demographics, including religious affiliation and marital status, differed more heavily and were less diverse for the MTurk respondents (Berinsky et al. 358). Overall, MTurk samples were more representative of the U.S population than a typical convenience sample (Berinsky et al. 366). It should also be noted that while Berinsky et al. mention how Amazon MTurk uses financial incentives to drive respondents, they discuss it only in the context of quickly and cheaply getting results (Berinsky et al. 353). The final paper reviewed is by Ryu et al., which examines the effects of incentives, both in terms of response rate and, more importantly, sample composition (Ryu et al. 90). In general, the effects were minimal, and might even be beneficial in some cases by potentially drawing in traditionally underrepresented groups, especially those with lower income (Ryu et al. 95). The main concern was that incentives might yield more favorable responses towards the survey sponsor than there would be normally (Ryu et al. 95). Overall though, the quality of data is generally unaffected.

Methods:

Pre-survey design:

In summary, the papers reviewed unanimously agree that question wording impacts answers given. These results are not very surprising as asking different questions should yield different answers. What none of these papers examine however, is whether there is consistency in which way the responses skew when asking about specific policies. They also focus on less controversial issues or short questions that people probably do not put as much thought into. Where research in this paper differs is that it specifically selects for controversial issues that people are probably aware of and already have an opinion on. For example, Kalton et al. took mundane issues where there was a consensus, took surveys, and made them appear almost evenly split. This paper intends to do the opposite and attempt to harness the effect of question wording to generate some desired results.

The reason for this pre-survey is that we cannot simply look at opinion polls for various political issues because while Amazon MTurk results might be more representative of the United States population than a convenience sample, they nevertheless do not generalize to it (Berinsky et al. 356). Therefore, we need to establish our own baseline for MTurk users' beliefs in order to make comparisons, rather than referring to existing polling data from, say, Gallup, which has a different group of respondents. Since the survey itself has nothing to do with Oregon State University, issues with incentives creating more positive responses, as

examined by Ryu et al., should be non-existent. Furthermore, with an evenly split issue, skewing from social desirability bias is unlikely because there should not be any response that people consider more socially desirable on average.

As required by Oregon State's Institutional Review Board (IRB), the survey's first part was a consent form. If participants did not consent to be studied, they were disqualified and no data were recorded. If they consented, they then received mandatory background and demographic questions. These were mainly included to make the survey look more like a typical survey that appears on these sites, but also serve as a method for comparing typical respondents in this survey compared to the skewed follow-up survey. After this, they were informed that they would not be disqualified from the survey for anything and asked to answer honestly. After this was when they reached the questions on our selected subjects.

In order to reasonably compare the subjects, they needed to be somewhat similar. Since this survey is intended to measure general attitudes towards issues and policies, we implemented questions designed to measure peoples' interest in said issues, including whether they were aware of them at all. Before answering questions on a subject, participants were asked whether they were familiar with it. If a participant stated that they were not familiar with a subject, the survey skipped questions about it and moved to the next one. This is for two reasons. First, this is a pre-survey, which serves only as a metric for selecting which question to focus on in a follow-up survey. Second, since MTurk's surveys are paid, respondents want to complete them quickly. If someone was unaware of the issues surrounding a subject, it is not in their interest to be quizzed on it. We preferred to mitigate risk of their losing interest in the survey rather than forcing them to answer everything. Note that this does introduce potential non-response bias by allowing people to skip questions, even if they actually are aware of the subjects. However, the survey is short enough that subjects should not lose all interest before finishing, minimizing this bias. Furthermore, as cynical as this sounds, while the survey explicitly states that it does not disqualify consenting participants, I personally have taken numerous paid surveys which made and broke this same promise, so participants might be inclined to answer all questions anyway to avoid this. While it should be noted that none of the papers reviewed examined any format quite like this, the design is rooted in creating as user-friendly a survey as possible.

For the actual questions, they all followed the same formula. There were two questions about opposing, though not necessarily opposite, policies in the following form: "Would you support policy X in your state of residence?" State of residence is specified because the United States has state and federal levels for laws and people might be more inclined to implement their ideas at the state level while not necessarily feeling comfortable having them enforced federally. Furthermore, state level is the level where people usually vote on specific policies with ballot measures. As for possible responses, we kept them neutral by giving people an option to say that they would support a policy, they would not

support a policy, that they had no strong opinion, or that they prefer not to answer. There were optional free response questions afterwards which invited participants to voice any issues, confusion, or complaints about the question wording, and whether they felt their views were represented.

The pre-survey consisted of questions on the following topics: abortion, the death penalty, drug decriminalization, euthanasia, gay marriage, gun control, and vaccination. This is not very many issues, but including more could have been problematic because respondents might have started losing interest in a longer survey. The motivation for picking these topics is due to them, with a couple exceptions, having a combination of being simple to understand, having relatively binary positions associated with them, and a tendency to draw controversy. In my questioning, I'm assuming that participants likely have some familiarity with the subject and issues at hand. Therefore, quizzing people on more obscure issues, or even particularly nuanced ones, is not consistent with our goals. As a result, most of the issues are not, for example, economic in nature. Getting an informed opinion on a tax law is difficult since tax laws are incredibly complex, have far reaching consequences, and have potential conflicts of interest for anyone who benefits. Getting an informed opinion on euthanasia however is much less so, due in part to it being a primarily ethical issue, so we should avoid seeing a repeat of swaying opinions like with Kalton et al.

Not all of these questions were fully intended to be candidates for the most controversial issue. The question on mandatory vaccination is intended to serve as an indicator for how non-typical the sample population is. To elaborate, vaccination rates in the United States are generally high, with around 90% of children receiving full sets of immunizations by age 3 (CDC 2016). Therefore, if there is an unusual lack of support for vaccination programs, there is reason to suspect that either our questions were flawed, or our sample population is atypical. In a similar vein, but less extreme, the question about gay marriage was expected to draw a clear majority of support because a majority of U.S voters have supported it for years now and Berinsky stated that MTurk users are more liberal than average (Pew Research Center 2017).

Skewed Survey Design:

The follow-up survey was designed to be much shorter. While the question variety was greater, with 9 total policy questions which could be answered, participants were randomly assigned only one policy question after answering the same mandatory demographic and consent questions as in the pre-survey. Each of these, besides a single control question, was designed to skew the responses.

With the death penalty being the most split issue overall in the pre-survey, as will be shown in the next section, it became the topic focused on. Half of the biased questions were

intended to skew results for supporting a ban on the death penalty, and the rest were intended to skew results against supporting a ban on the death penalty. To control for unintentional bias in the question writing, the 'for' and 'against' questions were designed to mirror one another, and the text and questions leading up to the policy questions were almost identical to the pre-survey. The only difference was some extra text in the survey's consent document stating that the participants were being given incomplete information, which was an IRB requirement since we did not initially disclose that the survey was intended to skew responses. There was also a slight shift in the familiarity questions as participants stating that they were unfamiliar with controversies surrounding the death penalty no longer skipped the next policy question. While this introduces a new population of respondents relative to the pre-survey, it will be shown in the analysis section that the two groups are not very different overall.

In order to maintain consistency between surveys and be able to use the results gathered as a baseline, the control question in the follow-up survey was one of the two death penalty questions from the pre-survey. The one we picked was, "Would you support a ban on the death penalty as a punishment for criminals in your state of residence?" The reason for this is that death penalty policy has variation, including how frequently it's used, how difficult securing a death sentence is, how easy it is to appeal, and the execution method. This nuance cannot be fully captured in a question which is at maximum 2 sentences long with only four possible responses. However, this question still represents a common view on the death penalty, that we should not have it, and does not prompt respondents one way or another. Furthermore, the other question, "Would you support having the death penalty for violent felonies in your state of residence?", used language about violent felonies specifically and is not as neutral a question.

Rather than focusing on specific types of logic or various synonyms in the wording, the focus was on broad tweaks which could influence peoples' opinions and then basing wording on those. The tweaks selected were as follows:

1. Adding descriptors to the question.
2. Changing the question scope.
3. Adding citations.
4. Introducing blatant bias.

With our chosen issue being the death penalty, along with our interest in trying to skew opinions both ways, we came up with the following question set templates. The descriptors would introduce additional imagery that would either imply that a prisoner truly deserves the death penalty, or that they might be innocent. In this survey, our choices of imagery were violent criminals to try to sway people toward saying they deserve the death penalty, and people who didn't get proper legal representation as the imagery of those who did

not deserve the death penalty. Altering scope is trickier. For the pre-survey, we specified a survey participant’s state of residence as the level for banning death penalty. To try to dissuade people from stating that they wanted to ban the death penalty, we broadened this scope to the national level, reasoning that a more drastic policy would draw less support. For trying to get people to support banning the death penalty, we narrowed the scope down to whether they supported prosecutors actively seeking the death penalty. It is not a perfect comparison since this itself is not a policy, but there is no ‘lower level’ of death penalty to use instead. With the citations, we hoped that providing research which supported a particular answer would sway people towards selecting that answer. Finally, in the case of blatant bias, we wanted to see whether it was possible to take these ideas too far. With an obvious appeal to peoples’ emotions, would they still lean towards the answer that the question suggests, or would they rebel and indicate no opinion, decline to answer, or even go with the opposing answer? It should be noted that while these blatantly biased questions were included in the survey and analyzed, they are not intended to be considered viable options for wording in an actual survey. If this were a survey which we wanted to have taken seriously as a reflection of peoples’ opinions, such questions would be easily called out as bogus. As a result, the questions would need to be at least plausibly unbiased such that criticism appears no different from routine criticism that every methodology receives. Like the vaccination and gay marriage questions in the pre-survey, they serve as indicators.

Results:

Pre-Survey Results:

This pre-survey was approved by the Oregon State IRB and posted on Amazon MTurk at around midnight on January 10th, 2019 and concluded the following morning after about twelve hours. It had a preset goal of 500 respondents and each respondent was paid \$0.10 for completing the survey and inputting a randomly generated code. The familiarity results for each issue came back as follows. Note that this is not exactly 500 respondents. This means that a couple people completed the survey and received their incentive, but opted out of having their responses recorded.

Subject	Familiarity Percentage		Familiarity Count	
	Yes	No	Yes	No
Abortion	96.18%	3.82%	479	19
Death Penalty	92.77%	7.23%	462	36
Drug Decriminalization	86.35%	13.65%	430	68
Euthanasia	80.92%	19.08%	403	95
Gay Marriage	93.17%	6.83%	464	34
Gun Control	93.78%	6.22%	467	31
Vaccination	87.95%	12.05%	438	60

The questions about abortion had the most respondents, 479 out of 498, so it was the issue that participants were most familiar with. The familiarity gaps were not particularly

large overall, with only the issue with the least familiarity, euthanasia, having distinctly fewer, and it still had 403 out of 498 respondents stating that they were familiar with it.

The raw quantities and percentages for the multiple choice parts of the policy questions are as follows. For information on what exactly these questions and responses were, reference the Pre-Survey Questions section of the appendix. ‘For’ indicates that the question asked about a policy supporting an issue, i.e. allowing abortion, and ‘Against’ indicates that the question asked about a policy opposing an issue, i.e. not having gun control. ‘Yes’ indicates support for a policy, ‘No’ indicates opposition to a policy.

Subject	Response Percentage				Response Count				
	Yes	No	No opinion	Decline	Yes	No	No opinion	Decline	Total
Abortion For	64.93%	25.05%	8.56%	1.46%	311	120	41	7	479
Abortion Against	30.69%	60.75%	6.68%	1.88%	147	291	32	9	479
Death Penalty For	53.46%	33.98%	12.12%	0.43%	247	157	56	2	462
Death Penalty Against	36.80%	48.27%	13.64%	1.30%	170	223	63	6	462
Drug Decriminalization For	53.26%	36.05%	9.07%	1.63%	229	155	39	7	430
Drug Decriminalization Against	28.60%	60.23%	8.84%	2.33%	123	259	38	10	430
Euthanasia For	67.99%	21.59%	9.18%	1.24%	274	87	37	5	403
Euthanasia Against	23.57%	63.52%	11.17%	1.74%	95	256	45	7	403
Gay Marriage For	70.26%	21.12%	8.62%	0.00%	326	98	40	0	464
Gay Marriage Against	21.98%	69.83%	7.76%	0.43%	102	324	36	2	464
Gun Control For	58.89%	34.69%	5.78%	0.64%	275	162	27	3	467
Gun Control Against	19.06%	73.02%	6.85%	1.07%	89	341	32	5	467
Vaccination 1	74.20%	17.35%	7.53%	0.91%	325	76	33	4	438
Vaccination 2	85.39%	6.39%	6.85%	1.37%	374	28	30	6	438

While not exact, these responses tend to mirror each other pretty closely in the for/against format, though there are some discrepancies, not counting vaccination where a majority of respondents were expected to answer yes to both questions. Some of this can probably be explained by the question wording being subpar. For drug decriminalization, the question about criminalizing non-medical drug use explicitly excluded alcohol, tobacco, and caffeine from the ban, but could also be reasonably be interpreted to cover banning over-the-counter painkillers or allergy medication if participants thought that non-medical meant non-pharmaceutical. See the appendix for the exact wording. Similarly, with that interpretation, they could also be supportive of permitting those drugs, but also be against medical usage of opioids, which do not neatly fit into either of the policies presented. The question was purposely kept brief in order to prevent too much confusion, but the issue is too complicated to fit any reasonable policy into one multiple choice question. In a way, it ended up unintentionally being another indicator question for overly complicated issues.

The desired outcome was an even split on a topic, preferably where few participants expressed no opinion. The results do not quite give this. Almost every issue had a clear majority supporting the corresponding policies, and all but one of them leaned towards more liberal policies, as shown. The death penalty was the one exception and it was not by as large a margin as the others. An even split also wasn't the only desired trait. We also wanted an issue that people were familiar with and felt strongly about. On this front however, recall that most people were aware of each issue. Since drug decriminalization has that complexity and wording issue, the death penalty having the greatest split overall meant that it was selected as our topic to focus on in the follow-up survey.

Skewed Survey Results:

The skewed follow-up survey was approved by the Oregon State IRB and posted on Amazon MTurk at midnight on April 1st 2019. It had a preset goal of 2000 respondents and each respondent was paid \$0.05 for completing the survey and inputting a randomly generated code. Gathering results took far longer than anticipated. With projections that we would not gather all 2000 respondents by May 14th 2019, the deadline for submitting review copies of this thesis, we closed the survey on the morning of May 1st 2019 for a month of total time spent gathering results. The raw quantities and percentages for the multiple choice parts of the policy questions are as follows. Note that unlike the pre-survey where the yes and no responses were expected to mirror each other on each topic's questions because the questions were directly opposed to each other, except for vaccination, here the yes and no responses always indicate the same opinion, support for and opposition to a ban on the death penalty respectively. For information on what exactly these questions and responses were, reference the Skewed Survey Questions section of the appendix.

Question	Response Percentage				Response Count				
	Yes	No	No opinion	Decline	Yes	No	No opinion	Decline	Total
Control	40.68%	38.42%	20.90%	0.00%	72	68	37	0	177
Descriptor: Oppose	38.38%	42.70%	18.92%	0.00%	71	79	35	0	185
Descriptor: Support	58.28%	23.93%	15.95%	1.84%	95	39	26	3	163
Scope: Oppose	40.11%	45.76%	13.56%	0.56%	71	81	24	1	177
Scope: Support	44.24%	34.55%	21.21%	0.00%	73	57	35	0	165
Citation: Oppose	44.59%	33.76%	19.75%	1.91%	70	53	31	3	157
Citation: Support	39.77%	42.69%	16.37%	1.17%	68	73	28	2	171
Bias: Oppose	37.34%	56.33%	6.33%	0.00%	59	89	10	0	158
Bias: Support	46.37%	39.66%	13.41%	0.56%	83	71	24	1	179
Familiarity Results	91.38%	8.62%	N/a	N/a	1400	132	N/a	N/a	1532

One notable difference between this group and the one in the pre-survey is that, again, the pre-survey results were gathered in about 12 hours, most of which was at night, while these results took a month to collect. As a result, it is likely that we got different respondents on average. This will be examined further in the Analysis section. It should also be noted that these groups are not independent. Out of 1551 respondents to the follow-up survey, 77 also took the first survey. Note that this count doesn't match the total in the bottom right corner of the table. This is because survey respondents had the option to opt out of having their responses recorded at the end of the survey without forfeiting payment, so they were counted on MTurk, but the responses are not included in this table. While this introduces potential non-response bias, it appears that fewer than 20 people took this option, so any effect should be minor.

As for the results themselves, as detailed in the skewed survey design section, we're expecting to see that the control question is evenly split, the questions marked as, 'Oppose,' have a higher percentage of no responses, and that the questions marked as, "Support," have a higher percentage of yes responses. A cursory glance at the table above seems to indicate that this does indeed occur for most questions. The control question is evenly split as hoped. The opposing and supporting questions based on descriptors and scope have higher proportions of

no and yes responses, respectively, compared to the control. Interestingly, the opposite pattern is observed for the citation questions. Both of them see more support on average for the option that they were trying to steer people away from. Finally, the heavily biased questions also prompted more responses in the direction that they were trying to skew people. Another interesting observation is that the blatantly biased question opposed to banning the death penalty had by far the lowest rate of people expressing no opinion at about 6% rather than between 15 to 20%. While these observations do appear to support this paper's hypothesis, further analysis is needed to determine whether these results are statistically significant.

Analysis:

The first, albeit brief question, is whether the demographics for the pre-survey and follow-up survey are significantly different. The demographic results for gender, age range, having children under 18, education level, and income range are all compiled into the following tables. ‘Full’ indicates the respondents in the skewed, follow-up survey, and ‘Pre’ indicates the respondents in the pre-survey, while the rows correspond to the answers selected. The questions and possible answers were consistent across surveys.

Gender/Survey	Full	Pre	Age/Survey	Full	Pre	Children Under 18/Survey	Full	Pre
Male	604	251	18 – 24	180	52	No	935	301
Female	911	242	25 – 34	554	187	Yes	588	193
Other	12	2	35 – 44	351	100	Declined	9	4
Declined	5	3	45 – 54	235	84			
			55 – 64	142	51			
			65 – 74	65	21			
			75 – 84	2	2			
			85 or older	1	0			
			Declined	2	1			

Education/Survey	Full	Pre	Income/Survey	Full	Pre
<HighSchool	12	3	Under \$10,000	92	22
HighSchool	160	38	\$10,000 - \$19,999	122	39
PostHSVoc	29	16	\$20,000 - \$29,999	184	47
SomeCollege	303	110	\$30,000 - \$39,999	176	57
Associates	142	54	\$40,000 - \$49,999	151	62
Bachelors	617	189	\$50,000 - \$74,999	322	126
Masters	230	77	\$75,000 - \$99,999	190	61
Doctorate	36	9	More than \$100,000	259	72
None	1	0	Declined	36	12
Declined	2	2			

To test whether any demographics differ significantly between surveys, we performed a Chi-squared test on the numerical portions of each table. The null hypothesis of the Chi-squared test is that the distribution of values in the columns is independent of the corresponding row. In this case, this means that we are testing the null hypothesis that the demographic responses that people gave were independent of the survey that they took. The test results are stated below. Issues with performing multiple tests for significance, as well as methods which correct for them, will be discussed later in this section.

Chi-squared test for whether respondents’ genders are independent of survey taken.

data: gendertable

X-squared = 20.09, df = 3, p-value = 0.0001626

Chi-squared test for whether respondents’ ages are independent of survey taken.

data: Agetable

X-squared = 4.9009, df = 8, p-value = 0.7681

Chi-squared test for whether respondents' having children under 18 is independent of survey taken.

data: CHU18table

X-squared = 0.30827, df = 2, p-value = 0.8572

Chi-squared test for whether respondents' education levels are independent of survey taken.

data: Eductable

X-squared = 10.907, df = 9, p-value = 0.2821

Chi-squared test for whether respondents' incomes are independent of survey taken.

data: Incometable

X-squared = 10.809, df = 8, p-value = 0.2128

From this, we conclude that the two surveys have similar demographics, with the exception that the follow-up survey had significantly more female respondents compared to the pre-survey. This should not matter too much, as it will be shown later that men and women do not appear to have a statistically significant difference in how they answer questions. Therefore, the control question being evenly split appears to be consistent, so we appear to have a good baseline for the level of support for a death penalty ban among Amazon MTurk users. With the demographics examined and lack of independence disclosed, there are two questions we seek to answer with these data.

1. Do these questions produce different answers as the background research indicates they should?
2. Do these questions skew results how they were intended to?

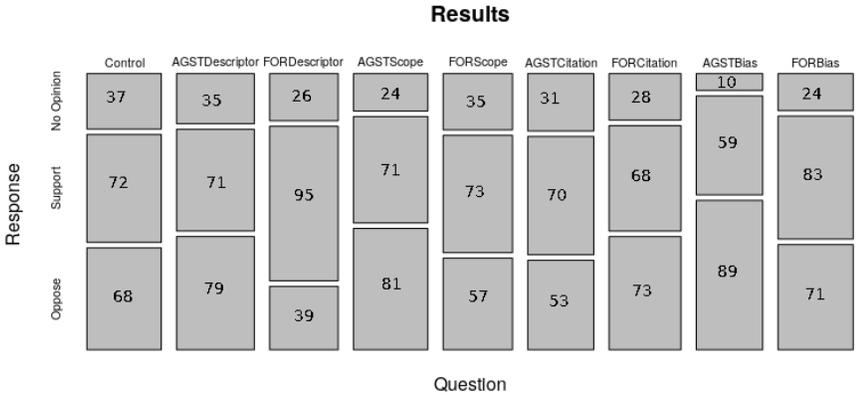
For the purposes of this analysis, 'support' refers to support for a ban on the death penalty. It is the anti death penalty option. In every question, even those meant to bias results towards having the death penalty, the first option for response was the one that expressed support for a ban. Similarly, the second option was the one which always expressed support for having the death penalty. The third and fourth questions expressed no strong opinion and a refusal to answer respectively. These options had the exact same wording throughout the entire survey regardless of question wording because they were not the focus of the experiment. Since they are very similar options and very few people refused questions anyway, they have been grouped together in the analysis. The table used to analyze the first major question is shown below. For reference, AGST indicates that the question was intended to skew results against supporting a ban on the death penalty, so we expect questions with that prefix to have a higher proportion of respondents in the 'Oppose' category relative to the control question. FOR indicates that they were intended to skew results towards supporting a ban on the death penalty, so we expect questions with that prefix to have a higher proportion

of respondents in the ‘Support’ category relative to the control question. This key is restated below. Finally, the phrase after these prefixes indicates the approach that we used to skew the survey. See the Skewed Survey Design section for more details on those.

Full Results	No Opinion	Support	Oppose	Total
Control	37	72	68	177
AGSTDescriptor	35	71	79	185
FORDescriptor	29	95	39	163
AGSTScope	25	71	81	177
FORScope	35	73	57	165
AGSTCitation	34	70	53	157
FORCitation	30	68	73	171
AGSTBias	10	59	89	158
FORBias	25	83	71	179

Support = Support Death Penalty Ban
Oppose = Oppose Death Penalty Ban
AGST → More Oppose Responses Expected
FOR → More Support Responses Expected

Each row indicates a particular survey question, the first three columns indicate the number of a type of response to it, and the final column is the total number of respondents to a question. Since the questions were randomly assigned, the totals are not equal, though they are reasonably close together. These data may also be visualized using the following mosaic plot with each bar’s count printed on it.



The question order presented in this table is the same as the order of the questions presented in the Skewed Survey section of the appendix. All the following tests in this section are answered with a Chi-squared test on the first three columns of the above table, with the rows included only if they are in the comparison that we are examining. For the first test, we examine all the rows, which checks whether the question that a survey respondent was asked is independent from what they answered on the survey, in order to answer our first question. The results are as follows.

Chi-squared test for whether respondents’ support for a ban on the death penalty are independent of the question received.

data: testtable

X-squared = 56.876, df = 16, p-value = 1.743e-06

The null hypothesis of the Chi-squared test is that the distribution of values in the columns is independent of the corresponding row. In this case, this means that we are testing the null hypothesis that the responses that people gave were independent of the questions asked. With such a tiny p-value, we have strong evidence to reject the null hypothesis that the questions and responses are independent in favor of the alternative that they are not. This answer is unsurprising as the papers examined during background research unanimously agreed that question wording has an impact.

Moving on to the second question, the third and eighth survey questions show that while it is possible to significantly skew survey results in a desired direction using question wording, most of the questions have no significant impact on the results. Performing Chi-squared tests which individually compare each question to the control, so the only rows in the table examined correspond to the control and that question, yields the following results.

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of receiving the AGSTDescriptor question compared to the control.

data: testtable[c(1, 2),]

X-squared = 0.70923, df = 2, p-value = 0.7014

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of receiving the FORDescriptor question compared to the control.

data: testtable[c(1, 3),]

X-squared = 11.44, df = 2, p-value = 0.00328

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of receiving the AGSTScope question compared to the control.

data: testtable[c(1, 4),]

X-squared = 3.4638, df = 2, p-value = 0.1769

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of receiving the FORScope question compared to the control.

data: testtable[c(1, 5),]

X-squared = 0.61015, df = 2, p-value = 0.7371

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of receiving the AGSTCitation question compared to the control.

data: testtable[c(1, 6),]

X-squared = 0.81977, df = 2, p-value = 0.6637

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of receiving the FORCitation question compared to the control.

data: testtable[c(1, 7),]

X-squared = 0.91976, df = 2, p-value = 0.6314

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of receiving the AGSTBias question compared to the control.

data: testtable[c(1, 8),]

X-squared = 18.592, df = 2, p-value = 9.18e-05

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of receiving the FORBias question compared to the control.

data: testtable[c(1, 9),]

X-squared = 3.1568, df = 2, p-value = 0.2063

This isn't evidence against the feasibility of influencing results. Again, out of eight previously untested questions, two had remarkable results which almost certainly didn't occur from random chance. It should be noted, however, that question 8 was included for comparison rather than a serious attempt at skewing results because such a question would be easily called out as biased. This shows that skewing survey results is not as simple as asking a biased question. There are other potential factors to consider, which will be examined further in the discussion section.

Now to address that doing multiple tests like this increases our probability of making a Type 1 error, i.e. we find statistical significance where there shouldn't be any, if we do not adjust the p-value that we consider statistically significant. For this, I am going to use a Bonferroni adjusted p-value. I did not set a certain number of comparisons to do ahead of time, so the Bonferroni adjusted p-value needed for significance will have to fall below the threshold that we would normally consider significant, which is the standard $\alpha = 0.05$, divided by the number of comparisons that we make. Bonferroni is known to be a conservative estimate, so if the results remain statistically significant, then we probably haven't made a Type 1 error. Not all the multiple comparisons made will be testing for the same thing, so only those that are will be counted as part of the number of comparisons to divide by. In the above set of comparisons to the control, there were 8 separate comparisons made, so the Bonferroni adjusted p-value is $\alpha = 0.05/8 = 0.00625$. Even with this adjustment, the p-values for the third and eighth questions are still statistically significant while the rest obviously remain insignificant. In short, there is no change to the conclusions.

Another multiple comparison test to check is whether the pairs of questions with the same tweak differ from each other meaningfully. While only the third and eighth questions were statistically significant, recall that this was in comparison to the control question. Since the pairs were in theory supposed to skew away from the control question in opposite directions, which appears to occur in the raw results despite lack of statistical significance, then perhaps they differ more significantly from each other than the control. Performing Chi-squared tests which compare each pair of questions yields the following results.

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of receiving the AGSTDescriptor question compared to the FORDescriptor question.

data: testtable[c(2, 3),]

X-squared = 16.266, df = 2, p-value = 0.0002937

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of receiving the AGSTScope question compared to the FORScope question.

data: testtable[c(4, 5),]

X-squared = 5.454, df = 2, p-value = 0.06541

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of receiving the AGSTCitation question compared to the FORCitation question.

data: testtable[c(6, 7),]

X-squared = 2.8612, df = 2, p-value = 0.2392

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of receiving the AGSTBias question compared to the FORBias question.

data: testtable[c(8, 9),]

X-squared = 11.245, df = 2, p-value = 0.003616

Once again, only the comparisons with the third and eight questions have statistical significance at significance level $\alpha = 0.05$, though the comparison between the fourth and fifth questions is borderline significant, and this statistical significance holds even when using the corresponding Bonferroni adjusted p-value, $\alpha = 0.05/4 = 0.0125$. As a result, this additional set of tests doesn't add any new information. A couple questions effectively skewed results, but this did not occur consistently. While the scope questions almost have statistical significance in the differences between their responses, it is not enough to conclude anything in a multiple comparison. The immediate conclusion is that perhaps using broad themes as a starting point for skewing surveys is not a good strategy, so the two successes likely can be attributed to those particular questions' writing only. Overall, while question wording does have an impact as expected, we were unable to consistently skew results.

Discussion:

While all the questions and possible responses are included in the appendix, for convenience while reading, I will reproduce questions 3 and 8, the only questions which showed statistical significance, in their entirety below.

“Question 3: Would you support a ban on the death penalty as a punishment for criminals who lacked proper legal representation in your state of residence?”

1. I would support a ban on the death penalty as a punishment for criminals who lacked proper legal representation in my state of residence.
2. I would not support a ban on the death penalty as a punishment for criminals who lacked proper legal representation in my state of residence.
3. I have no strong opinion.
4. Decline to answer.

Question 8: Would you support a ban on the death penalty as a punishment for criminals in your state of residence, even though it would mean that killers will escape justice and live in comfort that they never provided their victims?

1. I would support a ban on the death penalty as a punishment for criminals in my state of residence.
2. I would not support a ban on the death penalty as a punishment for criminals in my state of residence.
3. I have no strong opinion.
4. Decline to answer.”

The main tests compare these questions to the control question, included both in the appendix and below, which had the same four possible answers as question 8.

“Would you support a ban on the death penalty as a punishment for criminals in your state of residence?”

Regarding why these questions produced statistically significant results and their corresponding questions did not, it is difficult to say. My impression is that being somewhat vague and leaving things up to respondents' imaginations appears impactful, as with highly skewed results examined in the background research, especially the papers by Schumann, Presser, and Rasinski. Improper legal representation could refer to a variety of things, including a bad lawyer or the prosecution withholding exonerating evidence, leaving respondents to decide for themselves what it means. The biased question brings up victims, which also generates imagery. This does not however, explain why the corresponding questions 2 and 9, which are available in the appendix, were statistically insignificant, as question 2 mentions violent and dangerous criminals, another comparatively vague description with suggestive connotations, and question 9 explicitly mentions, “... wrongly convicted people.” Perhaps respondents already assumed that death row inmates were violent or dangerous, so mentioning that had little impact, while mentioning victims directly appealed to an idea of justice, however punitive, for them.

The questions using scope or citations, which were more analytical, didn't sway respondents how we wanted them to, though why citing research had an opposite, if statistically insignificant, effect is unclear, especially since they were the only questions to do so. With the questions on scope, it could be explained as peoples' views on the death penalty

not varying much on a state or federal level. For the citations, perhaps the questions seemed condescending, as though we were implying that we knew how to answer the question better than the respondents did, causing them to rebel and select the other option. I assumed that the biased questions would generate a reaction like this, but they did not. All of this is, of course, speculation, which highlights one of the limitations of statistical analysis. While we know that these correlations exist, we have no precise explanation for why they exist, and we cannot discern causality without a controlled experiment, which is hardly feasible considering how people develop their opinions over their lifetimes.

For my final thoughts on analysis, while tests were also made to compare all questions to their corresponding question in the same category, I see little need to do an in-depth look into them. The only statistically significant comparisons were with questions 3 and 8. These two questions gave results which differed from the control results with high statistical significance, and their corresponding questions gave results which did not differ from the control results with any statistical significance. Therefore, it makes sense that the corresponding questions were shown to be significantly different from each other.

As for future research paths, one focus could be to examine the questions themselves more deeply in order to figure out why a couple questions had significant results and the rest had middling results which ultimately didn't impact surveys' conclusions. The reason for doing this is that while the survey showed that it is possible to skew results in a chosen direction, doing so consistently requires more strategy than simply nudging a respondent one way. With an effective question though, it could be possible to develop a survey around it which then gets the consistently skewed responses desired. Similarly, the failed questions could be used to determine which strategies to avoid in the future.

Future research could also focus more heavily on the demographics. The average survey respondent for this research was an educated middle class female aged 25 to 44, and it is perfectly reasonable to speculate that a different average respondent would have a different average response. Would the question still skew results away from what a control survey would generate, or would it fail entirely? Future research might try distributing these questions to different targeted groups and compare how their responses differ. For the record, the following tables are a distribution of people who expressed support for a death penalty ban compared to all the demographic information gathered as well as a Chi-squared test for each.

Gender/Response	Support	Oppose	No Opinion	Age/Response	Support	Oppose	No Opinion
Male	262	255	87	18 - 24	79	63	38
Female	391	348	172	25 - 34	243	202	109
Other	8	4	0	35 - 44	149	140	62
Declined	1	3	1	45 - 54	102	109	24
				55 - 64	62	62	18
				65 - 74	25	31	9
Children Under 18/Response	Support	Oppose	No Opinion	75 - 84	1	1	0
No	414	356	165	85 or older	1	0	0
Yes	247	251	90	Declined	0	2	0
Declined	1	3	5				

Education/Response	Support	Oppose	No Opinion
<HighSchool	4	4	4
HighSchool	49	81	30
PostHSVoc	8	14	7
SomeCollege	127	121	55
Associates	61	60	21
Bachelors	282	231	104
Masters	107	88	35
Doctorate	24	9	3
None	0	0	1
Not State	0	2	0

Income/Response	Support	Oppose	No Opinion
Under \$10,000	41	31	20
\$10,000 - \$19,999	63	41	18
\$20,000 - \$29,999	79	67	38
\$30,000 - \$39,999	84	66	26
\$40,000 - \$49,999	63	66	22
\$50,000 - \$74,999	137	138	47
\$75,000 - \$99,999	79	78	33
More than \$100,000	104	112	43
Declined	12	11	13

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of their gender.

X-squared = 10.65, df = 6, p-value = 0.09981

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of their age.

data: agedem

X-squared = 24.464, df = 16, p-value = 0.07985

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of whether they have children under 18.

data: chudem

X-squared = 13.721, df = 4, p-value = 0.00824

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of their education.

data: edudem

X-squared = 36.141, df = 18, p-value = 0.006768

Chi-squared test for whether respondents' support for a ban on the death penalty is independent of their income.

data: incdem

X-squared = 23.182, df = 16, p-value = 0.109

Both education and having children under 18 have extremely significant p-values which hold up even with a Bonferroni adjusted p-value, $\alpha = 0.05/5 = 0.01$, as it appears that more educated people and those without younger children tend to support bans on the death penalty more often. It should also be noted that while the other demographic information is not statistically significant, the corresponding p-values are all still fairly small, which suggests that there might be some minor effects occurring with them as well, which could indicate a larger effect collectively. On the other hand, it is possible that this is due to correlation with other, more important, variables, which have not been analyzed or controlled for. These results indicate potential for further research in this direction. Finally, in an extension of that, it may be worthwhile to add another factor where the questions that people answer depend on what demographics they belong to. We live in an era of targeted advertising, and distributing surveys is one result that could come from that. Admittedly, this

does move away from randomness, which is key in most experimental designs, but that is why this would be considered in future research and not here.

Conclusion:

To recap briefly, we created a survey which polled 498 Amazon MTurk users to determine a political issue which evenly split opinions on that website. From those results, we created a survey about the death penalty with four question pairs and a control question. Each pair member attempted to skew the survey results in a particular direction, and each pair used the same method to try to influence the results. We used this survey to gather data from 1532 Amazon MTurk respondents in order to answer two questions. First, does question wording affect answers received, and second, can we harness this effect to skew results in the way that we want? While we were able to successfully demonstrate question wordings' potential to skew survey results in a desired direction, we were unable to do so consistently. Furthermore, we do not know why the only two questions which significantly skewed support in the direction we wanted were successful while the rest were not, especially since one of those two questions was blatantly biased and would not appear in any survey attempting to appear legitimate. The presence of a bad question is not guaranteed to create atypical results. According to our results, it is not even particularly likely. That said, the fact that we got significant results in the directions that we wanted suggests that this is possible to do. Further refinement and testing could result in effective public opinion manipulation.

Works Cited:

Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. "Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk." *Political analysis* 20, no. 3 (2012): 351-368.

Kalton, Graham, Martin Collins, and Lindsay Brook. "Experiments in wording opinion questions." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 27, no. 2 (1978): 149-161.

Nulty, Duncan D. "The adequacy of response rates to online and paper surveys: what can be done?." *Assessment & evaluation in higher education* 33, no. 3 (2008): 301-314.

Phillips, Derek L., and Kevin J. Clancy. "Some effects of 'social desirability' in survey studies." *American Journal of Sociology* 77, no. 5 (1972): 921-940.

Rasinski, Kenneth A. "The effect of question wording on public support for government spending." *Public Opinion Quarterly* 53, no. 3 (1989): 388-394.

Ryu, Erica, Mick P. Couper, and Robert W. Marans. "Survey incentives: Cash vs. in-kind; face-to-face vs. mail; response rate vs. non response error." *International Journal of Public Opinion Research* 18, no. 1 (2005): 89-106.

Schuman, Howard, and Stanley Presser. "Question wording as an independent variable in survey analysis." *Sociological Methods & Research* 6, no. 2 (1977): 151-170.

"Changing Attitudes on Gay Marriage" Pew Research Center. Last Modified June 26, 2017. <https://www.pewforum.org/fact-sheet/changing-attitudes-on-gay-marriage/>

"Vaccination coverage for selected diseases among children aged 19–35 months, by race, Hispanic origin, poverty level, and location of residence in metropolitan statistical area: United States, selected years 1998–2016" CDC. Last Modified August 9, 2018. <https://www.cdc.gov/nchs/fastats/immunize.htm>

Appendix:

Pre-Survey Questions:

Do you consider yourself to be familiar with any issues and controversies surrounding the following subject: Abortion

I am familiar with issues and controversies surrounding this subject.

I am unfamiliar with issues and controversies surrounding this subject.

Would you support legal abortion in your state of residence?

I would support legal abortion in my state of residence.

I would not support legal abortion in my state of residence.

I have no strong opinion on this issue.

Decline to answer.

Would you support an abortion ban, with possible exceptions in cases of rape, incest, or danger to the mother's health, in your state of residence?

I would support an abortion ban, with possible exceptions in cases of rape, incest, or danger to the mother's health, in my state of residence.

I would not support an abortion ban, with possible exceptions in cases of rape, incest, or danger to the mother's health, in my state of residence.

I have no strong opinion on this issue.

Decline to answer.

Do you consider yourself to be familiar with any issues and controversies surrounding the following subject: Death Penalty

I am familiar with issues and controversies surrounding this subject.

I am unfamiliar with issues and controversies surrounding this subject.

Would you support having the death penalty for violent felonies in your state of residence?

I would support implementing the death penalty for violent felonies in my state of residence.

I would not support implementing the death penalty for violent felonies in my state of residence.

I have no strong opinion on this issue.

Decline to answer.

Would you support a ban on the death penalty as a punishment for criminals in your state of residence?

I would support a ban on the death penalty as a punishment for criminals in my state of residence.

I would not support a ban on the death penalty as a punishment for criminals in my state of residence.

I have no strong opinion on this issue.

Decline to answer.

Do you consider yourself to be familiar with any issues and controversies surrounding the following subject: Drug Decriminalization

I am familiar with issues and controversies surrounding this subject.

I am unfamiliar with issues and controversies surrounding this subject.

Would you support decriminalizing all recreational drug use in your state of residence?

I would support decriminalizing all recreational drug use in my state of residence.

I would not support decriminalizing all recreational drug use in my state of residence.

I have no strong opinion on this issue.

Decline to answer.

Would you support criminalizing all non-medical drug use in your state of residence (Excluding caffeine, alcohol, and tobacco)?

I would support criminalizing all non-medical drug use in my state of residence.

I would not support criminalizing all non-medical drug use in my state of residence.

I have no strong opinion on this issue.

Decline to answer.

Do you consider yourself to be familiar with any issues and controversies surrounding the following subject: Euthanasia

I am familiar with issues and controversies surrounding this subject.

I am unfamiliar with issues and controversies surrounding this subject.

Would you support people having a legal right to be euthanized in your state of residence?

I would support people having a legal right to be euthanized in my state of residence.
I would not support people having a legal right to be euthanized in my state of residence.
I have no strong opinion on this issue.
Decline to answer.

Would you support a ban on human euthanasia in your state of residence?

I would support a ban on human euthanasia in my state of residence.
I would not support a ban on human euthanasia in my state of residence.
I have no strong opinion on this issue.
Decline to answer.

Do you consider yourself to be familiar with any issues and controversies surrounding the following subject: Gay Marriage

I am familiar with issues and controversies surrounding this subject.
I am unfamiliar with issues and controversies surrounding this subject.

Would you support legal gay marriage in your state of residence?

I would support legal gay marriage in my state of residence.
I would not support legal gay marriage in my state of residence.
I have no strong opinion on this issue.
Decline to answer.

Would you support a gay marriage ban in your state of residence?

I would support a gay marriage ban in my state of residence.
I would not support a gay marriage ban in my state of residence.
I have no strong opinion on this issue.
Decline to answer.

Do you consider yourself to be familiar with any issues and controversies surrounding the following subject: Gun Control

I am familiar with issues and controversies surrounding this subject.
I am unfamiliar with issues and controversies surrounding this subject.

Would you support a ban on assault rifles in your state of residence?

I would support a ban on assault rifles in my state of residence.
I would not support a ban on assault rifles in my state of residence.
I have no strong opinion on this issue.

Decline to answer.

Would you support removal of any restrictions on gun purchases in your state of residence?

I would support removal of any restrictions on gun purchases in my state of residence.

I would not support removal of any restrictions on gun purchases in my state of residence.

I have no strong opinion on this issue.

Decline to answer.

Do you consider yourself to be familiar with any issues and controversies surrounding the following subject: Vaccination

I am familiar with issues and controversies surrounding this subject.

I am unfamiliar with issues and controversies surrounding this subject.

Would you support vaccination for everyone, excluding those unable to be vaccinated due to medical issues or religious/ethical objections, in your state of residence?

I would support vaccination for everyone, excluding those unable to be vaccinated due to medical issues or religious/ethical objections, in my state of residence.

I would not support vaccination for everyone, excluding those unable to be vaccinated due to medical issues or religious/ethical objections, in my state of residence.

I have no strong opinion on this issue.

Decline to answer.

Would you support people having access to vaccination in your state of residence?

I would support people having access to vaccination in my state of residence.

I would not support people having access to vaccination in my state of residence.

I have no strong opinion on this issue.

Decline to answer.

Skewed Survey Questions:

Do you consider yourself to be familiar with any issues and controversies surrounding the following subject: Death Penalty

I am familiar with issues and controversies surrounding this subject.

I am unfamiliar with issues and controversies surrounding this subject.

Control

Would you support a ban on the death penalty as a punishment for criminals in your state of residence?

I would support a ban on the death penalty as a punishment for criminals in my state of residence.

I would not support a ban on the death penalty as a punishment for criminals in my state of residence.

I have no strong opinion.

Decline to answer.

Criminal Type: Oppose

Would you support a ban on the death penalty as a punishment for violent or dangerous criminals in your state of residence?

I would support a ban on the death penalty as a punishment for violent or dangerous criminals in my state of residence.

I would not support a ban on the death penalty as a punishment for violent or dangerous criminals in my state of residence.

I have no strong opinion.

Decline to answer.

Criminal Type: Support

Would you support a ban on the death penalty as a punishment for criminals who lacked proper legal representation in your state of residence?

I would support a ban on the death penalty as a punishment for criminals who lacked proper legal representation in my state of residence.

I would not support a ban on the death penalty as a punishment for criminals who lacked proper legal representation in my state of residence.

I have no strong opinion.

Decline to answer.

Scope: Oppose

Would you support a nationwide ban on the death penalty as a punishment for criminals?

I would support a nationwide ban on the death penalty as a punishment for criminals.

I would not support a nationwide ban on the death penalty as a punishment for criminals.

I have no strong opinion.

Decline to answer.

Scope: Support

Would you support having prosecutors avoid seeking the death penalty as a punishment for criminals in your state of residence?

I would support having prosecutors avoid seeking the death penalty as a punishment for criminals in my state of residence.

I would not support having prosecutors avoid seeking the death penalty as a punishment for criminals in my state of residence.

I have no strong opinion.

Decline to answer.

Citation: Oppose

Would you support a ban on the death penalty as a punishment for criminals in your state of residence, given that, according to Marquis in the Journal of Criminal Law & Criminology (2005), key criticisms of the death penalty, including the innocence of numerous inmates, are vastly overstated?

<https://scholarlycommons.law.northwestern.edu/cgi/viewcontent.cgi?article=7185&context=jclc>

I would support a ban on the death penalty as a punishment for criminals in my state of residence.

I would not support a ban on the death penalty as a punishment for criminals in my state of residence.

I have no strong opinion.

Decline to answer.

Citation: Support

Would you support a ban on the death penalty as a punishment for criminals in your state of residence, given that, according to Gross et al. in the Proceedings of the National Academy of Sciences (2013), at least 4% of death row inmates are innocent?

<https://www.pnas.org/content/pnas/111/20/7230.full.pdf>

I would support a ban on the death penalty as a punishment for criminals in my state of residence.

I would not support a ban on the death penalty as a punishment for criminals in my state of residence.

I have no strong opinion.

Decline to answer.

Bias: Oppose

Would you support a ban on the death penalty as a punishment for criminals in your state of residence, even though it would mean that killers will escape justice and live in comfort that they never provided their victims?

I would support a ban on the death penalty as a punishment for criminals in my state of residence.

I would not support a ban on the death penalty as a punishment for criminals in my state of residence.

I have no strong opinion.

Decline to answer.

Bias: Support

Would you support a ban on the death penalty as a punishment for criminals in your state of residence and stop the state-sanctioned murder of numerous wrongly convicted people?

I would support a ban on the death penalty as a punishment for criminals in my state of residence.

I would not support a ban on the death penalty as a punishment for criminals in my state of residence.

I have no strong opinion.

Decline to answer.

Source Code:

Stata .do file for cleaning raw data:

```
drop v1
drop v2
drop v3
drop v4
drop v5
drop v6
drop v7
drop v8
drop v9
drop v10
drop v11
drop v27
drop v28
drop v29
keep in 23/1554

unab allvars: *

destring *, replace

rename v12 Gender
rename v13 Age
rename v14 CHU18
rename v15 Education
rename v16 Income
rename v17 Familiar
rename v18 Control
rename v19 AGSTCriminalType
rename v20 FORCriminalType
rename v21 AGSTScope
rename v22 FORScope
rename v23 AGSTCitation
rename v24 FORCitation
rename v25 AGSTBias
rename v26 FORBias

replace Control = 0 if missing(Control)
```

replace AGSTCriminalType = 0 if missing(AGSTCriminalType)
replace FORCriminalType = 0 if missing(FORCriminalType)
replace AGSTScope = 0 if missing(AGSTScope)
replace FORScope = 0 if missing(FORScope)
replace AGSTCitation = 0 if missing(AGSTCitation)
replace FORCitation = 0 if missing(FORCitation)
replace AGSTBias = 0 if missing(AGSTBias)
replace FORBias = 0 if missing(FORBias)

gen SurveyType = 0
gen SupportBan = 0
gen AgainstBan = 0
gen OverallResponses = 0
gen Control_ANSWERED = 0
gen AGSTCriminalType_ANSWERED = 0
gen FORCriminalType_ANSWERED = 0
gen AGSTScope_ANSWERED = 0
gen FORScope_ANSWERED = 0
gen AGSTCitation_ANSWERED = 0
gen FORCitation_ANSWERED = 0
gen AGSTBias_ANSWERED = 0
gen FORBias_ANSWERED = 0

gen Control_SUPPORT = 0
gen AGSTCriminalType_SUPPORT = 0
gen FORCriminalType_SUPPORT = 0
gen AGSTScope_SUPPORT = 0
gen FORScope_SUPPORT = 0
gen AGSTCitation_SUPPORT = 0
gen FORCitation_SUPPORT = 0
gen AGSTBias_SUPPORT = 0
gen FORBias_SUPPORT = 0

gen Control_AGAINST = 0
gen AGSTCriminalType_AGAINST = 0
gen FORCriminalType_AGAINST = 0
gen AGSTScope_AGAINST = 0
gen FORScope_AGAINST = 0
gen AGSTCitation_AGAINST = 0
gen FORCitation_AGAINST = 0
gen AGSTBias_AGAINST = 0
gen FORBias_AGAINST = 0

replace SupportBan = 1 if Control == 1
replace SupportBan = 1 if AGSTCriminalType == 1
replace SupportBan = 1 if FORCriminalType == 1

replace SupportBan = 1 if AGSTScope == 1
replace SupportBan = 1 if FORScope == 1
replace SupportBan = 1 if AGSTCitation == 1
replace SupportBan = 1 if FORCitation == 1
replace SupportBan = 1 if AGSTBias == 1
replace SupportBan = 1 if FORBias == 1

replace AgainstBan = 1 if Control == 2
replace AgainstBan = 1 if AGSTCriminalType == 2
replace AgainstBan = 1 if FORCriminalType == 2
replace AgainstBan = 1 if AGSTScope == 2
replace AgainstBan = 1 if FORScope == 2
replace AgainstBan = 1 if AGSTCitation == 2
replace AgainstBan = 1 if FORCitation == 2
replace AgainstBan = 1 if AGSTBias == 2
replace AgainstBan = 1 if FORBias == 2

replace OverallResponses = 1 if Control == 1
replace OverallResponses = 1 if AGSTCriminalType == 1
replace OverallResponses = 1 if FORCriminalType == 1
replace OverallResponses = 1 if AGSTScope == 1
replace OverallResponses = 1 if FORScope == 1
replace OverallResponses = 1 if AGSTCitation == 1
replace OverallResponses = 1 if FORCitation == 1
replace OverallResponses = 1 if AGSTBias == 1
replace OverallResponses = 1 if FORBias == 1

replace OverallResponses = 2 if Control == 2
replace OverallResponses = 2 if AGSTCriminalType == 2
replace OverallResponses = 2 if FORCriminalType == 2
replace OverallResponses = 2 if AGSTScope == 2
replace OverallResponses = 2 if FORScope == 2
replace OverallResponses = 2 if AGSTCitation == 2
replace OverallResponses = 2 if FORCitation == 2
replace OverallResponses = 2 if AGSTBias == 2
replace OverallResponses = 2 if FORBias == 2

replace OverallResponses = 3 if Control >= 3
replace OverallResponses = 3 if AGSTCriminalType >= 3
replace OverallResponses = 3 if FORCriminalType >= 3
replace OverallResponses = 3 if AGSTScope >= 3
replace OverallResponses = 3 if FORScope >= 3
replace OverallResponses = 3 if AGSTCitation >= 3
replace OverallResponses = 3 if FORCitation >= 3
replace OverallResponses = 3 if AGSTBias >= 3
replace OverallResponses = 3 if FORBias >= 3

```
replace Control_ANSWERED = 1 if Control > 0
replace AGSTCriminalType_ANSWERED = 1 if AGSTCriminalType > 0
replace FORCriminalType_ANSWERED = 1 if FORCriminalType > 0
replace AGSTScope_ANSWERED = 1 if AGSTScope > 0
replace FORScope_ANSWERED = 1 if FORScope > 0
replace AGSTCitation_ANSWERED = 1 if AGSTCitation > 0
replace FORCitation_ANSWERED = 1 if FORCitation > 0
replace AGSTBias_ANSWERED = 1 if AGSTBias > 0
replace FORBias_ANSWERED = 1 if FORBias > 0
```

```
replace Control_SUPPORT = 1 if Control == 1
replace AGSTCriminalType_SUPPORT = 1 if AGSTCriminalType == 1
replace FORCriminalType_SUPPORT = 1 if FORCriminalType == 1
replace AGSTScope_SUPPORT = 1 if AGSTScope == 1
replace FORScope_SUPPORT = 1 if FORScope == 1
replace AGSTCitation_SUPPORT = 1 if AGSTCitation == 1
replace FORCitation_SUPPORT = 1 if FORCitation == 1
replace AGSTBias_SUPPORT = 1 if AGSTBias == 1
replace FORBias_SUPPORT = 1 if FORBias == 1
```

```
replace Control_AGAINST = 1 if Control == 2
replace AGSTCriminalType_AGAINST = 1 if AGSTCriminalType == 2
replace FORCriminalType_AGAINST = 1 if FORCriminalType == 2
replace AGSTScope_AGAINST = 1 if AGSTScope == 2
replace FORScope_AGAINST = 1 if FORScope == 2
replace AGSTCitation_AGAINST = 1 if AGSTCitation == 2
replace FORCitation_AGAINST = 1 if FORCitation == 2
replace AGSTBias_AGAINST = 1 if AGSTBias == 2
replace FORBias_AGAINST = 1 if FORBias == 2
```

```
replace SurveyType = 1 if Control_ANSWERED == 1
replace SurveyType = 2 if AGSTCriminalType_ANSWERED == 1
replace SurveyType = 3 if FORCriminalType_ANSWERED == 1
replace SurveyType = 4 if AGSTScope_ANSWERED == 1
replace SurveyType = 5 if FORScope_ANSWERED == 1
replace SurveyType = 6 if AGSTCitation_ANSWERED == 1
replace SurveyType = 7 if FORCitation_ANSWERED == 1
replace SurveyType = 8 if AGSTBias_ANSWERED == 1
```

R code for analysis:

```
#I changed 'CriminalType' to 'Descriptor' in the labels, but not the variable names.
survey <- read_csv("~/Desktop/Skewed Survey Results.csv")
```

```
testtable <- table(survey$SurveyType, survey$Control_SUPPORT)
```

```

varlist = data.frame(survey$AGSTCriminalType_SUPPORT,
survey$FORCriminalType_SUPPORT, survey$AGSTScope_SUPPORT,
survey$FORScope_SUPPORT, survey$AGSTCitation_SUPPORT,
survey$FORCitation_SUPPORT, survey$AGSTBias_SUPPORT,
survey$FORBias_SUPPORT)
varlist2 = data.frame(survey$AGSTCriminalType_AGAINST,
survey$FORCriminalType_AGAINST, survey$AGSTScope_AGAINST,
survey$FORScope_AGAINST, survey$AGSTCitation_AGAINST,
survey$FORCitation_AGAINST, survey$AGSTBias_AGAINST,
survey$FORBias_AGAINST)

testtable = cbind(testtable, 0)
colnames(testtable) <- c("No Opinion", "Support", "Oppose")
rownames(testtable) <- c("Control", "AGSTDescriptor", "FORDescriptor", "AGSTScope",
"FORScope", "AGSTCitation", "FORCitation", "AGSTBias", "FORBias")

testtable[1, 1] = testtable[1, 1] - sum(survey$Control_AGAINST)
testtable[1, 3] = sum(survey$Control_AGAINST)

for(i in 1:8){
  testtable[1 + i, 1] = testtable[1 + i, 1] - sum(varlist[i]) - sum(varlist2[i])
  testtable[1 + i, 2] = sum(varlist[i])
  testtable[1 + i, 3] = sum(varlist2[i])
}

testtable
mosaicplot(testtable, main="Results", xlab="Question", ylab="Response")
chisq.test(testtable)

for(i in 2:9){
  print(chisq.test(testtable[c(1, i),]))
}

chisq.test(testtable[c(1, 2, 4, 5, 6, 7, 8, 9),])
chisq.test(testtable[c(1, 2, 4, 5, 6, 7, 9),])
chisq.test(testtable[c(2, 3),])
chisq.test(testtable[c(4, 5),])
chisq.test(testtable[c(6, 7),])
chisq.test(testtable[c(8, 9),])

#Unused Comparisons excluding "No Opinion"
chisq.test(testtable[, c(2, 3)])
chisq.test(testtable[c(1, 2, 4, 5, 6, 7, 8, 9), c(2, 3)])
chisq.test(testtable[c(2, 3), c(2, 3)])
chisq.test(testtable[c(4, 5), c(2, 3)])

```

```
chisq.test(testtable[c(6, 7), c(2, 3)])
chisq.test(testtable[c(8, 9), c(2, 3)])
```

```
for(i in 2:9){
  print(chisq.test(testtable[c(1, i), c(2, 3)]))
}
```

```
for(i in 2:9){
  print(testtable[c(1, i), c(2, 3)])
}
```

```
#Demographics Comparison
```

```
gendertable = table(survey$Gender)
gendertable = cbind(gendertable, 0)
colnames(gendertable) <- c("Full", "Pre")
rownames(gendertable) <- c("Male", "Female", "Other", "Declined")
holder = table(presurvey$Gender)
holder = cbind(holder, 0)
gendertable[1,2] = holder[1]
gendertable[2,2] = holder[2]
gendertable[3,2] = holder[3]
gendertable[4,2] = holder[4]
```

```
Agetable = table(survey$Age)
Agetable = cbind(Agetable, 0)
colnames(Agetable) <- c("Full", "Pre")
rownames(Agetable) <- c("18 - 24", "25 - 34", "35 - 44", "45 - 54", "55 - 64",
  "65 - 74", "75 - 84", "85 or older", "Declined")
holder = table(presurvey$Age)
holder = cbind(holder, 0)
Agetable[1,2] = holder[1]
Agetable[2,2] = holder[2]
Agetable[3,2] = holder[3]
Agetable[4,2] = holder[4]
Agetable[5,2] = holder[5]
Agetable[6,2] = holder[6]
Agetable[7,2] = holder[7]
Agetable[8,2] = 0 #This had no respondents in the presurvey, so we have to skip it.
Agetable[9,2] = holder[8]
```

```
CHU18table = table(survey$CHU18)
CHU18table = cbind(CHU18table, 0)
colnames(CHU18table) <- c("Full", "Pre")
rownames(CHU18table) <- c("No", "Yes", "Declined")
holder = table(presurvey$CHU18)
```

```
holder = cbind(holder, 0)
CHU18table[1,2] = holder[1]
CHU18table[2,2] = holder[2]
CHU18table[3,2] = holder[3]
```

```
Eductable = table(survey$Education)
Eductable = cbind(Eductable, 0)
colnames(Eductable) <- c("Full", "Pre")
rownames(Eductable) <- c("<HighSchool", "HighSchool", "PostHSVoc", "SomeCollege",
"Associates", "Bachelors", "Masters", "Doctorate", "None", "Declined")
holder = table(presurvey$Education)
holder = cbind(holder, 0)
Eductable[1,2] = holder[1]
Eductable[2,2] = holder[3] #10 showed up where I thought 2 should be.
Eductable[3,2] = holder[4]
Eductable[4,2] = holder[5]
Eductable[5,2] = holder[6]
Eductable[6,2] = holder[7]
Eductable[7,2] = holder[8]
Eductable[8,2] = holder[9]
Eductable[9,2] = 0 #This had no respondents in the presurvey, so we have to skip it.
Eductable[10,2] = holder[2]
```

```
Incometable = table(survey$Income)
Incometable = cbind(Incometable, 0)
colnames(Incometable) <- c("Full", "Pre")
rownames(Incometable) <- c("Under $10,000", "$10,000 - $19,999", "$20,000 - $29,999",
"$30,000 - $39,999", "$40,000 - $49,999", "$50,000 - $74,999", "$75,000 - $99,999", "More
than $100,000", "Declined")
holder = table(presurvey$Income)
holder = cbind(holder, 0)
Incometable[1,2] = holder[1]
Incometable[2,2] = holder[2]
Incometable[3,2] = holder[3]
Incometable[4,2] = holder[4]
Incometable[5,2] = holder[5]
Incometable[6,2] = holder[6]
Incometable[7,2] = holder[7]
Incometable[8,2] = holder[8]
Incometable[9,2] = holder[9]
```

```
gendertable
chisq.test(gendertable)
```

```
Agetable
chisq.test(Agetable)
```

```
CHU18table
chisq.test(CHU18table)
```

```
Eductable
chisq.test(Eductable)
```

```
Incometable
chisq.test(Incometable)
```

```
gendem <- table(survey$Gender, survey$OverallResponses)
rownames(gendem) <- c("Male", "Female", "Other", "Declined")
colnames(gendem) <- c("Support", "Oppose", "No Opinion")
gendem
chisq.test(gendem)
```

```
agedem <- table(survey$Age, survey$OverallResponses)
colnames(agedem) <- c("Support", "Oppose", "No Opinion")
rownames(agedem) <- c("18 - 24", "25 - 34", "35 - 44", "45 - 54", "55 - 64",
"65 - 74", "75 - 84", "85 or older", "Declined")
agedem
chisq.test(agedem)
```

```
chudem <- table(survey$CHU18, survey$OverallResponses)
colnames(chudem) <- c("Support", "Oppose", "No Opinion")
rownames(chudem) <- c("No", "Yes", "Declined")
chudem
chisq.test(chudem)
```

```
edudem <- table(survey$Education, survey$OverallResponses)
colnames(edudem) <- c("Support", "Oppose", "No Opinion")
rownames(edudem) <- c("<HighSchool", "HighSchool", "PostHSVoc", "SomeCollege",
"Associates", "Bachelors", "Masters", "Doctorate", "None", "Not Stated")
edudem
chisq.test(edudem)
```

```
incdem <- table(survey$Income, survey$OverallResponses)
colnames(incdem) <- c("Support", "Oppose", "No Opinion")
rownames(incdem) <- c("Under $10,000", "$10,000 - $19,999", "$20,000 - $29,999",
"$30,000 - $39,999", "$40,000 - $49,999", "$50,000 - $74,999", "$75,000 - $99,999", "More
than $100,000", "Declined")
incdem
chisq.test(incdem)
```

