Primer on Polling:
A tad bit more than just asking questions.

by Scott H. Huffmon, PhD

Professor of Political Science

Director, Center for Public Opinion & Policy Research at

WINTHROP UNIVERSITY

Which is the BEST Pain Reliever?

Which is the BEST Pain Reliever?

Which is the BEST Pain Reliever?

Which is the BEST Pain Reliever?

The choice is clear!

Or is it ...The BEST Pain Reliever?

	Special Grip Bottle	Easy on the Stomach	Pain Relief Strength
Brand W	yes	good	very good
Brand X	yes	good	good
Brand Y	yes	poor	excellent
Brand Z	no	excellent	excellent

HOW you ask questions matters!

Priming

- **Abortion**
- Abortion
- **Abortion**

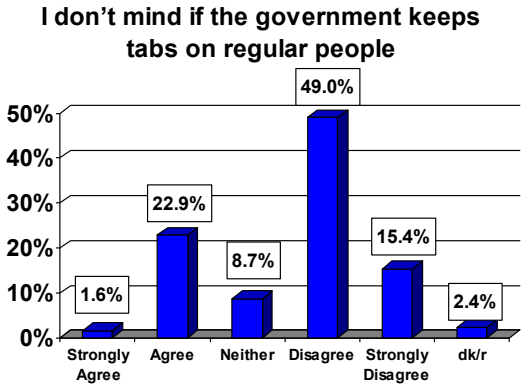
Priming

- Putting something in the forefront of the respondent's mind so it becomes more salient to him or her when producing an opinion
- Usually unintentional
- Often related to question order

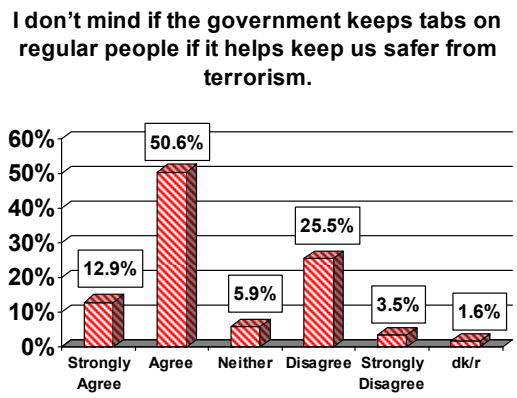
Framing

- When something about the structure or presentation of the question predispose respondents to answer in a certain way

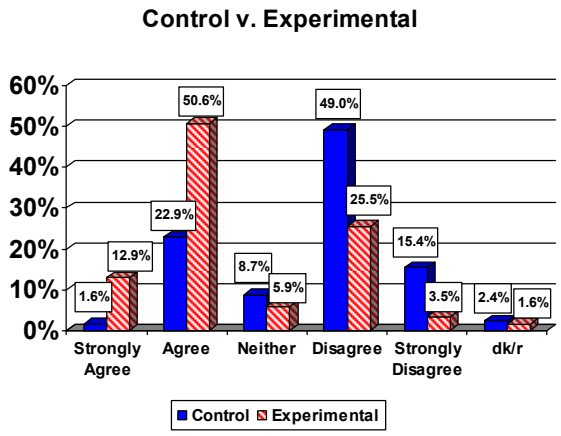
Framing



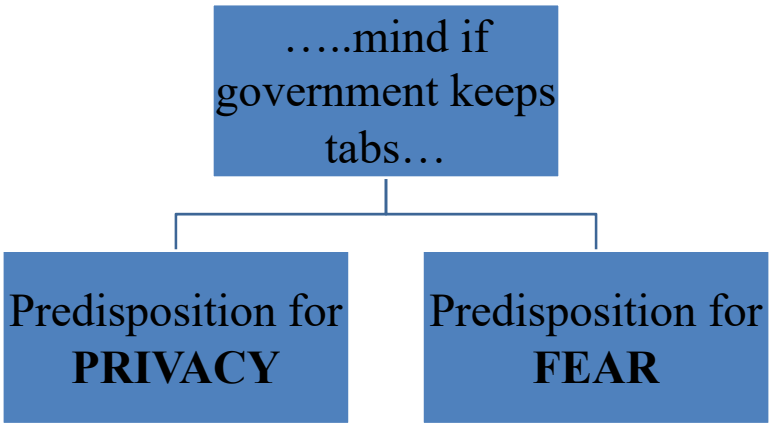
Framing



Framing



Producing an Opinion



Producing an Opinion

- "Do you think the United States should let communist reporters from other countries come in here and send back to their papers the news as they see it?"
- In one survey experiment from the 1970s, only 36.5% of respondents were willing to let Communist reporters in the US.

Producing an Opinion

- However, among those who were **first** asked, "Do you think a communist country like Russia should let American newspaper reporters come in and send back to America the news as they see it?"
- …… 73.1% said Communist reporters should be allowed in the US when asked *that* question.

Producing an Opinion

Word Choice

- Terms can be loaded – often unintentionally so.
- Using a certain term my trigger a gut reaction or make one attitude more salient than another

Word Choice

- *Pew: "As you may know, the government is potentially investing billions to try and keep financial institutions and markets secure. Do you think this is the right thing or the wrong thing for the government to be doing?" 9/08*
- *Result? Right 57%, Wrong 30%*

Word Choice

- *LAT/Bloomberg: "Do you think the government should use taxpayers' dollars to rescue ailing private financial firms whose collapse could have adverse effects on the economy and market, or is it not the government's responsibility to bail out private companies with taxpayers' dollars?" 9/08*
- *Result? Should 31%, Should not 55%*

Word Choice

- *ABC/Post: "Do you approve or disapprove of the steps the Federal Reserve and the Treasury Department have taken to try to deal with the current situation involving the stock market and major financial institutions?" 9/08*
- *Result? Approve 44%, Disapprove 42%*

Word Choice

Poll	Wording	Favor	Don't Favor
Pew	"invest"	57%	30%
LAT/Bloomberg	"bailout"	31%	55%
ABC News/Post	"steps"	44%	42%

Holocaust Denial?

- November 1992 Roper poll for American Jewish Committee
- "Does it seem possible or does it seem impossible to you that the Nazi extermination of the Jews never happened?“
  - Possible: 22.1%
  - Impossible: 65.4%
  - Don’t Know: 12.4%

Holocaust Denial?

- Gallup reproduced the awkward question, BUT with a follow up:
  - "Just to clarify, in your opinion, did the Holocaust definitely happen, probably happen, probably not happen, or definitely not happen?"
  - Of those who had said it was possible that the Holocaust never happened, 98.0 percent changed their position, with 96.7 percent saying it did happen and 1.3 percent saying "don't know."

Holocaust Denial?

- Gallup also asked a simpler version to half of their sample: “Do you doubt that the Holocaust actually happened, or not?”
- Do not doubt: 87%
- Doubt: 9%
- Unsure: 4%
- Further probing showed only 3-5% *actually* doubted; the rest were just ignorant.

Double-Barreled Questions

- Is the consumption of drugs or alcohol prevalent in your community?
- Does your company have pension and health insurance benefits?
- How satisfied are you with your pay and job conditions?
- Should government spend less money on the military and more on education?
- Do you agree with the Tea Party or do you think they’re hurting the Republican Party?

But What About Polls

Good Polls

- Size?
- Sample?
- Questions? (which we've already covered with framing, etc.)

Size

- Question: You have a population of 25,000. How many people do you need to sample to get a margin of error of +/- 4%?
- Answer: 586

Size

- Question: You have a population of 75,000. How many people do you need to sample to get a margin of error of +/- 4%?
- Answer: 595

Size

- Question: You have a population of 750,000. How many people do you need to sample to get a margin of error of +/- 4%?
- Answer: 600

Size

- Question: You have a population of 3,000,000. How many people do you need to sample to get a margin of error of +/- 4%?
- Answer: 600

Size

- Question: To go from a margin of error of +/- 5% for 3,000,000 to a margin of error of +/- 4%, you need to go from a sample of _____ to 600?
- Answer: 384

Size

- Question: To go from a margin of error of +/- 4% for 3,000,000 to a margin of error of +/- 3%, you need to go from a sample of 600 to a sample of ____
- Answer: 1067

Size

- Question: To go from a margin of error of +/- 3% for 3,000,000 to a margin of error of +/- 2%, you need to go from a sample of 1067 to a sample of ____
- Answer: 2399

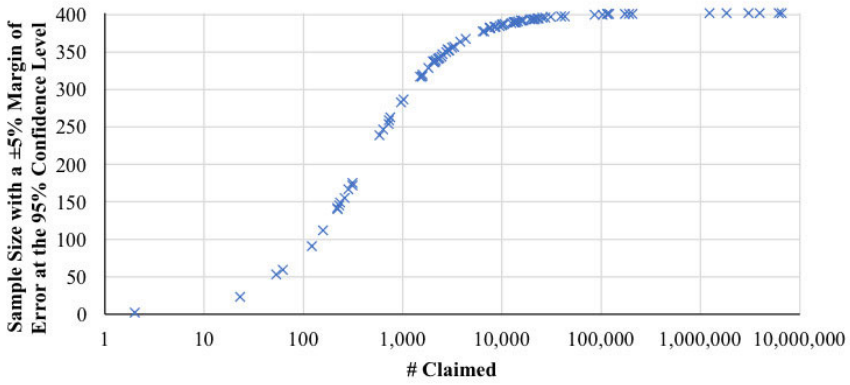
Why

$$MOE = Z\ value * \sqrt{\frac{p(1-p)}{n}}$$

This is the Confidence Interval Formula for a sample proportion.

And there's no reason to know this…. yet

Why?

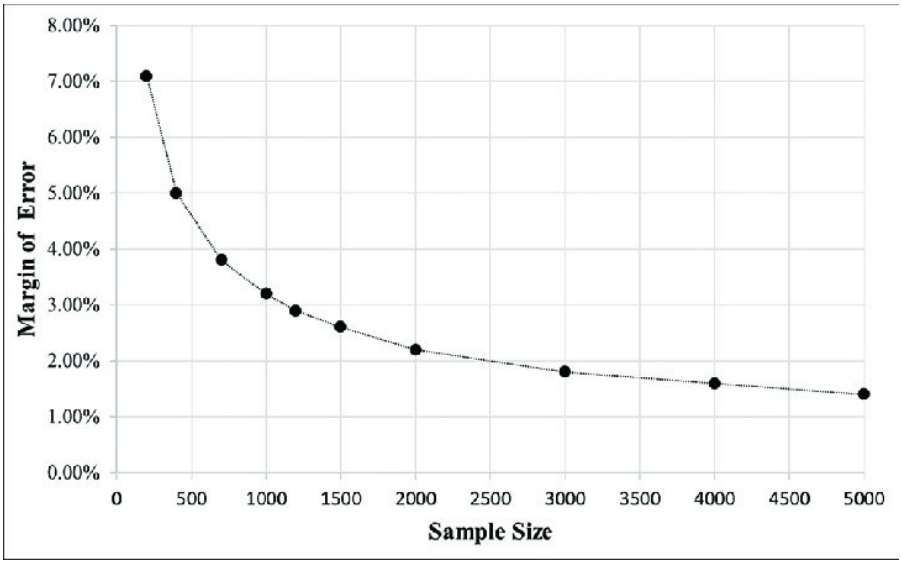


But

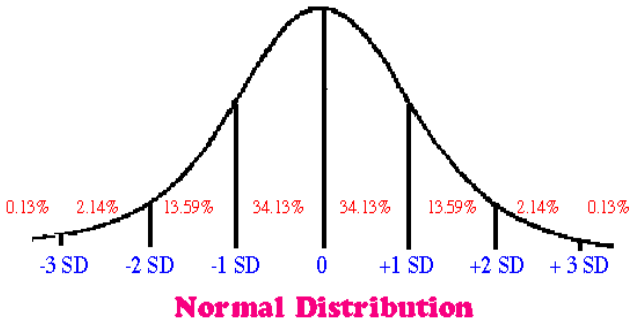
$$\frac{\frac{z^2 \times p\,(1-p)}{e^2}}{1+\left(\frac{z^2 \times p\,(1-p)}{e^2 N}\right)}$$

This is the formula for calculating required sample size for a given margin of error

Why?



Why: Normal Distribution



Sample

- When you go to the physician's office for blood tests, does she or he take all of your blood?
- (if yes, DIE, ZOMBIE, DIE!)

Sample

- No! He or she only draws a few drops because every element that appears in all of your blood is accurately (and proportionately) represented in each drop.
- Each drop perfectly *samples* your entire volume of blood.

Sample

- So…we know how LARGE the sample has to be….
- How do we go about making it as representative as a drop of blood?
- Answer: every element must have an equal probability of appearing in the sample.

Sample

- Key: random selection
- For a general population, this means Random Digit Dialed (RDD)
- For Landlines: RDD samples are generated using a database of "working blocks." A block (also known as a 100-bank or a bank) is a set of 100 contiguous numbers identified by the first two digits of the last four digits of a telephone number. For example, in the telephone number 203-567-7200, "72" is the block. A block is termed to be working if some specified number of listed telephone numbers are found in that block.
- RDD sample is distributed based on population density within working block

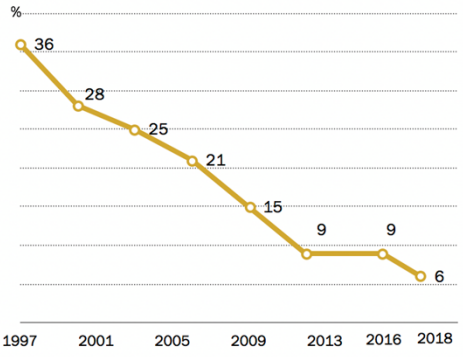
HOWEVER

- 61.9% of adults in South Carolina and 71.7% of all adults nationally are in wireless-only households and have NO land line. **Only 4% in South Carolina are reachable by land line only.**
- You MUST include wireless sample to get an accurate representation.
- IVR (Interactive Voice Response aka Automated polls aka Robopolls) are FORBIDDEN by law from calling wireless phones.

Plus, response rates are plummeting.

After brief plateau, telephone survey response rates have fallen again

Response rate by year (%)



Note: Response rate is AAPOR RR3. Only landlines sampled 1997-2006. Rates are typical for surveys conducted in each year.
Source: Pew Research Center telephone surveys conducted 1997-2018.

PEW RESEARCH CENTER

Non-Response Bias

- Are people who answer polls fundamentally different than those who do not?
- Thankfully, for the most part, no.
- No partisan or general demographic differences.
- Poll respondents are more likely to exhibit civic, and to a lesser extent, political engagement.

Mixed-Mode Polling

- In order to reduce non-coverage bias, a majority of polls have moved to mixed-mode surveys.
- These include an online sample in addition to phone sample.

Likely Voters

- Here is where the “pros” do things different.
- For likely voters, you need to start with RBS – Registration Based Sample.
- Now, your party leaders can tell you, you can buy the voter lists from the SC Election Commission.
- These make great walking lists, but they do NOT have phone numbers.

Likely Voters

- So to get numbers, you have to purchase them from a company that has matched names and addressed to phone numbers.
- Here’s the problem: you can only get “high confidence” matches for about 65% of registered voters.
- Many pros, and some ‘academic’ polls, stop there and only call RBS

Likely Voters

- This problem is not apparent in a state where one party is frequently winning by 10 points, but those polls ARE inaccurate.
- RBS must be “topped” with RDD and wireless sample.
- Additionally, likely voter screening is required.

Additional Methodology

- Calling times
- Within household randomization
- Weighting

	Census	Survey	
Male	48.4	50.4	0.960317
Female	51.6	49.6	1.040323

Can we trust polling?

- Short answer: yes
- Long answer: yeeeeeeeesssss*

- Is polling reliable?
- Short answer: yes
- Long answer: yeeeeeeesssss …. Mostly*

Polling in 2016

- Conventional wisdom: polling was way off in predicting Hillary Clinton would win.
- Reality: national polls were accurate …. VERY accurate in how much more of the popular vote Clinton would get
- But! We don't elect presidents by popular votes and state level polling in a few key states was off by just enough to throw things.

Polling in 2016

- From the AAPOR Report: “Eight states with more than a third of the electoral votes needed to win the presidency had polls showing a lead of three points or less (Trende 2016). As Sean Trende noted, “The final RealClearPolitics Poll Averages in the battleground states had Clinton leading by the slimmest of margins in the Electoral College, 272-266.” The polls on average indicated that Trump was one state away from winning the election.”
- “In four of those battleground states, a subset, which included Florida and Pennsylvania, the average poll margin was less than a point – signaling that either candidate could win.”

Polling in 2016

- “Polls showed Hillary Clinton leading, if narrowly, in Pennsylvania, Michigan and Wisconsin, which had voted Democratic for president six elections running. Those leads fed predictions that the Democratic Blue Wall would hold. Come Election Day, however, Trump edged out victories in all three.”
- “About 13 percent of voters in Wisconsin, Florida and Pennsylvania decided on their presidential vote choice in the final week, according to the best available data. These voters broke for Trump by near 30 points in Wisconsin and by 17 points in Florida and Pennsylvania.”

Polling in 2020

- ".. the average topline support for Trump in the polls understated Trump's share in the certified vote by 3.3 percentage points and overstated Biden's share in the certified vote by 1.0 percentage point. When undecided voters are excluded from the base, the two-candidate support in the polls understated Trump's certified vote share by 1.4 percentage points and overstated Biden's vote share by 3.1 percentage points."
- ".. it is unknown if Republicans who responded to polls voted differently than those who did not respond. If the voters most supportive of Trump were least likely to participate in polls then the polling error may be explained as follows: Self-identified Republicans who choose to respond to polls are more likely to support Democrats and those who choose not to respond to polls are more likely to support Republicans. Even if the correct percentage of self-identified Republicans were polled, differences in the Republicans who did and did not respond could produce the observed polling error."

Polling in 2020

- "Weighting to a reasonable target for partisanship and past 2016 vote does not fully correct the polling error. Reweighting the polls to reproduce the 2020 outcome requires a much larger margin for Trump in 2016 than actually occurred among respondents who report voting in 2016. The larger 2016 margin for Trump among those who reported voting for Trump in 2016 could be caused by the following: an issue with the weighting targets, i.e., the implied vote share among 2016 voters who voted in 2020 was different from the 2016 actual outcome; or differences in opinion within groups that responded, e.g., the 2016 Trump supporters who responded to polls were more likely to vote for Democrats than those who did not. It is impossible to know which caused the larger 2016 margin."

Polling in 2020

- "It is possible that 2020 pre-election polls were not successful in correctly accounting for new voters who participated in the 2020 election. There were many new voters in 2020 and it is unclear whether the proportion of new voters in the polls matched the proportion of actual new voters. It is also unclear whether the new voters who responded to polls had similar opinions to those who did not respond. Given the relative proportion and self-reported voting behavior of these new voters in the data available to the Task Force, this group of voters pushed the overall polling margins in the Democratic direction. Error in polling this group could have produced the observed polling error."

Polling in 2022

- The "Red Wave" that wasn't ….
- The overwhelming majority of polls – legitimate, non-partisan polls – were correct in predicting the "Red Wave" would be smaller than anticipated.
- However, poll aggregators included a slew of polls from partisan polling outfits that were significantly off … in one direction.
- For example, one, which put out a significant number of polls close to the election, was off by an average of 7.5%

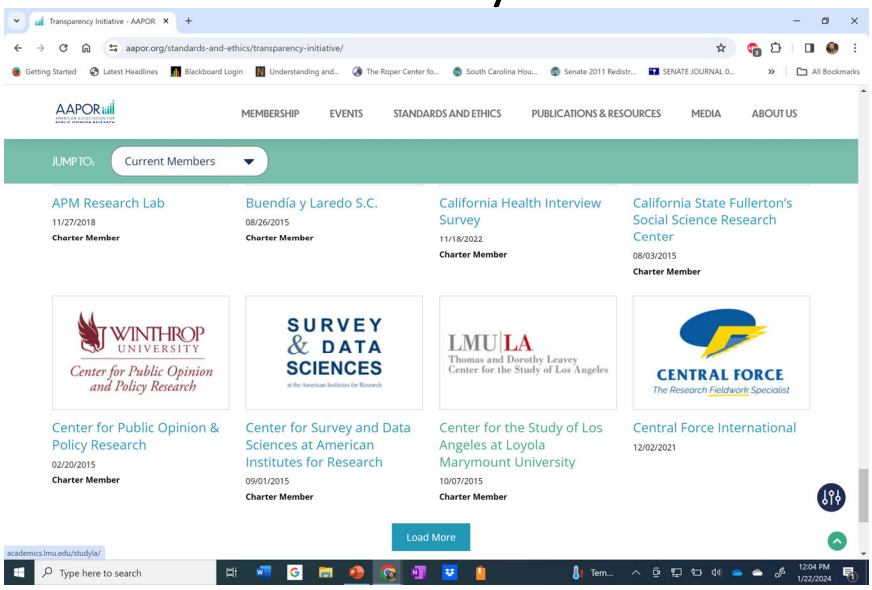
What polls to trust?

- Most polls sponsored or co-sponsored by a legitimate news source and/or an academic institution strive to use the best methodology and get it right.
  - Even when they get it wrong, the error is not intentional.

What polls to trust

- American Association of Public Opinion Transparency Initiative:
- Items for Immediate Disclosure
- 1. Data Collection Strategy
- 2. Who Sponsored the Research and Who Conducted It
- 3. Measurement Tools/Instruments
- 4. Population Under Study
- 5. Method Used to Generate and Recruit the Sample
- 6. Method(s) and Mode(s) of Data Collection
- 7. Dates of Data Collection
- 8. Sample Sizes (by sampling frame if more than one frame was used) and (if applicable) Discussion of the Precision of the Results
- 9.How the Data Were Weighted
- 10. How the Data Were Processed and Procedures to Ensure Data Quality
- 11. A General Statement Acknowledging Limitations of the Design and Data Collection

Well … looky there



Methodology and Funding
November 2023 Winthrop Poll

- **Quick Methodology**

- The November 2023 Winthrop Poll was a mixed-mode online and phone survey conducted and paid for by Winthrop University. The survey is a sample of adult residents of South Carolina who are registered to vote. The sample includes a probability sample of cell phones, a probability sample of land lines, a non-probability sample from an online panel curated by Dynata, and a non-probability sample of cell phones verified by Dynata. 65.7% of unweighted respondents came from the online panel. Among respondents reached by phone, 81.2% of unweighted respondents were contacted via cell phone. Data from 1655 (weighted) respondents were collected between November 4 and November 12, 2023. Results using all respondents have a margin of error of +/- 2.41% at the 95% confidence level (weighted data). All subgroups will have a higher margin of error. The margin of error for Republicans only is +/- 3.92% (n=626 weighted data) and the margin of error for Republicans and GOP leaning Independents is +/- 3.51% (n=780 weighted data).

AAPOR Transparency Initiative Statement

- The November 2023 Winthrop Poll was a conducted online and via telephone and paid for by Winthrop University (1,2). The online sample is a non-probability sample drawn from panels of adult residents of South Carolina curated by and purchased from Dynata (4,5). The sample includes a probability sample of cell phones, a probability sample of land lines, a non-probability sample from an online panel curated by Dynata, and a non-probability sample of cell phones verified by Dynata. 65.7% of unweighted respondents came from the online panel. Among respondents reached by phone, 81.2% of unweighted respondents were contacted via cell phone. (4,5). Data were weighted on sex, race and age based on data from the US Census Bureau (9). Full question wording for released results are included; full instrument available upon request (3). Participants responded using CAPI via computer or smartphone for the online sample and were hand-dialed by trained interviewers using a CATI system for the telephone sample; the survey averaged 5-8 minutes (6). Data were collected between November 4 and November 12, 2023. (7). Weighted data include a sample of 1655 respondents which translates to a margin of error of +/- 2.41% at the 95% confidence level for results using all respondents; all subgroups will have a higher margin of error (8). Center for Public Opinion & Policy Research personnel conducted quality control by eliminating data with inappropriate responses (didn't match the question asked), "line" data (strings of the same number punched in in a long row), or completions recorded too quickly for an individual to have actually read the questions for data collected online and through direct supervision and full monitoring for responses collected via telephone (10). All survey research contains unmeasured error and results should be seen as informative, not definitive (11).
- Numbers correspond to the AAPOR Transparency Initiative Disclosure Elements found HERE.

Thank you!

Questions?