

POLITICAL POLLING – PART II: PRACTICE OF POLLING

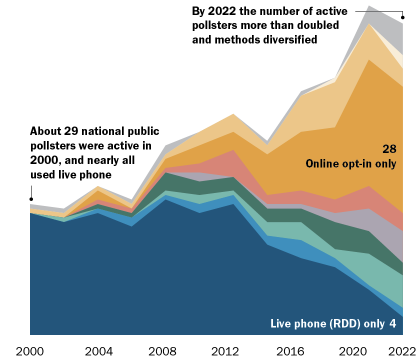
Data Analysis for Journalism and Political Communication
(Fall 2025)

Prof. Bell

TYPES OF SURVEY SAMPLES

As the number of public pollsters in the U.S. has grown, survey methods have become more diverse

Number of national pollsters using method(s)



Note: RDD refers to random-digit dial sampling. Refer to "How Public Polling Has Changed in the 21st Century" for a breakdown of other methods analyzed.

Source: Pew Research Center analysis of external data.

PEW RESEARCH CENTER

TYPES OF SURVEY SAMPLES

- **Systematic Sampling:** Select every n^{th} unit of a population (e.g. exit polls)

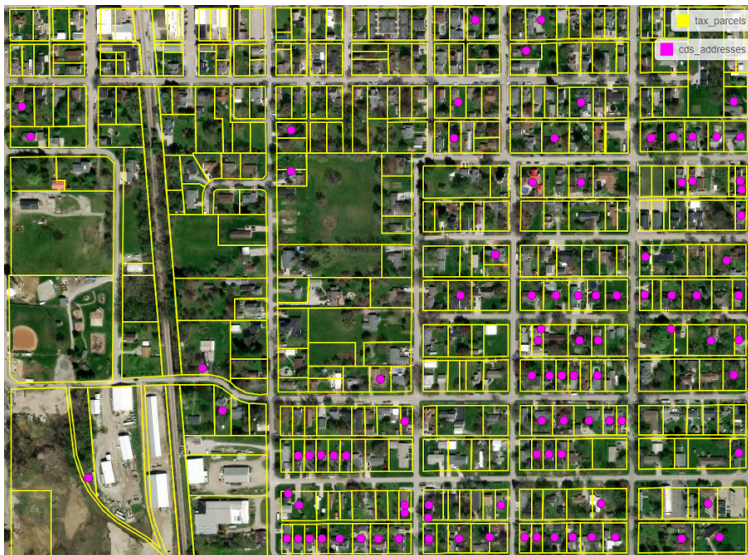
TYPES OF SURVEY SAMPLES

- **Systematic Sampling:** Select every n^{th} unit of a population (e.g. exit polls)
- **List-based Sampling:** Randomly select units from an existing list of the population (e.g. registered voter lists)

TYPES OF SURVEY SAMPLES

- **Systematic Sampling:** Select every n^{th} unit of a population (e.g. exit polls)
- **List-based Sampling:** Randomly select units from an existing list of the population (e.g. registered voter lists)
- **Address-based Sampling (ABS):** Randomly select households from a list of addresses provided by the U.S. Postal Service

TYPES OF SURVEY SAMPLES



TYPES OF SURVEY SAMPLES

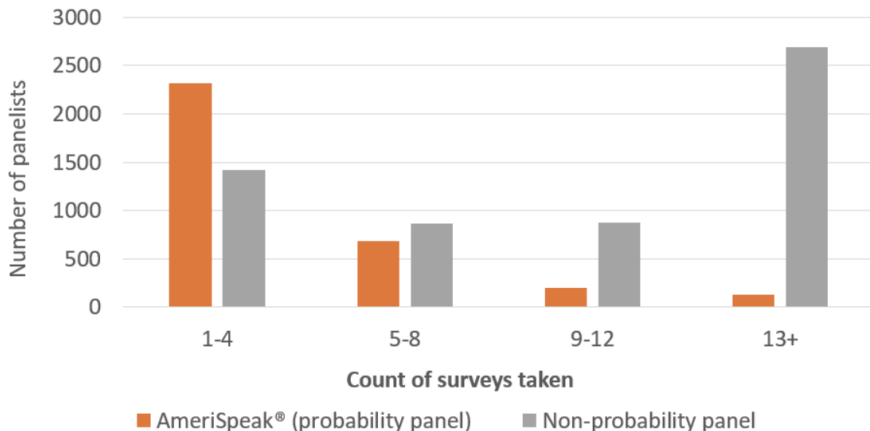
- **Systematic Sampling:** Select every n^{th} unit of a population (e.g. exit polls)
- **List-based Sampling:** Randomly select units from an existing list of the population (e.g. registered voter lists)
- **Address-based Sampling (ABS):** Randomly select households from a list of addresses provided by the U.S. Postal Service
- **Random-digit Dialing (RDD):** Randomly select area codes, and then random digits are added to the end to create 10-digit phone numbers

TYPES OF SURVEY SAMPLES

- **Systematic Sampling:** Select every n^{th} unit of a population (e.g. exit polls)
- **List-based Sampling:** Randomly select units from an existing list of the population (e.g. registered voter lists)
- **Address-based Sampling (ABS):** Randomly select households from a list of addresses provided by the U.S. Postal Service
- **Random-digit Dialing (RDD):** Randomly select area codes, and then random digits are added to the end to create 10-digit phone numbers
- **Non-probability/Quota Sampling:** Pseudo-randomly selecting, from an opt-in pool of respondents, a sample that approximates the make-up of the general population

TYPES OF SURVEY SAMPLES

Table 1. Surveys taken in past 30 days by panel type (self-report)



AI AND SURVEYS

- AI has been successfully applied in survey research, such as transforming open-ended responses into data or evaluating the quality of survey questions (e.g., AI agents can pre-test surveys)

AI AND SURVEYS

- AI has been successfully applied in survey research, such as transforming open-ended responses into data or evaluating the quality of survey questions (e.g., AI agents can pre-test surveys)
- A nascent area of research is the use of “synthetic personas” to imitate different types of survey respondents

AI AND SURVEYS

The screenshot displays a digital marketing analytics dashboard. On the left is a dark sidebar with navigation icons and labels: 'MyAgency' (with a dropdown and plus icon), 'Customer Persona', 'Website Persona' (selected, with a dropdown), 'Social Persona', and 'Competitor Persona'. Below these are icons for a list, people, settings, and a search icon. The main content area is titled 'Website Persona > All traffic > Segment 2' and includes an 'Export' button. It features three tabs: 'Persona' (selected), 'Distribution', and 'Journeys'. The 'Persona' tab shows a profile for 'Heather McCabe', a 33-year-old urban millennial from Las Vegas. A quote from her reads: 'I believe in building strong relationships, staying fit, and indulging in magical entertainment to brighten my days.' To the right is a map of Las Vegas, Nevada. Below the profile are sections for 'PROFILE INFORMATION' (summary, personal aspirations), 'PSYCHOLOGICAL DRIVERS' (goals), 'BUYING BEHAVIOR' (triggers), and 'WORK' (Travel Agent, Digital Lizard Idaho Publishing). A 'LIFESTYLE' section at the bottom shows icons for a married couple, a school, and a house, with text: 'Married', '1 grade-schooler', 'Has horses', and 'Rents house'. The top right of the dashboard shows 'Last updated: 28 Feb 2025'. The bottom of the image shows a navigation bar with icons for back, forward, and other controls.

MyAgency ▼ +

Customer Persona >

Website Persona ▼

Keith Mahoney
Time Buyer (High-value)
17%

Heather McCabe
Website Visitor
53%

Colin Scott
Segment Report
30%

Social Persona >

Competitor Persona >

Website Persona > All traffic > Segment 2

Persona Distribution Journeys

Last updated: 28 Feb 2025

Export ▼

Heather McCabe
33 yrs
Urban, Millennial
Website Visitor

"I believe in building strong relationships, staying fit, and indulging in magical entertainment to brighten my days."

PROFILE INFORMATION

Summary: Heather McCabe is a 33-year-old married woman residing in urban Las Vegas, Nevada. She rents a house, enjoys riding her bicycle, and has a passion for horses.

Personal aspirations: Heather aspires to create a strong sense of community, maintain a healthy and fit lifestyle, and foster deep relationships with her loved ones.

PSYCHOLOGICAL DRIVERS

Goals: Heather aims to explore new forms of entertainment like art, theater, and magic, indulge in reading children's literature, and continue pursuing her hobbies in books and biographies.

BUYING BEHAVIOR

Triggers: Heather is motivated by self-help and motivational content, community issues that impact her surroundings, and stories related to healthy living and fitness.

WORK

Travel Agent
Digital Lizard Idaho
Publishing

LIFESTYLE

Married
1 grade-schooler
Has horses

Rents house

Las Vegas, Nevada, United States

AI AND SURVEYS

- AI has been successfully applied in survey research, such as transforming open-ended responses into data or evaluating the quality of survey questions (e.g., AI agents can pre-test surveys)
- A nascent area of research is the use of “synthetic personas” to imitate different types of survey respondents
- With declining survey response rates making representative samples difficult, it may be a viable tradeoff

“If you’re going to pay for polling data that gets the wrong result, you might as well use AI and save money. While surveying real people seems to be getting less accurate over time, the question is whether AI polling will improve.”

- Reed Albergotti (Semafor)

AI AND SURVEYS

- AI has been successfully applied in survey research, such as transforming open-ended responses into data or evaluating the quality of survey questions (e.g., AI agents can pre-test surveys)
- A nascent area of research is the use of “synthetic personas” to imitate different types of survey respondents
- With declining survey response rates making representative samples difficult, it may be a viable tradeoff

“If you’re going to pay for polling data that gets the wrong result, you might as well use AI and save money. While surveying real people seems to be getting less accurate over time, the question is whether AI polling will improve.”

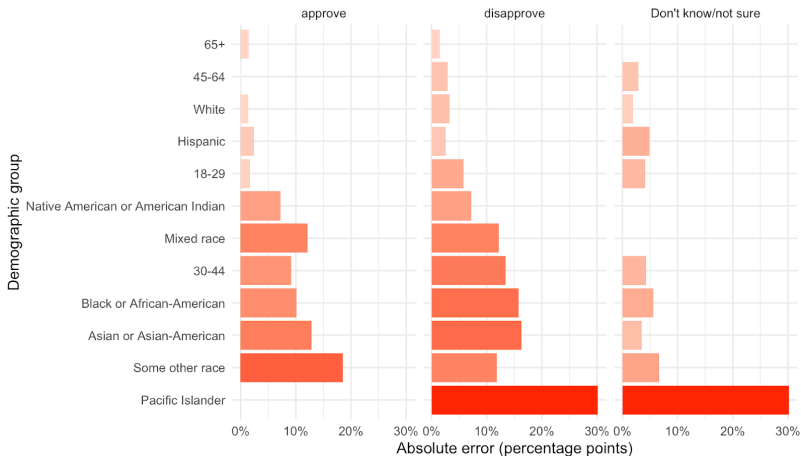
- Reed Albergotti (Semafor)

- However, AI models are still not good at “out-of-sample” inference, which means these synthetic personas may not generate good data

AI AND SURVEYS

LLMs produce large errors across important demographic groups

The absolute difference in the proportion of each group that says they ___ of Trump's job performance, minus the percent among each group that gets the same response from the LLM.



WRITING GOOD SURVEYS

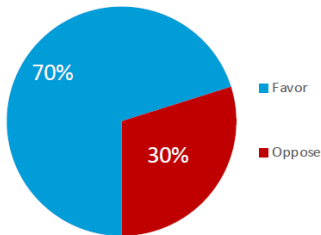
- As with data visualization, we have to assume that we have a limited amount of the respondent's attention

WRITING GOOD SURVEYS

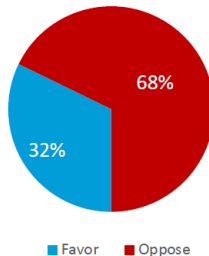
- As with data visualization, we have to assume that we have a limited amount of the respondent's attention
- The goal of survey design is to *minimize* cognitive load and *maximize* specificity, but these two goals are often in tension

WRITING GOOD SURVEYS

Do you favor or oppose a permanent ceasefire in Gaza?



Would you favor or oppose that ceasefire if it meant that Hamas was allowed to continue to hold hostages and Hamas were to continue to run Gaza?



Source: Harvard IOP Youth Poll

WRITING GOOD SURVEYS

Do you support or oppose a ban on each of the following:

	Strongly Support	Somewhat Support	Somewhat Oppose	Strongly Oppose
TikTok	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Diversity, Equity, and Inclusion (DEI) Initiatives	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Transgender collegiate athletes participating on sports teams different than their sex assigned at birth	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dynamic ticket pricing, where companies change ticket prices based on demand	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Daylight Saving Time	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Do you support or oppose a **ban** on each of the following:

	Strongly Support	Somewhat Support	Somewhat Oppose	Strongly Oppose
TikTok	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Diversity, Equity, and Inclusion (DEI) Initiatives	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Transgender collegiate athletes participating on sports teams different than their sex assigned at birth	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dynamic ticket pricing, where companies change ticket prices based on demand	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Daylight Saving Time	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

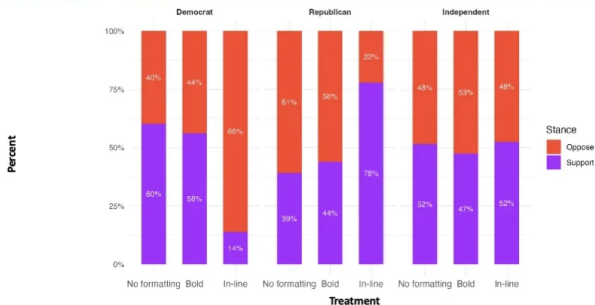
Do you support or oppose each of the following:

	Strongly Support	Somewhat Support	Somewhat Oppose	Strongly Oppose
A ban on TikTok	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A ban on Diversity, Equity, and Inclusion (DEI) Initiatives	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A ban on transgender collegiate athletes participating on sports teams different than their sex assigned at birth	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A ban on dynamic ticket pricing, where companies change ticket prices based on demand	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A ban on Daylight Saving Time	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

WRITING GOOD SURVEYS

YouGov

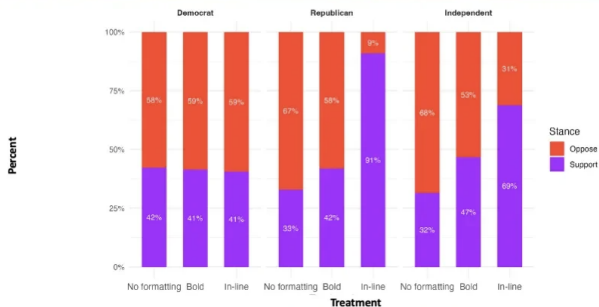
Ban DEI: Support vs oppose by political party and treatment



WRITING GOOD SURVEYS

YouGov

Transgender athletes in college sports: Support vs oppose by political party and treatment



WRITING GOOD SURVEYS

- As with data visualization, we have to assume that we have a limited amount of the respondent's attention
- The goal of survey design is to *minimize* cognitive load and *maximize* specificity, but these two goals are often in tension
- When the cognitive load on respondents is too high, they are likely to engage in **satisficing** or exit the survey entirely (known as survey attrition).

WRITING GOOD SURVEYS

Satisficing

Occurs when respondents do not expend the mental effort necessary to generate optimal answers to survey questions

WRITING GOOD SURVEYS

Satisficing

Occurs when respondents do not expend the mental effort necessary to generate optimal answers to survey questions

Jon Krosnick (1991) identifies several types of satisficing:

- 1 Picking the first or last answer

WRITING GOOD SURVEYS

Satisficing

Occurs when respondents do not expend the mental effort necessary to generate optimal answers to survey questions

Jon Krosnick (1991) identifies several types of satisficing:

- 1 Picking the first or last answer
- 2 Agreeing/acquiescing

WRITING GOOD SURVEYS

Satisficing

Occurs when respondents do not expend the mental effort necessary to generate optimal answers to survey questions

Jon Krosnick (1991) identifies several types of satisficing:

- 1 Picking the first or last answer
- 2 Agreeing/acquiescing
- 3 “Straight-lining”

WRITING GOOD SURVEYS

Satisficing

Occurs when respondents do not expend the mental effort necessary to generate optimal answers to survey questions

Jon Krosnick (1991) identifies several types of satisficing:

- 1 Picking the first or last answer
- 2 Agreeing/acquiescing
- 3 “Straight-lining”
- 4 Saying “don’t know”

WRITING GOOD SURVEYS

Satisficing

Occurs when respondents do not expend the mental effort necessary to generate optimal answers to survey questions

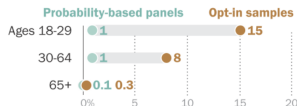
Jon Krosnick (1991) identifies several types of satisficing:

- 1 Picking the first or last answer
- 2 Agreeing/acquiescing
- 3 “Straight-lining”
- 4 Saying “don’t know”
- 5 Mental coin-flipping

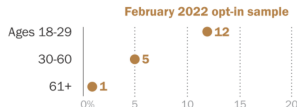
WRITING GOOD SURVEYS

Prior studies of 'bogus respondents' show large errors among young people

Average % of U.S. adults who answered "yes" to at least 10 of 16 yes/no questions



% of U.S. adults who say they are licensed to operate a class SSGN submarine



Source: Results for yes/no questions represent the mean values across three probability-based panels and across three online opt-in samples, originally reported in Pew Research Center's "Comparing Two Types of Online Survey Samples." Results for adults licensed to operate a class SSGN submarine are based on a Center survey experiment conducted Feb. 1, 2022, among an opt-in sample of U.S. adults.

"Online opt-in polls can produce misleading results, especially for young people and Hispanic adults"

PEW RESEARCH CENTER

HOW CAN WE AVOID SATISFICING?

Methods 101: Question Wording (Pew Research Center)

HOW CAN WE AVOID SATISFICING?

- 1 Keep question and survey length to a minimum

HOW CAN WE AVOID SATISFICING?

- 1 Keep question and survey length to a minimum
- 2 Put high-cognition questions earlier in the survey, and low-cognition questions (like demographics) at the end of the survey

HOW CAN WE AVOID SATISFICING?

- 1 Keep question and survey length to a minimum
- 2 Put high-cognition questions earlier in the survey, and low-cognition questions (like demographics) at the end of the survey
- 3 Use simple, unambiguous language

HOW CAN WE AVOID SATISFICING?

- 1 Keep question and survey length to a minimum
- 2 Put high-cognition questions earlier in the survey, and low-cognition questions (like demographics) at the end of the survey
- 3 Use simple, unambiguous language
- 4 Avoid leading questions or putting questions in an order that “primes” the respondent to think a particular way

HOW CAN WE AVOID SATISFICING?

- 1 Keep question and survey length to a minimum
- 2 Put high-cognition questions earlier in the survey, and low-cognition questions (like demographics) at the end of the survey
- 3 Use simple, unambiguous language
- 4 Avoid leading questions or putting questions in an order that “primes” the respondent to think a particular way
- 5 Avoid double-barreled questions (“To what extent do you agree with the Biden Administration’s plan to forgive \$20,000 of student loans for Pell Grant recipients and \$10,000 of student loans for most other borrowers?”)

HOW CAN WE AVOID SATISFICING?

- 1 Keep question and survey length to a minimum
- 2 Put high-cognition questions earlier in the survey, and low-cognition questions (like demographics) at the end of the survey
- 3 Use simple, unambiguous language
- 4 Avoid leading questions or putting questions in an order that “primes” the respondent to think a particular way
- 5 Avoid double-barreled questions (“To what extent do you agree with the Biden Administration’s plan to forgive \$20,000 of student loans for Pell Grant recipients and \$10,000 of student loans for most other borrowers?”)
- 6 Use a concise, mutually-exclusive, and complete set of response options (e.g., Likert scale)

HOW CAN WE AVOID SATISFICING?

- 1 Keep question and survey length to a minimum
- 2 Put high-cognition questions earlier in the survey, and low-cognition questions (like demographics) at the end of the survey
- 3 Use simple, unambiguous language
- 4 Avoid leading questions or putting questions in an order that “primes” the respondent to think a particular way
- 5 Avoid double-barreled questions (“To what extent do you agree with the Biden Administration’s plan to forgive \$20,000 of student loans for Pell Grant recipients and \$10,000 of student loans for most other borrowers?”)
- 6 Use a concise, mutually-exclusive, and complete set of response options (e.g., Likert scale)
- 7 Use open-ended questions judiciously

HOW CAN WE AVOID SATISFICING?

- 1 Keep question and survey length to a minimum
- 2 Put high-cognition questions earlier in the survey, and low-cognition questions (like demographics) at the end of the survey
- 3 Use simple, unambiguous language
- 4 Avoid leading questions or putting questions in an order that “primes” the respondent to think a particular way
- 5 Avoid double-barreled questions (“To what extent do you agree with the Biden Administration’s plan to forgive \$20,000 of student loans for Pell Grant recipients and \$10,000 of student loans for most other borrowers?”)
- 6 Use a concise, mutually-exclusive, and complete set of response options (e.g., Likert scale)
- 7 Use open-ended questions judiciously
- 8 Pre-test your survey

HOW CAN WE AVOID SATISFICING?




AJ Thurston, PhD
@AJThurston

...

So that later in my survey where I use a different measure that has 1 = low and 5 = high my survey will look even better by contrast, then, just to crush participants spirits, we ask them to enter their phone number for a follow-on survey using a slider input:

Please enter your phone number:



2158559745

FINDING POLLING DATA

Link to Roper iPoll

IN-CLASS EXERCISE

- 1 Search for a topic that you are interested in and select a question
- 2 Calculate the margin of error for the response options of interest (i.e., ignore “don’t know”). Are you confident that there is a difference between these proportions?
 - ▶ Recall the formula for margin of error from last week:
$$1.96 * \sqrt{(p * (1 - p)) / n}$$
- 3 Evaluate the question wording. Is this a well-written survey question? Why or why not?