Political Polling – Part II: Practice of Polling
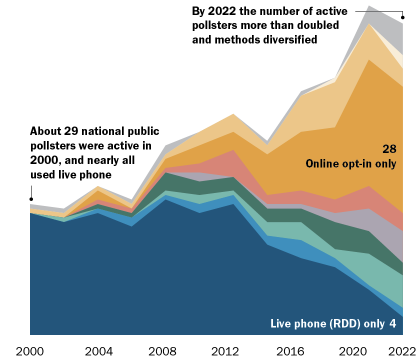
Data Analysis for Journalism and Political Communication (Fall 2025)

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)

By 2022 the number of active pollsters more than doubled and methods diversified

About 29 national public pollsters were active in 2000, and nearly all used live phone

28
Online opt-in only

Live phone (RDD) only 4

2000   2004   2008   2012   2016   2020   2022

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)
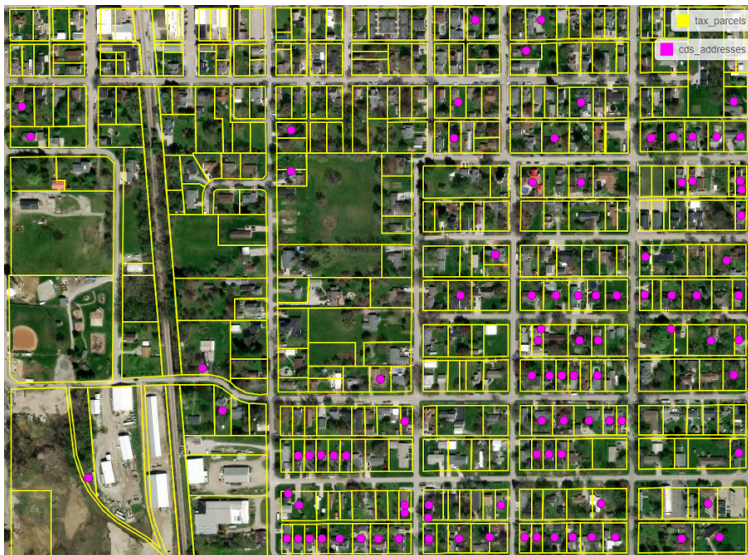
Types of Survey Samples

- **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)
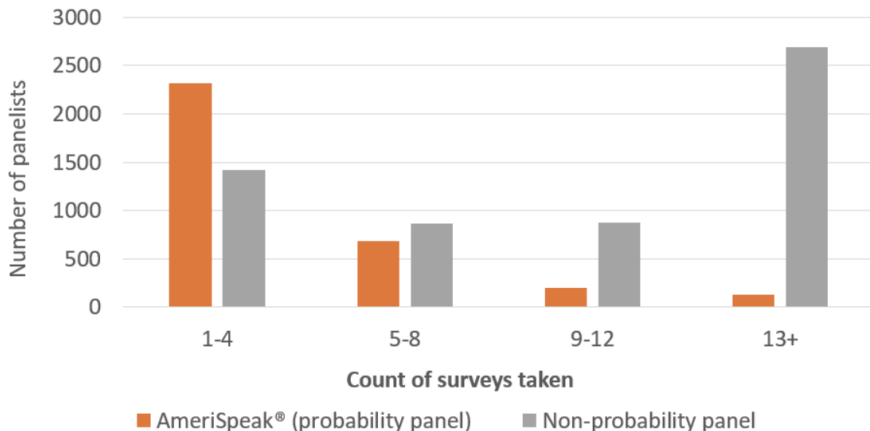
AI and Surveys

- 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

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 say they ___ of Trump's job performance, minus the percent among each group that gets the same response from the LLM.
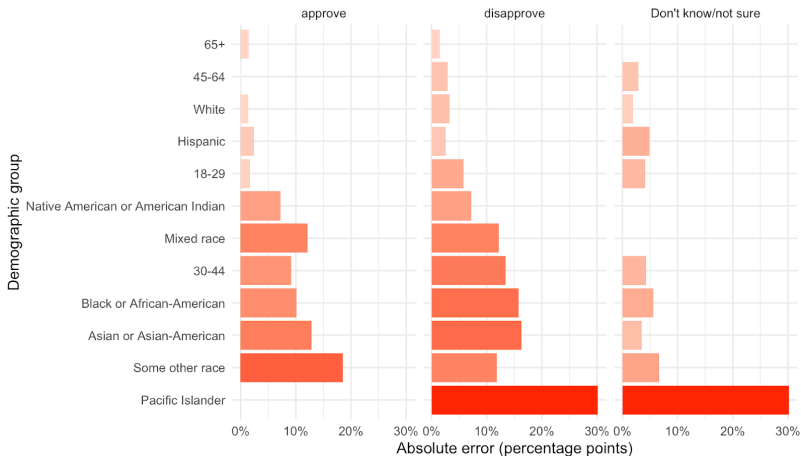
Writing Good Surveys

- 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

Source: Harvard IOP Youth Poll

Writing Good Surveys

Do you support or oppose a ban on each of the following:
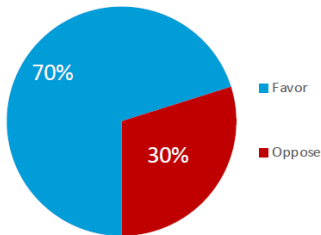
	Strongly Support	Somewhat Support	Somewhat Oppose	Strongly Oppose
TikTok	○	○	○	○
Diversity, Equity, and Inclusion (DEI) initiatives	○	○	○	○
Transgender collegiate athletes participating on sports teams different than their sex assigned at birth	○	○	○	○
Dynamic ticket pricing, where companies change ticket prices based on demand	○	○	○	○
Daylight Saving Time	○	○	○	○

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

	Strongly Support	Somewhat Support	Somewhat Oppose	Strongly Oppose
TikTok	○	○	○	○
Diversity, Equity, and Inclusion (DEI) initiatives	○	○	○	○
Transgender collegiate athletes participating on sports teams different than their sex assigned at birth	○	○	○	○
Dynamic ticket pricing, where companies change ticket prices based on demand	○	○	○	○
Daylight Saving Time	○	○	○	○

Do you support or oppose each of the following:

	Strongly Support	Somewhat Support	Somewhat Oppose	Strongly Oppose
A ban on TikTok	○	○	○	○
A ban on Diversity, Equity, and Inclusion (DEI) initiatives	○	○	○	○
A ban on transgender collegiate athletes participating on sports teams different than their sex assigned at birth	○	○	○	○
A ban on dynamic ticket pricing, where companies change ticket prices based on demand	○	○	○	○
A ban on Daylight Saving Time	○	○	○	○

Writing Good Surveys

Writing Good Surveys

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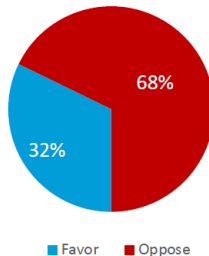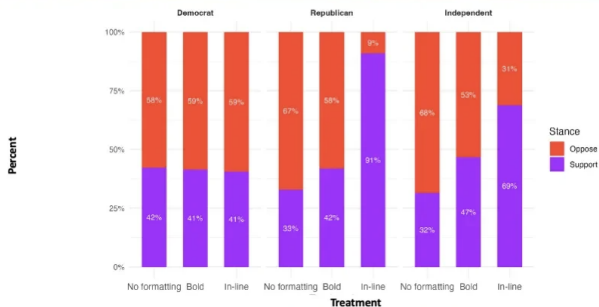
Writing Good Surveys

- 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

	Probability-based panels	Opt-in samples
Ages 18-29	1	15
30-64	1	8
65+	0.1	0.3
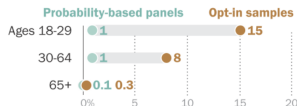
% of U.S. adults who say they are licensed to operate a class SSGN submarine

	February 2022 opt-in sample
Ages 18-29	12
30-60	5
61+	1
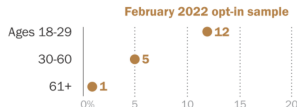
Source: Results for yes/no questions represent the mean values across three probability-based panels and across three online opt-in samples, originally reported in Pew Research Center's "Comparing Two Types of Online Survey Samples." Results for adults licensed to operate a class SSGN submarine are based 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?