Survey Research and Design

Measuring Sensitive and Implicit Attitudes

William Marble

November 16, 2022

Final Paper

Goal

- ▶ Pose a research question, contextualize why it's important, present evidence from your survey, interpret how that evidence helps answer the question

Logistics

- ▶ Your paper should be roughly 10 pages double spaced, including figures and tables but excluding citations
- ▶ Paper due by 11:59 on 12/20
- ▶ Last two days of class (12/5, 12/7): 7-minute presentations explaining briefly explaining your research and findings to the class (sign-ups open after class)
- ▶ I will give feedback on one rough draft if requested

Writing Advice

Content

- ▶ Part lab report: What hypotheses are you testing? What data and methods do you use to test the hypotheses? Do you find evidence in support or against the hypotheses? What are strength and limitations of the evidence?
- ▶ Part discussion of findings: Why is the research important? What does it say about society or politics more broadly? What follow-up questions does your research raise?

Writing Advice

Content

- ▶ Part lab report: What hypotheses are you testing? What data and methods do you use to test the hypotheses? Do you find evidence in support or against the hypotheses? What are strength and limitations of the evidence?
- ▶ Part discussion of findings: Why is the research important? What does it say about society or politics more broadly? What follow-up questions does your research raise?

Structure:

- ▶ **Introduction**: statement of research question, preview of methods and findings
- ▶ **Theory and Hypotheses**: discuss why the research question is important, what's already known about the topic, and what specific expectations or hypotheses you have (and why)
- ▶ **Data**: describe the data you use to answer the question
- ▶ **Methods** and analysis: describe the methods you use and how the analysis strategy maps on to the hypotheses
- ▶ **Results**: describe your findings
- ▶ **Discussion and Conclusion**: discuss the research, its strengths and limitations, follow-up questions it raises, and your broader conclusions about the findings

Deviations from this structure should be carefully thought out!

More Writing Advice

Write for an audience of your future self

- ▶ Someone who has taken this class but forgets some of it: explain the details!
- ▶ Write so that it's as easy as possible for your future self to remember what you did, why you did it, and what you found

More Writing Advice

Write for an audience of your future self

- ▶ Someone who has taken this class but forgets some of it: explain the details!
- ▶ Write so that it's as easy as possible for your future self to remember what you did, why you did it, and what you found

Write for clarity, not poetry

- ▶ Clearly state research question and a preview of the results in the introduction
- ▶ Don't keep your reader in suspense: tell 'em what you're gonna tell 'em, then tell 'em, then tell 'em what you told 'em ↝ need a clear thesis statement in the introduction (!!)
- ▶ Simple, direct language is better than flowery language (minimize the use of adjectives, avoid metaphors, write short sentences, etc.)
- ▶ Every paragraph should begin with a topic sentence for the paragraph
- ▶ Use sections and subsections liberally

More Writing Advice

Write for an audience of your future self

- ▶ Someone who has taken this class but forgets some of it: explain the details!
- ▶ Write so that it's as easy as possible for your future self to remember what you did, why you did it, and what you found

Write for clarity, not poetry

- ▶ Clearly state research question and a preview of the results in the introduction
- ▶ Don't keep your reader in suspense: tell 'em what you're gonna tell 'em, then tell 'em, then tell 'em what you told 'em $\rightsquigarrow$ need a clear thesis statement in the introduction (!!)
- ▶ Simple, direct language is better than flowery language (minimize the use of adjectives, avoid metaphors, write short sentences, etc.)
- ▶ Every paragraph should begin with a topic sentence for the paragraph
- ▶ Use sections and subsections liberally

Include figures and tables

- ▶ Should have at least one figure and one table in your paper
- ▶ Plots make it easy to see the findings
- ▶ Include a caption: should be able to understand the figure/table without reading anything else in the paper

Data Analysis

- ▶ Include survey weights
- ▶ Include as much analysis as needed to answer core question and clear follow-ups — no more, no less
- ▶ Describe the analysis clearly
- ▶ You are welcome to use supplemental data (e.g. other surveys, census data, data on policies)
- ▶ Submit all code for analysis along with paper

Measuring Sensitive Attitudes

Measuring Sensitive Attitudes

Often survey researchers may want to ask about sensitive topics:

- ▶ drug use, other illicit activities
- ▶ vote buying
- ▶ prejudice and discrimination toward outgroups
- ▶ attitudes toward authoritarian regimes

Measuring Sensitive Attitudes

Often survey researchers may want to ask about sensitive topics:

▶ drug use, other illicit activities

▶ vote buying

▶ prejudice and discrimination toward outgroups

▶ attitudes toward authoritarian regimes

These topics may elicit **social desirability bias**: a tendency to respond in a way that will be viewed favorably by others (including the interviewer).

▶ e.g., refusing to admit to illegal behavior

▶ obscuring taboo beliefs (e.g. racist attitudes)

Measuring Sensitive Attitudes

Often survey researchers may want to ask about sensitive topics:

- ▶ drug use, other illicit activities
- ▶ vote buying
- ▶ prejudice and discrimination toward outgroups
- ▶ attitudes toward authoritarian regimes

These topics may elicit **social desirability bias**: a tendency to respond in a way that will be viewed favorably by others (including the interviewer).

- ▶ e.g., refusing to admit to illegal behavior
- ▶ obscuring taboo beliefs (e.g. racist attitudes)

Upshot: directly asking about sensitive topics may lead to measurement error.

Example: Measuring Racial Prejudice

One way of measuring anti-Black prejudice is to directly ask. For example, Jackman (1978) asks:

Which of these statements would you agree with:

1. *White people have a right to keep Black people out of their neighborhoods if they want to.*
2. *Black people have a right to live wherever they can afford to, just like anybody else.*

Example: Measuring Racial Prejudice

One way of measuring anti-Black prejudice is to directly ask. For example, Jackman (1978) asks:

Which of these statements would you agree with:

1. *White people have a right to keep Black people out of their neighborhoods if they want to.*
2. *Black people have a right to live wherever they can afford to, just like anybody else.*

Even if people agree with these statements, they not want to acknowledge it to interviewers because they are socially unacceptable opinions.

Example: Measuring Racial Prejudice

One way of measuring anti-Black prejudice is to directly ask. For example, Jackman (1978) asks:

Which of these statements would you agree with:

1. *White people have a right to keep Black people out of their neighborhoods if they want to.*
2. *Black people have a right to live wherever they can afford to, just like anybody else.*

Even if people agree with these statements, they not want to acknowledge it to interviewers because they are socially unacceptable opinions.

Is there a less obtrusive way to measure racial prejudice?

List Experiments: Measuring Racial Prejudice

A popular method to measure sensitive attitudes or behaviors is known as a **list experiment** (aka the **item count** technique).

List Experiments: Measuring Racial Prejudice

A popular method to measure sensitive attitudes or behaviors is known as a **list experiment** (aka the **item count** technique).

Example from Kuklinski et al. (1997):

Now I'm going to read you three things that sometimes make people angry or upset. After I read all three, just tell me HOW MANY of them upset you. I don't want to know which ones, just HOW MANY.

1. *the federal government increasing the tax on gasoline*
2. *professional athletes getting million-dollar salaries*
3. *large corporations polluting the environment*
4. *a Black family moving in next door*

Half of the sample is only presented with the first three lines; half the sample is presented with all four.

List Experiments: Analysis

- ▶ Random assignment guarantees that the treatment group — which sees item 4 — is upset about the same number of items in 1-3, on average, as the control group — which is not asked about item 4
- ▶ Thus, if the treatment group reports a higher number of things that make them angry it's because of the fourth item
- ▶ Can estimate the proportion of the population who are upset about a Black family moving in next door by taking the difference in counts between the two groups

List Experiments: Logic

- ▶ List experiments help generate anonymity: we ask respondents *how many*, not *which*
- ▶ People can therefore plausibly deny that they hold the socially unacceptable view $\rightsquigarrow$ lower social desirability bias
- ▶ Person answer they're upset about 2 of those items might be upset about the non-sensitive items
- ▶ Can estimates *how many* people hold the sensitive opinion but not (necessarily) *which* people

Table 1. Mean Level of Anger Toward A Black Family Moving in Next Door, by Region (Whites Only)

Region	Experimental Condition		Estimated Percent Angry
	Baseline	Black Family	
Non-South	2.28[a]	2.24	0
	(.07)	(.05)	
	425[b]	461	
South	1.95	2.37	42
	(.06)	(.08)	
	139	136	

[a]Standard error of the estimate.
[b]Number of cases.

List Experiments: Pitfalls

Satisficing: respondents may skim over the items, ignoring whether the sensitive item is included/not included

► Mitigate by only including a few items, making it easy to read

List Experiments: Pitfalls

Satisficing: respondents may skim over the items, ignoring whether the sensitive item is included/not included

▶ Mitigate by only including a few items, making it easy to read

Lack of anonymity: suppose a respondent is upset about all 4 items or by 0 items. Then if they answer honestly we know their feeling about every item.

▶ Mitigate by careful choice of other options
▶ Could include an item no one is likely to be upset at
▶ Can also include items that are likely negatively correlated: people who get upset about higher taxes not as likely to be upset about pollution

List Experiments: Pitfalls

Satisficing: respondents may skim over the items, ignoring whether the sensitive item is included/not included

- ▶ Mitigate by only including a few items, making it easy to read

Lack of anonymity: suppose a respondent is upset about all 4 items or by 0 items. Then if they answer honestly we know their feeling about every item.

- ▶ Mitigate by careful choice of other options
- ▶ Could include an item no one is likely to be upset at
- ▶ Can also include items that are likely negatively correlated: people who get upset about higher taxes not as likely to be upset about pollution

Low power: need a larger sample size to generate as precise an estimate from a list experiment as from direct elicitation

- ▶ Only use a list experiment if necessary
- ▶ Include both a list experiment and direct elicitation to compare results

Explaining 2016 Polling Errors

Coppock (2017):

The "Shy Trump Supporter" hypothesis holds that polls understated support for Donald Trump because some respondents were reluctant to admit their support for his candidacy. This downward bias is hypothesized to be particularly pronounced on some phone surveys because they may have caused "shy" Trump supporters to dissemble to the live operators.

How would you design a study to test this hypothesis?

Shy Trump Voters: Coppock (2017) Study

Online survey fielded by Reuters/IPSOS, September 2-13, 2016.

Direct question: "If the 2016 presidential election were being held today and the candidates were as below, for whom would you vote?"

- ▶ 32.5% Donald Trump
- ▶ 37.0% Hillary Clinton
- ▶ 30.5% other/undecided/would not vote

Shy Trump Voters: Coppock (2017) Study

Online survey fielded by Reuters/IPSOS, September 2-13, 2016.

Direct question: "If the 2016 presidential election were being held today and the candidates were as below, for whom would you vote?"

- ▶ 32.5% Donald Trump
- ▶ 37.0% Hillary Clinton
- ▶ 30.5% other/undecided/would not vote

Shy Trump Voters: Coppock (2017) Study

List experiment: "Here is a list of [three/four] things that some people would do and some people would not. Please tell me HOW MANY of them you would do. We do not want to know which ones of these you would do, just how many. Here are the [three/four] things:"

- ▶ If it were up for a vote, I would vote to raise the minimum wage to 15 dollars an hour
- ▶ If it were up for a vote, I would vote to repeal the Affordable Care Act, also known as Obamacare
- ▶ If it were up for a vote, I would vote to ban assault weapons
- ▶ **If the 2016 presidential election were being held today and the candidates were Hillary Clinton (Democrat) and Donald Trump (Republican), I would vote for Donald Trump.**
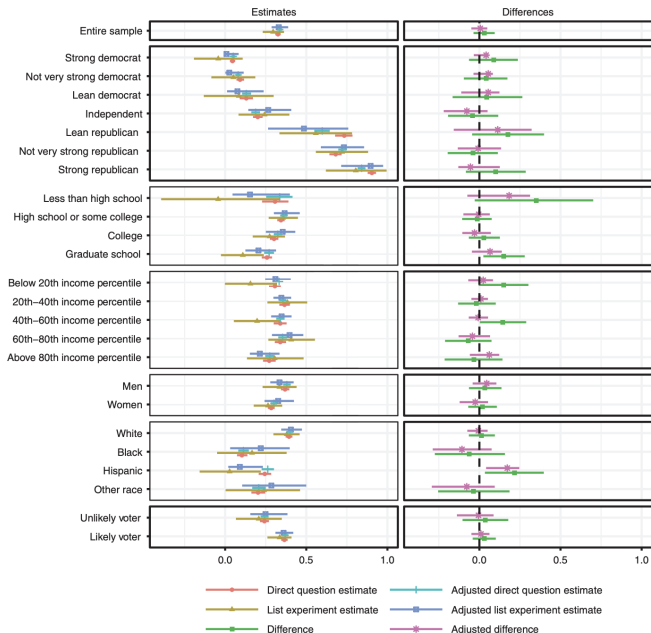
Shy Trump Voters: Results

	Control List	Treatment List
0 Items	0.11	0.11
1 Item	0.37	0.22
2 Items	0.40	0.46
3 Items	0.13	0.15
4 Items		0.06
Average	**1.548**	**1.843**
N	2,645	2,645

Treatment effect estimate: 1.843 - 1.548 = .295

$\rightsquigarrow$ Very similar to direct questioning!

Death Knell for the Shy Trump Voter Hypothesis?

Death Knell for the Shy Trump Voter Hypothesis?

Other Approaches to Avoiding Social Desirability Bias

Endorsement experiments

- ▶ "Randomly selected respondents are assigned to a treatment group and asked to express their opinion toward a policy endorsed by a specific actor whose support level we wish to measure . . . . These responses are then con- trasted with those from a control group of respondents that answered an identical question without the endorsement."
- ▶ Example application: measuring support for US forces and the Taliban in wartime Afghanistan (Lyall, Blair, and Imai, 2013)
- ▶ Logic?

Other Approaches to Avoiding Social Desirability Bias

Endorsement experiments

- ▶ "Randomly selected respondents are assigned to a treatment group and asked to express their opinion toward a policy endorsed by a specific actor whose support level we wish to measure . . . . These responses are then con- trasted with those from a control group of respondents that answered an identical question without the endorsement."
- ▶ Example application: measuring support for US forces and the Taliban in wartime Afghanistan (Lyall, Blair, and Imai, 2013)
- ▶ Logic?

Randomized response

- ▶ Respondent flips a coin. If heads, reply "yes" to the sensitive question. If tails, respond truthfully.
- ▶ Can calculate overall agreement with the sensitive question by subtracting out the noise
- ▶ Logic?

Other Approaches to Avoiding Social Desirability Bias

Crosswise model

- ▶ Ask respondents about two items at once, one sensitive and the other non-sensitive
- ▶ Then, unobtrusively ask the non-sensitive question in another way later in the survey
- ▶ Example: Jee and Zhang (2022): measuring regime criticism in China

  In your opinion, how many of the following statements are true?
  1. If I could choose the political regime of my country, the status quo authoritarian system would not be my ideal choice.
  2. I am currently between 25 and 30 years old (inclusive).
  You do not need to answer which statement is true, please select A or B below:
  A. Both statements are true OR neither statement is true
  B. Only one of the two statements is true

Later in the survey, ask birth year.

Indirect Questioning: Takeaways

▶ People may have reason to obscure their true attitudes from survey researchers

▶ These techniques all rely on *plausible deniability*: difficult/impossible to identify any given individual's responses

▶ But enable researchers to estimate *aggregate* opinions

Measuring Implicit Attitudes

Implicit Association Tests

Spend 10 minutes to take one of the implicit association tests on this website:

https://implicit.harvard.edu/implicit/takeatest.html

(just google "implicit association test")

Implicit and Explicit Cognition

Are you aware of all the thoughts, attitudes, and associations that influence your view of yourself? Of the world around you?

Implicit and Explicit Cognition

Are you aware of all the thoughts, attitudes, and associations that influence your view of yourself? Of the world around you?

Psychologists suggest that some information is not explicitly available in people's consciousness. Some information is available upon introspection, but other information is **implicit**.

Implicit and Explicit Cognition

Are you aware of all the thoughts, attitudes, and associations that influence your view of yourself? Of the world around you?

Psychologists suggest that some information is not explicitly available in people's consciousness. Some information is available upon introspection, but other information is **implicit**.

Memories, experience, knowledge, perceptions, and attitudes may all influence behavior, even if people aren't **explicitly** aware of it.

Implicit Cognition

Nosek, Greenwald, and Banaji (2007):

> [Researchers argued that] implicit cognition could reveal associative information that people were either unwilling or unable to report. In other words, implicit cognition could reveal traces of past experience that people might explicitly reject because it conflicts with values or beliefs, or might avoid revealing because the expression could have negative social consequences. Even more likely, implicit cognition can reveal information that is not available to introspective access even if people were motivated to retrieve and express it.

Implicit Cognition

Nosek, Greenwald, and Banaji (2007):

> [Researchers argued that] implicit cognition could reveal associative
> information that people were either unwilling or unable to report. In
> other words, implicit cognition could reveal traces of past experience
> that people might explicitly reject because it conflicts with values or
> beliefs, or might avoid revealing because the expression could have
> negative social consequences. Even more likely, implicit cognition can
> reveal information that is not available to introspective access even if
> people were motivated to retrieve and express it.

Some prominent applications of implicit cognition: attitudes, stereotypes, and
self-esteem

Implicit Cognition vs. Social Desirability Bias

▶ Implicit attitudes will be underreported when using explicit measures

▶ Social desirability bias also leads to mismeasurement of sensitive attitudes and behaviors

▶ Different reasoning, though: social desirability bias involves *intentionally* obscuring true attitudes, whereas implicit cognition is *unknown* even to the respondent

▶ Need new methods to measure implicit cognition

Enter the IAT

- ▶ The **implicit association test** was developed around 2000 to measure implicit associations
- ▶ Indirectly measures "strengths of associations among concepts" (Nosek et al.)
- ▶ Sorting examples from four concepts into just two response options
- ▶ Key intuition: easier to group together items that are strongly associated with each other

IAT: Setup

Four categories, usually contrasting sets of (**attributes** and **target concepts**)

- ▶ attributes: *good* and *bad*; *rich* and *poor*; *science* and *liberal arts*
- ▶ target concepts: *men* and *women*; *white* and *non-white*; *Democrat* and *Republican*

IAT: Setup

Four categories, usually contrasting sets of (**attributes** and **target concepts**)

- ▶ attributes: *good* and *bad*; *rich* and *poor*; *science* and *liberal arts*
- ▶ target concepts: *men* and *women*; *white* and *non-white*; *Democrat* and *Republican*

Respondents are shown examples and asked to place it into the corresponding category quickly

- ▶ attributes: "happy" and "sad"; "Scrooge" and "Tiny Tim"; "biology" and "history"
- ▶ target concepts: "Emily" and "George"; pictures of people of different races; "Obama" and "Trump"

Male

Female

Grandpa

If you make a mistake, a red **X** will appear. Press the other key to continue.

Liberal Arts

Science

History

If you make a mistake, a red **X** will appear. Press the other key to continue.

IAT: Core Task

- After some practice with categorizing single categories, researchers ask people to use all four categories at once
- One target category is paired with an attribute
- Respondents must place each item into the corresponding category — with only two response options

Male

or

Liberal Arts

Female

or

Science

Physics

If you make a mistake, a red **X** will appear. Press the other key to continue.

Male

or

Liberal Arts

Female

or

Science

Man

If you make a mistake, a red X will appear. Press the other key to continue.

IAT: Switching the Pairings

- After completing several categorization tasks, the pairings are flipped
- E.g., instead of pairing "female - science" and "male - liberal arts", group categories as "male - science" and "female - liberal arts"
- Then, repeat the same categorization task again

Male

or

Science

Female

or

Liberal Arts

Music

If you make a mistake, a red X will appear. Press the other key to continue.

Male

or

Science

Female

or

Liberal Arts

Boy

If you make a mistake, a red X will appear. Press the other key to continue.

Male

or

Science

Female

or

Liberal Arts

Astronomy

If you make a mistake, a red **X** will appear. Press the other key to continue.

IAT: Analysis

- Implicit association between concepts ("female" and "science", for example) is measured by differences in response time

- How long does it take to categorize "women - science" relative to "women - liberal arts"? And how does that difference compare to the difference between "men - science" and "men - liberal arts"? If first difference is larger than the second, suggests the association between "women - science" is strong

- Often-used measure is known as the "D-score" — controls for differences in average response time across respondents

IAT: Advantages

- ▶ Indirect measure that doesn't require *explicit* associations in the mind of the respondent
- ▶ Hard to fake responses, especially when respondents are unfamiliar with the IAT
- ▶ Relatively high test-retest reliability

IAT Critiques

Use of IAT to quantify *implicit bias* is pervasive. But what exactly is the IAT measuring?

IAT Critiques

Use of IAT to quantify *implicit bias* is pervasive. But what exactly is the IAT measuring?

- ▶ Taken literally, just measuring cognitive *associations*
- ▶ Do these associations necessarily translate to behavior?
- ▶ Is bias the same as discrimination?

Clayton, Horrillo, and Sniderman (2021) investigate correlation between IAT and explicit measures of racial prejudice

Table 1: Significant proportions of white respondents who openly express prejudice toward Black people are classified as free of implicit anti-Black prejudice.

	Non-prejudiced, IAT (panel)	Non-prejudiced, AMP (panel)	Non-prejudiced, AMP (time series)
Rate Black people lazier than white people	23%	24%	20%
Rate Black people less intelligent than white people	26%	28%	19%
Feel cooler toward Black people than white people	26%	28%	23%

Using the IAT

- IAT likely useful for measuring associations between concepts, and perhaps implicit stereotypes
- Implicit stereotypes and associations don't *necessarily* translate to explicit prejudice, but might point to underlying factors influencing society at large
- For explicit but socially undesirable attitudes, better to use other methods