

# **Survey Research and Design**

## Questionnaire Design

---

William Marble

September 21, 2023

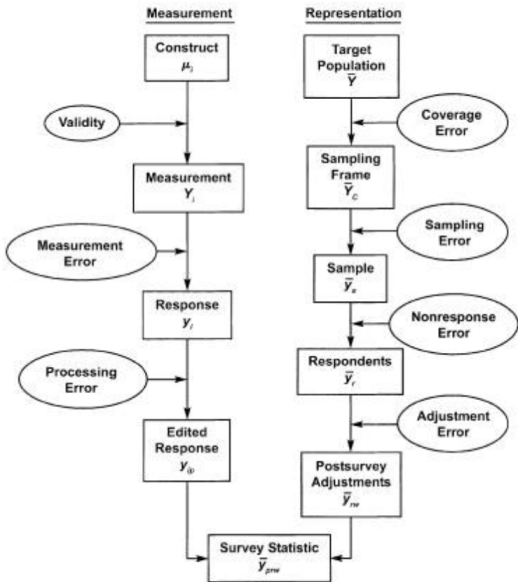


Figure 2.5 Survey life cycle from a quality perspective.

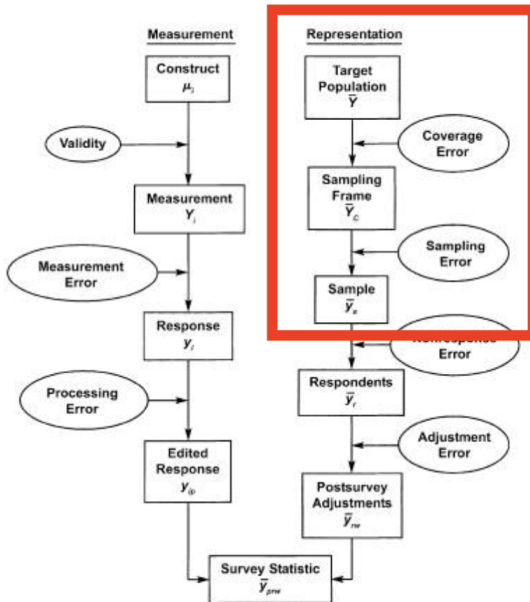


Figure 2.5 Survey life cycle from a quality perspective.

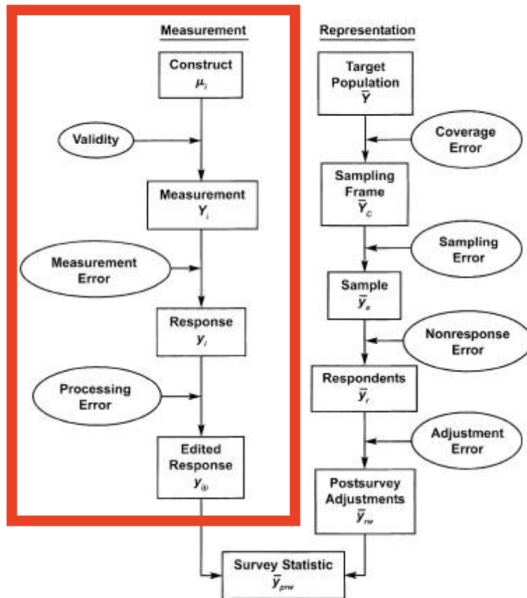


Figure 2.5 Survey life cycle from a quality perspective.

# Know What You're Trying to Measure

Important first step: know what the **goal** is. What **construct** are you trying to measure?

## Know What You're Trying to Measure

Important first step: know what the **goal** is. What **construct** are you trying to measure?

- ▶ Sometimes straightforward: income, who you voted for in the last election

# Know What You're Trying to Measure

Important first step: know what the **goal** is. What **construct** are you trying to measure?

- ▶ Sometimes straightforward: income, who you voted for in the last election
- ▶ Sometimes more difficult: ideology, personality traits, issue attitudes

# Know What You're Trying to Measure

Important first step: know what the **goal** is. What **construct** are you trying to measure?

- ▶ Sometimes straightforward: income, who you voted for in the last election
- ▶ Sometimes more difficult: ideology, personality traits, issue attitudes
- ▶ Without a clear idea of what you're trying to measure, it'll be hard to write good questions



# Overview of the Interview Process

- ▶ Think of a survey as a conversation between the interviewer and the respondent (even if online!)

# Overview of the Interview Process

- ▶ Think of a survey as a conversation between the interviewer and the respondent (even if online!)
- ▶ Trying to learn about the respondent while making minimal presuppositions

# Overview of the Interview Process

- ▶ Think of a survey as a conversation between the interviewer and the respondent (even if online!)
- ▶ Trying to learn about the respondent while making minimal presuppositions
- ▶ To keep respondent engaged, ask interesting questions that are not too complicated or intrusive

# Overview of the Interview Process

- ▶ Think of a survey as a conversation between the interviewer and the respondent (even if online!)
- ▶ Trying to learn about the respondent while making minimal presuppositions
- ▶ To keep respondent engaged, ask interesting questions that are not too complicated or intrusive
- ▶ Survey questions should be straightforward, easy to understand, and simple to answer

# Overview of the Interview Process

- ▶ Think of a survey as a conversation between the interviewer and the respondent (even if online!)
- ▶ Trying to learn about the respondent while making minimal presuppositions
- ▶ To keep respondent engaged, ask interesting questions that are not too complicated or intrusive
- ▶ Survey questions should be straightforward, easy to understand, and simple to answer
- ▶ Answer choices (if provided) should make sense in the context of the conversation

## **Framework for Thinking About Measurement**

---

## Measurement: A Framework

- ▶ We have some underlying concept,  $\mu_i$ 
  - ▶ Unobservable but there is a “true value” for each respondent

## Measurement: A Framework

- ▶ We have some underlying concept,  $\mu_i$ 
  - ▶ Unobservable but there is a “true value” for each respondent
- ▶ A measure  $Y_i$ 
  - ▶ Unobservable, but there is a “true value” for each respondent
  - ▶ We hope it is the same or closely related to  $\mu$



## Measurement: A Framework

- ▶ We have some underlying concept,  $\mu_i$ 
  - ▶ Unobservable but there is a “true value” for each respondent
- ▶ A measure  $Y_i$ 
  - ▶ Unobservable, but there is a “true value” for each respondent
  - ▶ We hope it is the same or closely related to  $\mu$
- ▶ People report a response  $y_{it}$ 
  - ▶ Observable response to a survey question for respondent  $i$  at time  $t$
  - ▶ May or may not be the same as  $Y_i$

## Measurement: A Framework

- ▶ We have some underlying concept,  $\mu_i$ 
  - ▶ Unobservable but there is a “true value” for each respondent
- ▶ A measure  $Y_i$ 
  - ▶ Unobservable, but there is a “true value” for each respondent
  - ▶ We hope it is the same or closely related to  $\mu$
- ▶ People report a response  $y_{it}$ 
  - ▶ Observable response to a survey question for respondent  $i$  at time  $t$
  - ▶ May or may not be the same as  $Y_i$

We observe  $y_{it}$ . We can write the relationship to  $\mu_i$  as:

$$\overbrace{Y_i}^{\text{measure}} = \overbrace{\mu_i}^{\text{construct}} + \overbrace{\varepsilon_i}^{\text{error}}$$

# Measurement: A Framework

- ▶ We have some underlying concept,  $\mu_i$ 
  - ▶ Unobservable but there is a “true value” for each respondent
- ▶ A measure  $Y_i$ 
  - ▶ Unobservable, but there is a “true value” for each respondent
  - ▶ We hope it is the same or closely related to  $\mu$
- ▶ People report a response  $y_{it}$ 
  - ▶ Observable response to a survey question for respondent  $i$  at time  $t$
  - ▶ May or may not be the same as  $Y_i$

We observe  $y_{it}$ . We can write the relationship to  $\mu_i$  as:

$$\begin{array}{ccccc} \text{measure} & & \text{construct} & & \text{error} \\ \underbrace{Y_i} & = & \underbrace{\mu_i} & + & \underbrace{\varepsilon_i} \\ \\ \underbrace{y_{it}} & = & \underbrace{Y_i} & + & \underbrace{u_{it}} \\ \text{response} & & \text{measure} & & \text{measurement error} \end{array}$$

## Validity

---

## Principles: Validity

Validity is the extent to which the measure is related to the construct.

Validity is the extent to which the measure is related to the construct.

$$Y_i = \mu_i + \varepsilon_i$$

- ▶  $\varepsilon_i = 0$  is the best-case scenario

Validity is the extent to which the measure is related to the construct.

$$Y_i = \mu_i + \varepsilon_i$$

- ▶  $\varepsilon_i = 0$  is the best-case scenario
- ▶ If  $\text{Var}(\varepsilon_i)$  is low and  $E(\varepsilon_i) = 0$ , then  $Y_i \approx \mu_i$
- ▶ Usually, there's some distinction between the measure and the construct,  $\varepsilon_i \neq 0$

## Principles: Validity

Validity is the extent to which the measure is related to the construct.

$$Y_i = \mu_i + \varepsilon_i$$

- ▶  $\varepsilon_i = 0$  is the best-case scenario
- ▶ If  $\text{Var}(\varepsilon_i)$  is low and  $E(\varepsilon_i) = 0$ , then  $Y_i \approx \mu_i$
- ▶ Usually, there's some distinction between the measure and the construct,  $\varepsilon_i \neq 0$
- ▶ Validity can be thought of as the correlation between  $\mu_i$  and  $Y_i$ : high correlation  $\rightarrow$  high validity; low correlation  $\rightarrow$  low validity



Validity is the extent to which the measure is related to the construct.

$$Y_i = \mu_i + \varepsilon_i$$

- ▶  $\varepsilon_i = 0$  is the best-case scenario
- ▶ If  $\text{Var}(\varepsilon_i)$  is low and  $E(\varepsilon_i) = 0$ , then  $Y_i \approx \mu_i$
- ▶ Usually, there's some distinction between the measure and the construct,  $\varepsilon_i \neq 0$
- ▶ Validity can be thought of as the correlation between  $\mu_i$  and  $Y_i$ : high correlation  $\rightarrow$  high validity; low correlation  $\rightarrow$  low validity

Why might a measure have low validity ( $\varepsilon_i \neq 0$ )?

Why might a measure have low validity ( $\varepsilon_i \neq 0$ )?

- ▶ Multiple interpretations of the measure
- ▶ Measure taps into unrelated constructs
- ▶ Only measures a portion of the construct, but leaves out important components

# Face Validity

## Face validity

- ▶ Would your respondents think the question is trying to measure the construct?
- ▶ Informal: does it ~feel~ right?

# Face Validity

## Face validity

- ▶ Would your respondents think the question is trying to measure the construct?
- ▶ Informal: does it ~feel~ right?

## Measuring age:

- ▶ “What is your age?”
- ▶ “How many gray hairs do you have?”

# Face Validity

## Face validity

- ▶ Would your respondents think the question is trying to measure the construct?
- ▶ Informal: does it ~feel~ right?

## Measuring age:

- ▶ “What is your age?”
- ▶ “How many gray hairs do you have?”

## Measuring vote intention in the next election:

- ▶ “In elections for Congress, do you usually vote for Democrats, Republicans, or some other party?”
- ▶ “If the election for the U.S. House of Representatives were held today, would you vote for for the Democratic candidate, the Republican candidate, or some other party?”

## Content validity

- ▶ Would subject-matter experts think the measure covers all relevant dimensions of the construct?
- ▶ Is there something the measure is leaving out?

## Content validity

- ▶ Would subject-matter experts think the measure covers all relevant dimensions of the construct?
- ▶ Is there something the measure is leaving out?

## Measuring health:

- ▶ “Do you have any chronic diseases?”



## Content validity

- ▶ Would subject-matter experts think the measure covers all relevant dimensions of the construct?
- ▶ Is there something the measure is leaving out?

## Measuring health:

- ▶ “Do you have any chronic diseases?”

## Validity: Example from Educational Testing

Consider the math SAT. What is the construct and what is the measure? Is there face validity? Content validity?

# Convergent Validity

## Convergent validity

- ▶ Does the measure correlate highly with other measures of the same construct?
- ▶ Especially important for latent (fundamentally unobservable) traits

## Measuring racial prejudice

- ▶ Do measures of implicit bias correlate with measures of explicit bias?
- ▶ If people endorse negative stereotypes of outgroups, does that correlate with behavior?

from Clayton, Horillo, and Sniderman (2021)

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

	<b>Non-prejudiced, IAT (panel)</b>	<b>Non-prejudiced, AMP (panel)</b>	<b>Non-prejudiced, AMP (time series)</b>
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%

## Reliability

---

## Principles: Reliability

Reliability is the extent to which a single respondent's answers to a question would vary over repeated trials.

$$\underbrace{y_{it}}_{\text{response}} = \underbrace{Y_i}_{\text{measure}} + \underbrace{u_{it}}_{\text{measurement error}}$$

## Principles: Reliability

Reliability is the extent to which a single respondent's answers to a question would vary over repeated trials.

$$\underbrace{y_{it}}_{\text{response}} = \underbrace{Y_i}_{\text{measure}} + \underbrace{u_{it}}_{\text{measurement error}}$$

- ▶ How do respondents go from hearing a question to providing an answer?
- ▶ Even if they have a “true” measure,  $Y_i$ , they might misreport (for many reasons)

## Principles: Reliability

Reliability is the extent to which a single respondent's answers to a question would vary over repeated trials.

$$\underbrace{y_{it}}_{\text{response}} = \underbrace{Y_i}_{\text{measure}} + \underbrace{u_{it}}_{\text{measurement error}}$$

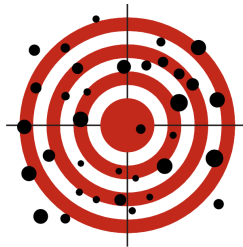
- ▶ How do respondents go from hearing a question to providing an answer?
- ▶ Even if they have a “true” measure,  $Y_i$ , they might misreport (for many reasons)
- ▶ If  $\text{Var}(u_{it}) \approx 0$  and  $E(u_{it}) \approx 0$ , then  $y_{it} \approx Y_i$  and we have high reliability
- ▶ Responses affected by context in the survey, things respondents happen to be thinking of, recent events, memory limitations, etc.



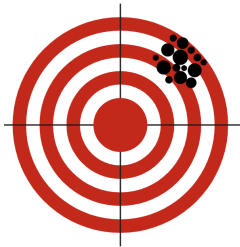
# Validity vs. Reliability



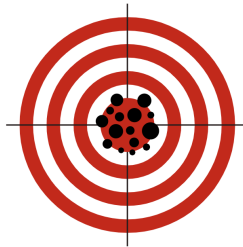
Unreliable & Unvalid



Unreliable, But Valid



Reliable, Not Valid



Both Reliable & Valid

## Measuring Reliability

- ▶ Most commonly: test-retest reliability
- ▶ Ask the same survey question to the same respondents at multiple points in time
- ▶ Correlation between answers is test-retest reliability
- ▶ Limitation: answers might genuinely change!

## **Cognitive Process of Survey-Taking**

---

# Survey Response Process

Comprehension of question



Retrieval of information



Estimation and judgment



Reporting an answer

## Comprehension of Question

- ▶ Can respondents easily understand what the question is asking?
- ▶ Are they familiar with all the words and terms used?
- ▶ Average American reads at 7th-8th grade level

## Example from Groves (NSDUH)

*Now think about the past 12 months, from [DATE] through today. We want to know how many days you've used any prescription tranquilizer that was not prescribed to you or that you only took for the experience or feeling it caused during the past 12 months.*

## Example from Groves (NSDUH)

*Now think about the past 12 months, from [DATE] through today. We want to know how many days you've used any prescription tranquilizer that was not prescribed to you or that you only took for the experience or feeling it caused during the past 12 months.*

What tasks does the respondent need to do to comprehend the question?

## Example from Groves (NSDUH)

*Now think about the past 12 months, from [DATE] through today. We want to know how many days you've used any prescription tranquilizer that was not prescribed to you or that you only took for the experience or feeling it caused during the past 12 months.*

What tasks does the respondent need to do to comprehend the question?

- ▶ Think about the time frame
- ▶ Assign meaning to key terms (“use,” “prescription tranquilizer”)
- ▶ Infer the purpose of the question
- ▶ Consider the meaning of the response options



## Example from Groves (NSDUH)

*Now think about the past 12 months, from [DATE] through today. We want to know how many days you've used any prescription tranquilizer that was not prescribed to you or that you only took for the experience or feeling it caused during the past 12 months.*

What information does the respondent need to retrieve?

## Example from Groves (NSDUH)

*Now think about the past 12 months, from [DATE] through today. We want to know how many days you've used any prescription tranquilizer that was not prescribed to you or that you only took for the experience or feeling it caused during the past 12 months.*

What information does the respondent need to retrieve?

- ▶ Have they ever used this type of drug?
- ▶ How often do they use it, on average?
- ▶ If remembering specific events, when did they occur?
- ▶ Some events/information are easier to remember than others
- ▶ Things that happened long ago are harder to remember

## Example from Groves (NSDUH)

*Now think about the past 12 months, from [DATE] through today. We want to know how many days you've used any prescription tranquilizer that was not prescribed to you or that you only took for the experience or feeling it caused during the past 12 months.*

How do respondents need to estimate and judge the information they retrieved?

- ▶ How to summarize information about their drug usage
- ▶ Probably don't keep a running tally → estimate based on response options
- ▶ Consider self-image or confidentiality concerns → do they want to report truthfully?

## **Survey Biases and What to Do About Them**

---

## Cognitive Biases

- ▶ Answering a survey question often involves different cognitive tasks
- ▶ These can be burdensome, leading to use of “shortcuts” that may induce biases

# Cognitive Biases

- ▶ Answering a survey question often involves different cognitive tasks
- ▶ These can be burdensome, leading to use of “shortcuts” that may induce biases
- ▶ General framework: **satisficing** versus **optimizing**

# Cognitive Biases

- ▶ Answering a survey question often involves different cognitive tasks
- ▶ These can be burdensome, leading to use of “shortcuts” that may induce biases
- ▶ General framework: **satisficing** versus **optimizing**
- ▶ Satisficers look for ways to provide a reasonable answer
  - ▶ Skim question, don't fully reflect on its meaning
  - ▶ Look through answers quickly
  - ▶ Pick the first answer option that seems reasonable
  - ▶ Often yields poor measurement

# Cognitive Biases

- ▶ Answering a survey question often involves different cognitive tasks
- ▶ These can be burdensome, leading to use of “shortcuts” that may induce biases
- ▶ General framework: **satisficing** versus **optimizing**
- ▶ Satisficers look for ways to provide a reasonable answer
  - ▶ Skim question, don't fully reflect on its meaning
  - ▶ Look through answers quickly
  - ▶ Pick the first answer option that seems reasonable
  - ▶ Often yields poor measurement
- ▶ Optimizers try to understand the question and provide “optimal” response
  - ▶ Read the question thoroughly
  - ▶ Carefully reflect on the information needed
  - ▶ Consider which response option best reflects their true answer



# Cognitive Biases

- ▶ Answering a survey question often involves different cognitive tasks
- ▶ These can be burdensome, leading to use of “shortcuts” that may induce biases
- ▶ General framework: **satisficing** versus **optimizing**
- ▶ Satisficers look for ways to provide a reasonable answer
  - ▶ Skim question, don't fully reflect on its meaning
  - ▶ Look through answers quickly
  - ▶ Pick the first answer option that seems reasonable
  - ▶ Often yields poor measurement
- ▶ Optimizers try to understand the question and provide “optimal” response
  - ▶ Read the question thoroughly
  - ▶ Carefully reflect on the information needed
  - ▶ Consider which response option best reflects their true answer

Our job as survey researchers is to design questionnaires that minimize satisficing and to minimize the bias satisficing causes.

## Order Effects: Primacy and Recency

Primacy bias and recency bias

- ▶ Primacy bias: when response options presented first are selected at a higher rate (often in written/online surveys)
- ▶ Recency bias: when response options presented last are selected at a higher rate (survey questions read aloud)
- ▶ Common in survey responses and even in elections!

# Order Effects: Primacy and Recency

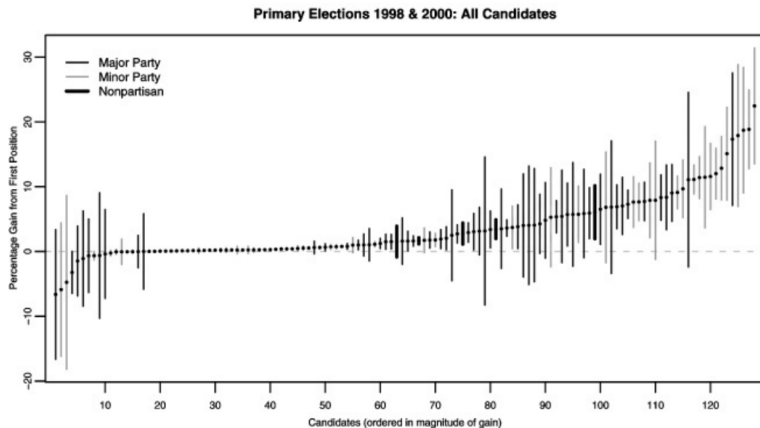
## Primacy bias and recency bias

- ▶ Primacy bias: when response options presented first are selected at a higher rate (often in written/online surveys)
- ▶ Recency bias: when response options presented last are selected at a higher rate (survey questions read aloud)
- ▶ Common in survey responses and even in elections!

## Solution: randomize response ordering

- ▶ Version 1: Do you agree or disagree with the following statement?  
Strongly agree, Agree, Neither agree nor disagree, Disagree, Strongly Disagree
- ▶ Version 2: Do you agree or disagree with the following statement?  
Strongly disagree, Disagree, Neither agree nor disagree, Agree, Strongly agree

# Ballot Order Effects



Source: Ho and Imai (2008)

## Acquiescence Bias

- ▶ Survey respondents sometimes say they “agree” with a statement, regardless of whether they actually do
- ▶ Potentially explained by personality traits (agreeableness or tendency to conform) or social desirability bias — assumption that interviewer endorses the statement

- ▶ Survey respondents sometimes say they “agree” with a statement, regardless of whether they actually do
- ▶ Potentially explained by personality traits (agreeableness or tendency to conform) or social desirability bias — assumption that interviewer endorses the statement

Solution: including construct-specific response options

## Acquiescence Bias: Example

ANES 2012: experiment to compare standard questions on political efficacy to revised questions with construct-specific response options.

## Acquiescence Bias: Example

ANES 2012: experiment to compare standard questions on political efficacy to revised questions with construct-specific response options.

### **Standard Version**

'Sometimes, politics and government seem so complicated that a person like me can't really understand what's going on.' Do you AGREE STRONGLY, AGREE SOMEWHAT, NEITHER AGREE NOR DISAGREE, DISAGREE SOMEWHAT, or DISAGREE STRONGLY with this statement?



## Acquiescence Bias: Example

ANES 2012: experiment to compare standard questions on political efficacy to revised questions with construct-specific response options.

### **Standard Version**

'Sometimes, politics and government seem so complicated that a person like me can't really understand what's going on.' Do you AGREE STRONGLY, AGREE SOMEWHAT, NEITHER AGREE NOR DISAGREE, DISAGREE SOMEWHAT, or DISAGREE STRONGLY with this statement?

**Revised Version** How often do politics and government seem so complicated that you can't really understand what's going on? [ALWAYS, MOST OF THE TIME, ABOUT HALF THE TIME, SOME OF THE TIME, or NEVER / NEVER, SOME OF THE TIME, ABOUT HALF THE TIME, MOST OF THE TIME, or ALWAYS]?

Problem: sometimes surveys ask for too much information at once

- ▶ Ask a manageable number of questions at a time
- ▶ Don't overload respondents with choices, questions, or information
- ▶ Leads to disengagement from the survey

Solution: keep it simple stupid!

- ▶ Don't use overly technical terms
- ▶ Ask one question at a time

# Don't Do This

Which of the following movies, TV shows, or entertainment brands have you ever heard of?

Please select all that apply.

Samurai Troopers (TV anime series)	Contra (Video game series)	Tokyo Ghoul (Animation, comics, and video game series)	DUNE (Film)
StarCraft (Video game series)	Harry Potter (Film series)	The Walking Dead (comic book series)	Attack on Titan (Manga series)
Pokemon (game series)	Star Wars (Film series)	Yakuza (Video game series)	The Witcher (Video game series)
Alien (Film series)	Pacific Rim (Film)	Silent Hill (Video game series)	James Bond (Film series)
Gundam (Animation, comics, and game series)	Resident Evil (Video game series)	Metal Gear (Video game series)	Dungeons & Dragons (Board game)
Dark Soul, Demon Soul, etc. (Video game series)	Castlevania (Animation, comics, and game series)	WWE (Sports Entertainment)	Star Trek
My Hero Academia (Animation, comics, and game series)	MARVEL (Comic book series)	Robotech (Anime series/franchise)	Macross (Anime series/franchise)
Nobunaga's Ambition (Video game series)	One Championship (sport)	Final Fantasy (Video game series)	None of these
Demon Slayer (Animation, comics, and game series)			

## Figure Out a Better Way to Do This

47. After the (2016) election and before the newly elected president and members of Congress are sworn into office, President (Barack) Obama is expected to submit the TPP agreement for a vote in Congress with the support of the Republican leaders of the House and Senate, Paul Ryan and Mitch McConnell. They say passing this agreement is important to our economy, our national security and our standing in the world. TPP opens the world's fastest growing markets to American exports, creating jobs here. And, it prevents China from setting the rules of the road. Failing to move ahead with this agreement would call into question America's leadership. But this agreement is opposed by most Democrats and many Republicans in Congress. They say TPP rigs the rules against American companies and workers. Democrats say TPP will tarnish President Obama's legacy, undermining the progress he made on the economy. TPP will lead to higher medicine prices and give corporations new powers over our lives, which is why Hillary Clinton opposes it too. Imagine that your member of Congress were to vote to pass the Trans-Pacific Partnership when Congress considers it after the election. Would you be more likely or less likely to vote for that member of Congress?

## **Other Advice on Writing Survey Questions**

---

## Avoiding Overly Technical or Unfamiliar Language

- ▶ Think of the conversation: wouldn't throw around acronyms or jargon that the other person doesn't understand
- ▶ Use common, easy-to-understand language
- ▶ Use conversational wording, even if grammar isn't technically "correct"

## Balanced Questions and Response Options

Want to ensure people can express both sides of an issue equally and that the question doesn't prime them one way or another.

Response categories:

- ▶ Approve vs. Disapprove
- ▶ Satisfied vs. Dissatisfied
- ▶ Gotten better vs. Stayed the same vs. Gotten worse
- ▶ Too much vs. About right vs. Too little

Question wording:

- ▶ Adding “or not” sometimes helps (e.g. Do you think the U.S. government should forgive federal student loan debt, or not?)

*I'm going to read you a list of personal characteristics or qualities. If 5 represents an absolutely essential quality in a president, and 1 a quality that is not too important, where on this scale of 5 to 1 would you rate... willingness to compromise?*



*I'm going to read you a list of personal characteristics or qualities. If 5 represents an absolutely essential quality in a president, and 1 a quality that is not too important, where on this scale of 5 to 1 would you rate... willingness to compromise?*

“Not too important” is not the same as “this is a negative quality.”

## Survey in Trump Email

1. How would you rate President Trump's first year in office (2017)?

- ☐ Great
- ☐ Good
- ☐ Okay
- ☐ Other

2. How would you rate President Obama's first year in office (2009)?

- ☐ Great
- ☐ Good
- ☐ Okay
- ☐ Poor

## Avoid Ambiguous and Vague Concepts

- ▶ Be specific when possible
- ▶ For example: Respondents might not know or remember if they are “civically engaged,” but they probably know if they voted or volunteered recently
- ▶ Problem: they may have a different idea of the construct (civic engagement) than you do
- ▶ How would you measure something like introversion?

## Avoid Double-Barrelled Questions

- ▶ A double-barrelled question is really two questions in one
- ▶ Example:  
*To reduce the federal deficit, do you support or oppose raising taxes and cutting spending?*
- ▶ Unclear how to answer this question
- ▶ Better: split it into two separate questions

## Avoid Presuppositions

- ▶ Don't assume things about the respondent that may not be true
- ▶ Example:  
*Last time you voted, did you vote in person on Election Day, by mailing your ballot, or in person before Election Day?*

Academic surveys:

- ▶ American National Election Studies
- ▶ Cooperative Election Survey
- ▶ General Social Survey

New surveys:

- ▶ Roper iPoll database
- ▶ Great resource

When in doubt: use questions that have been used before!