

Basic Survey Theory and Design

Constantine Boussalis

Harvard Law School

cboussalis@law.harvard.edu

Surveys: Not Just Asking People Questions

- The purpose of a **survey** is to produce accurate information that reflects the views and experiences of a given population.
- To achieve valid inferences from a survey, we need to reduce (as best we can) four distinct types of error:
 1. Coverage Error
 2. Sampling Error
 3. Non-response Error
 4. Measurement Error
- Theory and empirical evidence can help us tackle these sources of **bias** (misrepresentation of the population).

Some Key Terms & Concepts

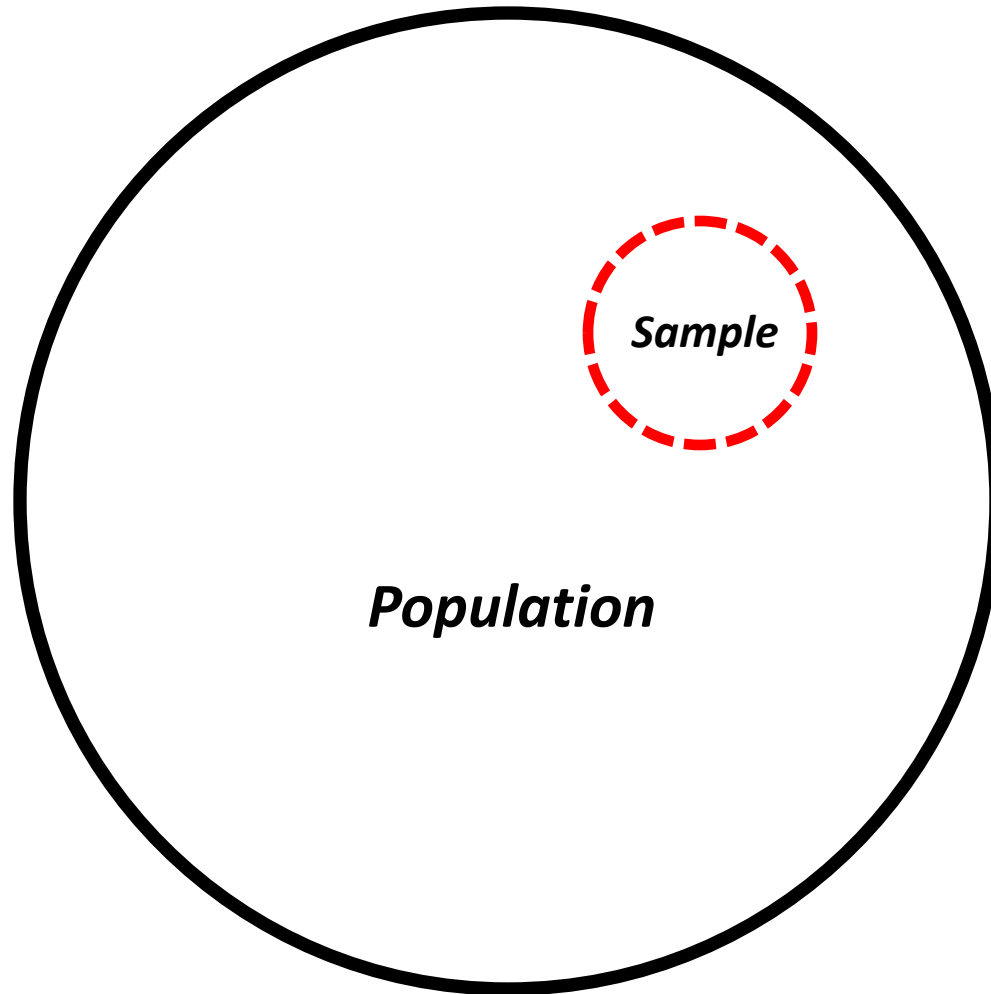
Population: All of the units (e.g. individuals, households, organizations) to which one desires to *generalize* survey results.

Sample: Units of the population that are drawn for inclusion in the survey.

Sample frame: The list from which a sample is to be drawn in order to represent the population.

Completed sample: The units that complete the questionnaire.

Generalizing to the Population



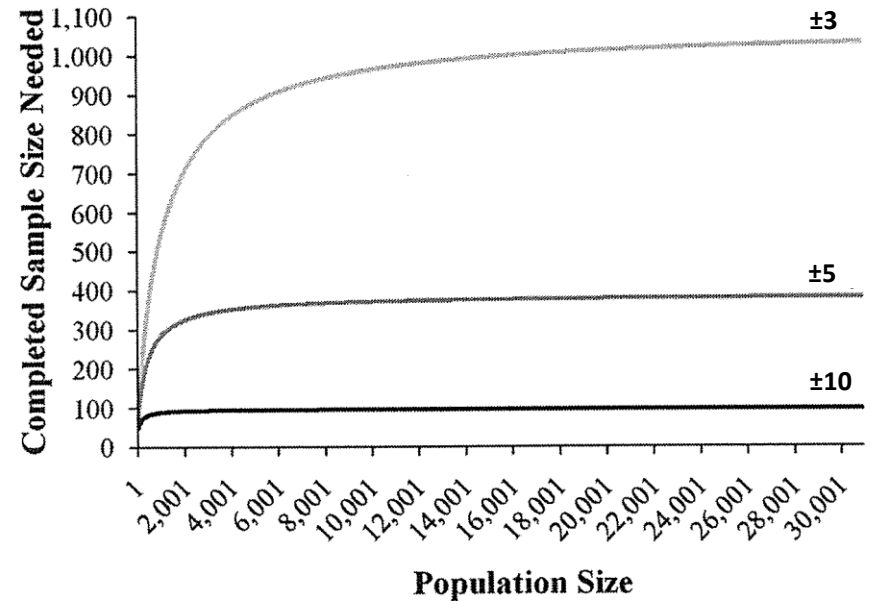
Coverage Error

- The results of a survey can be biased when:
 1. Not all members of the population have a known, non-zero chance of being included in the sample.
 2. The members of the population who were excluded are substantively different than the included members.
- Coverage error, therefore, stems from the **inability** of the survey to contact segments of the population.
- Strategies to avoid coverage error:
 1. Consider whether the survey mode (i.e. internet, telephone, mail, etc.) is applicable to all units.
 2. Be sure that units included in the sample frame are indeed members of the population of interest and not from another population
 3. Avoid unit duplication. This can lead to double-counting which, in turn, is a source of bias.

Sampling Error

- The number of *properly selected* completed responses determines the level of *precision* of our estimates (margin of error).
- As the sample size \rightarrow population size, the margin of error approaches zero
 - This is often too costly .
 - We do not *need* to sample the entire population.
- The size of the sample, **not** the proportion of the population sampled, affects precision.
- Diminishing returns on reduced margin of error as completed sample size increases.

Figure 3.2 Completed sample size needed by population size and desired margin of error (95% confidence level with 50/50 split).



Source: Dillman et al. (2009)

Non-Response Error

- Systematic error can also arise when not everyone selected for the survey completes the questionnaire.
- Crucially, bias occurs when the units who do not respond are *substantively different* from those who did complete the survey *on the item of interest*.
 - E.g. In a study estimating U.S. political opinion, if non-respondents are predominantly, say, Republican, the results of the study will be biased.
 - If the survey was estimating ice cream preferences, this heterogeneity might not be too problematic.
- Motivating (incentivizing) respondents to complete the survey is a good strategy for reducing non-response error.

Measurement Error

- Bias resulting from inaccurate answers to questions.
- How can a response be “inaccurate”?
 - Poor question wording and design
 - Improper questionnaire construction
- Some factors to consider when crafting a question:
 - Question format (e.g. open- vs. closed-ended, single vs. multiple answer, etc.)
 - Wording of the question stem

Question Formats: Open-ended

- Type answer in a blank box.
- Pros:
 - Allows respondent to freely answer question
 - Provides rich, detailed information
 - Appropriate for short, numerical responses (e.g. age)
- Cons:
 - Respondents more likely to skip open-ended questions
 - Coding answers into categories is resource intensive

Open-ended question

How would you describe the dean's leadership abilities?

How would you describe the dean's ability to innovate?

Question Formats: Closed-ended

- Respondents provide an answer after considering choices provided.
- **Nominal** scalar questions provide answers with no natural order.
 - E.g. “check-all-that-apply” and ranking questions .
 - Response difficulty increases as number of categories increases.
- **Ordinal** scalar questions provide an ordered set of answer categories.
 - Responses fall on some type of continuum.
 - E.g. intensity of satisfaction, frequency of behavior or events

Nominal, closed-ended

Which of these five statements best describes the dean?

- A strong leader
- A strong innovator
- Both a strong leader and innovator
- Neither a strong leader nor innovator

Ordinal, closed-ended for each concept

To what extent has the dean demonstrated strong leadership qualities?

- All of the time
- Most of the time
- Some of the time
- None of the time

To what extent has the dean demonstrated an ability to innovate?

- All of the time
- Most of the time
- Some of the time
- None of the time

Question Formats: Partially closed

- Hybrid of the open- and closed-ended formats.
- Includes set of responses and an “other” category.
- Pros:
 - Reduces number of items respondents have to consider simultaneously.
 - Allows for some freedom.
- Respondents more likely to answer provided answers than original answer.
 - We cannot accurately compare responses of listed to unlisted options

Partially closed question

Which of the following is your favorite college women’s sport?

- Basketball
- Gymnastics
- Soccer
- Softball
- Swimming
- Tennis
- Other: Please specify

Make every question require an answer



A question that does not require an answer from every respondent

If you made dinner at home last night, about how many minutes did it take to prepare the meal?

minutes



A revision that uses a filter question

Did you make dinner at home last night?

- Yes
- No



If yes, how many minutes did it take to prepare the meal?

minutes

Make sure the question is technically accurate



How many **feet** tall is your horse?

feet



How many **hands** tall is your horse?

hands

Ask one question at a time (double-barreled question)



A double-barreled question

Do you personally subscribe to, or regularly read, any periodicals, magazines, newsletters, etc. that are specifically related to your occupation?

- Yes
- No



A revision to ask each question separately and collect more info

Please indicate whether or not you personally subscribe to each of the following sources of information specifically related to your occupation.

- | | I do subscribe | I do not subscribe |
|-------------|--------------------------|--------------------------|
| Periodicals | <input type="checkbox"/> | <input type="checkbox"/> |
| Magazines | <input type="checkbox"/> | <input type="checkbox"/> |
| Newsletters | <input type="checkbox"/> | <input type="checkbox"/> |

Use simple and familiar words



1. Exhausted
2. Candid
3. Top priority
4. Leisure
5. Employment
6. Courageous
7. Rectify
8. "Occupants of this household"
9. "Your responses"
10. "Post-school extracurricular activities"
11. "Subnational region"



1. Tired
2. Honest
3. Most important
4. Free time
5. Work
6. Brave
7. Correct
8. "People who live here"
9. "Your answers"
10. "What you do after school"
11. "Area of the country"

Use specific and concrete words to specify concepts clearly



How many times did you eat together as a family last week?

of times



How many meals did you sit down to eat at home as a family last week?

of meals

Fewer words are better



1. Due to the fact that
2. At this point in time
3. A small number of
4. A considerable number of
5. Small in size
6. Has the ability
7. Ascertain the location of
8. Concerning the matter of
9. If conditions are such that
10. In the majority of instances
11. Make a decision
12. Take into consideration



1. Because
2. Now
3. A few
4. Many
5. Small
6. Can
7. Locate
8. About
9. If
10. Usually
11. Decide
12. Consider

Use complete sentences with simple sentence structures



Common use of incomplete sentences

Number of years lived in Massachusetts

Years

Your city or town

City or Town

Your county

County



A revision using complete sentences

How many years have you lived in Massachusetts?

Years

In what city or town do you live?

Name of City or Town

In what Massachusetts county do you live?

Name of Massachusetts County

Make sure “yes” means yes and “no” means no



A question containing a double negative

Do you favor or oppose not allowing the state to raise taxes without approval of 60% of the voters?

- Favor
- Oppose



A revision with no double negative

Do you favor or oppose requiring 60% approval by voters in order to raise taxes?

- Favor
- Oppose

Question must specify the response task



How many days did you exercise last week?

- All of the time
- Most of the time
- Some of the time
- None of the time



How many days did you exercise last week?

- Number of days (0-7)

State both positive and negative sides in the question stem



Do you favor congressional term limits of four years?

- Favor
- Oppose

To what extent do you agree with this statement "...”?

- Strongly agree
- Somewhat agree
- Somewhat disagree
- Strongly disagree



Do you favor or oppose congressional term limits of four years?

- Favor
- Oppose

To what extent do you agree or disagree with this statement "...”?

- Strongly agree
- Somewhat agree
- Somewhat disagree
- Strongly disagree

Answer categories must be mutually exclusive



What is your annual income?

- Less than \$25,000
- \$25,000 – \$50,000
- \$50,000 -- \$75,000
- More than \$75,000



What is your annual income?

- Less than \$25,000
- \$25,000 – \$49,999
- \$50,000 -- \$75,000
- More than \$75,000

Avoid check-all-that-apply questions



Which of the following varsity sports would you consider yourself to be a fan of? Please check all that apply

- Men's baseball
- Women's basketball
- Men's basketball
- Women's cross-country
- Men's cross-country
- Men's football
- Women's golf
- Men's golf
- Women's rowing
- Women's soccer
- Women's swimming
- Women's tennis
- Women's track and field
- Men's track and field
- Women's volleyball



Do you consider yourself to be a fan of each of the following varsity sports?

- | yes | no | |
|--------------------------|--------------------------|-------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | Men's baseball |
| <input type="checkbox"/> | <input type="checkbox"/> | Women's basketball |
| <input type="checkbox"/> | <input type="checkbox"/> | Men's basketball |
| <input type="checkbox"/> | <input type="checkbox"/> | Women's cross-country |
| <input type="checkbox"/> | <input type="checkbox"/> | Men's cross-country |
| <input type="checkbox"/> | <input type="checkbox"/> | Men's football |
| <input type="checkbox"/> | <input type="checkbox"/> | Women's golf |
| <input type="checkbox"/> | <input type="checkbox"/> | Men's golf |
| <input type="checkbox"/> | <input type="checkbox"/> | Women's rowing |
| <input type="checkbox"/> | <input type="checkbox"/> | Women's soccer |
| <input type="checkbox"/> | <input type="checkbox"/> | Women's swimming |
| <input type="checkbox"/> | <input type="checkbox"/> | Women's tennis |
| <input type="checkbox"/> | <input type="checkbox"/> | Women's track and field |
| <input type="checkbox"/> | <input type="checkbox"/> | Men's track and field |
| <input type="checkbox"/> | <input type="checkbox"/> | Women's volleyball |

State both positive and negative sides in the question stem



Do you favor congressional term limits of four years?

- Favor
- Oppose

To what extent do you agree with this statement “...”?

- Strongly agree
- Somewhat agree
- Somewhat disagree
- Strongly disagree



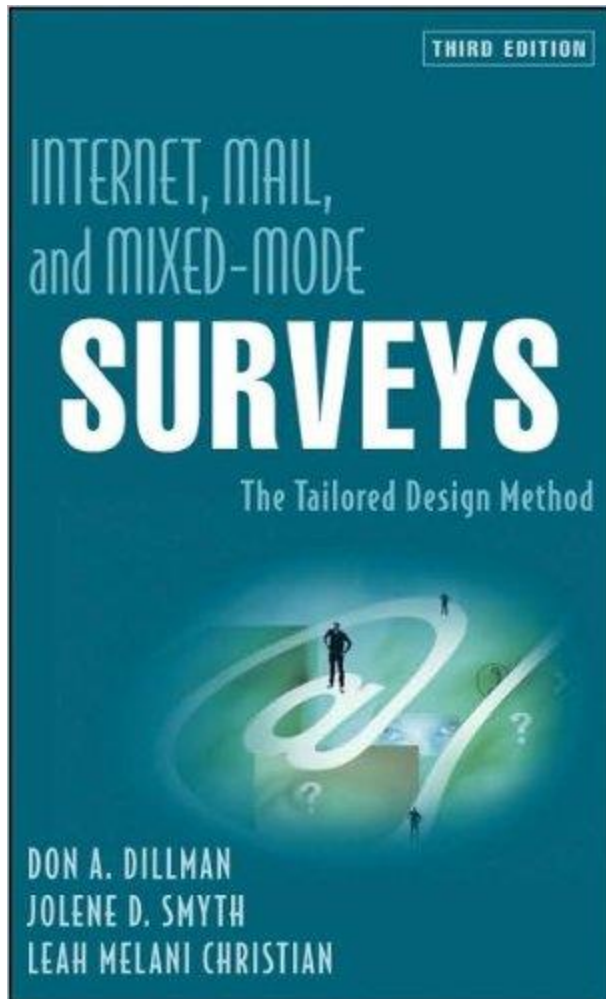
Do you favor or oppose congressional term limits of four years?

- Favor
- Oppose

To what extent do you agree or disagree with this statement “...”?

- Strongly agree
- Somewhat agree
- Somewhat disagree
- Strongly disagree

Recommended Reading



Dillman, D.A., Smyth, J.D., & Christian, L.M. (2009) *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method* (3rd Edition). Hoboken, N.J.: Wiley & Sons.

ISBN: 978-0-471-69868-5



harvard.qualtrics.com