Survey Experiments

SURVEY RESEARCH LABORATORY
UNIVERSITY OF ILLINOIS AT CHICAGO

Outline
1. Overview
2. What is a survey?
3. What is an experiment?
4. What are the advantages and disadvantages of experiments in surveys?
5. How have experiments been used in surveys?
6. Conclusions

What is a survey?
Goal of generalizing findings to a population of interest
External validity
Typically relies on self-report
Often involves large samples
Can cover a large geographic area (e.g., national telephone survey)

What is an experiment?
Manipulation of independent variables
- Within subjects – vignette design, pre-test/post-test design
- Between subjects
Often uses random assignment
- Between subjects manipulations or order of within subjects
- Random number generator, computer program; equality of groups
High internal validity
Campbell and Stanley – threats to internal validity and designs that address them
Often conducted in a laboratory
Often uses homogenous sample
Can use vivid manipulations and a broad array of measures – hot sauce measure!
Can and often does use deception (ability to debrief)

What are the advantages and disadvantages of experiments in surveys?

What are survey experiments?
Implement an experiment as part of a survey
Methodological or substantive question
Typical, but not required:
- Large sample size?
- Representative/probability sample – Mturk study
- Often not done in a laboratory setting
Limitations/typical design
- Manipulations often more subtle than some laboratory studies
- Limited array of measurements available (primarily self-report)
- Deception unusual

What is not a survey experiment?
Not a clear-cut line – matter of some debate
Goal of generalizing to population?
- Probability sampling methods
- Every unit in the sampling frame has a known, non-zero probability of selection
- Quota sampling
- Convenience sampling
- Nonprobability online panels
- Conveniencesampling
- Not a survey experiment?
Use survey to measure effect of intervention but the survey itself doesn’t include the intervention
- Matched design
- Match survey respondents with a particular characteristic or experience (e.g., breast cancer) with similar respondents (on demographic characteristics) who do not have the characteristic or have not had the experience
- Natural experiment
- E.g., event occurs midway through data collection
- Matched designs
Why are survey experiments valuable?

- Maximize internal and external validity
- Both generalizable and can draw causal inferences

How have experiments been used in surveys?

Methodological experiments
- Measurement – validity and/or reliability
- Question wording
- Question order
- Mode
- Nonresponse bias and/or cost effectiveness – survey design

Substantive experiments
- Vignette designs
- Can be used for both basic or applied research

Question wording: acquiescence

Do you agree or disagree with the following statement:

QW#1: Individuals are more to blame than social conditions for crime and lawlessness in this country.

QW#2: Social conditions are more to blame than individuals for crime and lawlessness in this country.

Has been manipulated both between and within subjects.

Between subjects: Individuals or social conditions? (Schuman and Presser, 1981)

<table>
<thead>
<tr>
<th>Agree</th>
<th>Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>70</td>
<td>30</td>
</tr>
</tbody>
</table>

Within subjects: Individuals or social conditions? (Schuman and Presser, 1981)

<table>
<thead>
<tr>
<th>Agree with both</th>
<th>Disagree with both</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>70</td>
</tr>
</tbody>
</table>
Question wording: Response Format
Number of options/midpoint

Would you rate Trump’s performance as President as...
- Very good,
- Somewhat good,
- Slightly good,
- Neither good nor bad,
- Slightly bad,
- Somewhat bad, or
- Very bad

Would you rate Trump’s performance as President as...
- Very good,
- Somewhat good,
- Slightly good,
- Neither good nor bad,
- Slightly bad,
- Somewhat bad, or
- Very bad

Would you rate Trump’s performance as President as...
- Good,
- Neither good nor bad, or
- Bad

Question Wording: No opinion responding

QW#1: Do you favor or oppose stricter gun control laws?
QW#2A: Do you favor or oppose stricter gun control laws, or do you not have an opinion?
QW#2B: Do you favor or oppose stricter gun control laws, or do you not have enough information to say?
QW#2C: Do you have an opinion about gun control laws?
[IF YES] Do you favor or oppose stricter gun control laws?

Question wording: Response Format
Response labeling

How would you rate your overall experience in this course?
- Very bad
- Somewhat bad
- Slightly bad
- Neither bad nor good
- Slightly good
- Somewhat good
- Very good

How would you rate your overall experience in this course?
- Very bad
- Somewhat bad
- Slightly bad
- Neither bad nor good
- Slightly good
- Somewhat good
- Very good

How would you rate your overall experience in this course?
- Very bad
- Somewhat bad
- Slightly bad
- Neither bad nor good
- Slightly good
- Somewhat good
- Very good

Question wording: Response Order

Which do you think is the more serious facing the country today: terrorism or the economy?
Which do you think is the more serious facing the country today: the economy or terrorism?

Estimate and control for question order effects

Question wording: Terminology

Do you favor or oppose the Affordable Care Act?
Do you favor or oppose Obamacare?

Another example: Welfare versus Assistance to the Poor
Question order:

Why does order matter:
- Priming
- Conversational norms
- Setting the scale for repeated questions

Question order: Priming: assimilation effect

How would you rate Trump’s tax policies? Would you say very good, somewhat good, slightly good, neither good nor bad, slightly bad, somewhat bad, or very bad?
How would you rate Trump’s performance overall? Would you say very good, somewhat good, slightly good, neither good nor bad, slightly bad, somewhat bad, or very bad?

Vs.

How would you rate Trump’s performance overall? Would you say very good, somewhat good, slightly good, neither good nor bad, slightly bad, somewhat bad, or very bad?
How would you rate Trump’s tax policies? Would you say very good, somewhat good, slightly good, neither good nor bad, slightly bad, somewhat bad, or very bad?

Question order: Conversational norms: contrast effect

How would you rate Trump’s tax policies? Would you say very good, somewhat good, slightly good, neither good nor bad, slightly bad, somewhat bad, or very bad?
How would you rate Trump’s performance overall? Would you say very good, somewhat good, slightly good, neither good nor bad, slightly bad, somewhat bad, or very bad?

Vs.

How would you rate Trump’s tax policies? Would you say very good, somewhat good, slightly good, neither good nor bad, slightly bad, somewhat bad, or very bad?
How would you rate Trump’s performance overall? Would you say very good, somewhat good, slightly good, neither good nor bad, slightly bad, somewhat bad, or very bad?

How have experiments been used in surveys?

Methodological experiments
- Measurement – validity and/or reliability
- Question wording
- Question order
- Nonresponse bias and/or cost effectiveness – survey design
- Mode
  - Sampling
  - Consent – e.g., advance letters or use of colored envelopes
  - Participation – e.g., reminders, refusal conversion

Substantive experiments
Vignette designs
Can be used for both basic or applied research

Survey Mode

Mode comparisons
- Telephone, face-to-face, Web, and mail
- Random assignment versus completely separate samples drawn from the same population (different sampling strategies)

Effects of mode combinations
- Mode choice
- Nonresponse follow-up in different mode
Mode can affect measurement as well as other aspects of survey design (e.g., participation)
### Other elements of survey design:

**Sampling:**
- E.g., ABS versus more traditional block-listing
- Within household sampling methods

**Contact and/or cooperation:**
- Advance letters
- Different colored or shaped envelopes in mailed questionnaires
- Interviewer training to address common questions/objections to participate
- Incentives
- Wording of introduction
- Leaving voicemail messages (in telephone surveys)

### Substantive research:

**Limitations**
- Less vivid or involved manipulations
- Confederates
- Deception

**Effects of TESS (Time Sharing Experiments for the Social Sciences):**
[http://www.tessexperiments.org](http://www.tessexperiments.org)

**Examples:**
- Priming
- Conformity (results of “survey”)

### Other methodological factors that have been studied using experiments:

**Visual presentation of mailed or Internet surveys**
- Screen order
  - Type of response format (e.g., circling versus checking a box)

**New response formats:**
- Slider bars versus traditional response scale
- Ranking questions that allow the respondent to move items around to physically place them in an order versus asking them to provide numerical responses to rank

### Vignette research:

**Typically within subjects.**
- Involves asking respondents to make a judgment or decision based on descriptions or information that varies.

**Information is often manipulated so that each dimension is independent of other dimensions.**
- Often used to study the effects of candidate attributes or policy positions on candidate evaluations:
  - Respondents are asked to evaluate a series of candidates who have different attributes and policy positions
  - Statistically assess the effect of candidate characteristics

**Limitations:**
- Number and order of vignettes – fatigue; respondents may only see a subset of potential vignettes
- Often lacks realism (hypothetical candidates)

### How have experiments been used in surveys?

**Methodological experiments**
- Measurement — validity and/or reliability
  - Question wording
  - Question order
- Nonresponse bias and/or cost effectiveness — survey design
  - Mode
  - Sampling
  - Contact — e.g., advance letters or use of colored envelopes
  - Participation — e.g., incentives, refusal conversions

**Substantive experiments**
- Vignette research:
  - Typically within subjects.
  - Involves asking respondents to make a judgment or decision based on descriptions or information that varies.
  - Information is often manipulated so that each dimension is independent of other dimensions.

**Examples:**
- Priming
- Conformity (results of “survey”)

**Vignette research:**
- Typical within subjects.
  - Involves asking respondents to make a judgment or decision based on descriptions or information that varies.
  - Information is often manipulated so that each dimension is independent of other dimensions.

**Often used to study the effects of candidate attributes or policy positions on candidate evaluations:**
- Respondents are asked to evaluate a series of candidates who have different attributes and policy positions
  - Statistically assess the effect of candidate characteristics

**Limitations:**
- Number and order of vignettes — fatigue; respondents may only see a subset of potential vignettes
- Often lacks realism (hypothetical candidates)

**Summary:**

Survey experiments valuable because they can have both high internal and high external validity.

Survey experiments have been used widely used for a long time to study methodological questions — increasing popularity of survey experiments to assess substantive questions:
- Concerns about external validity of laboratory studies, and the potential for respondent-level moderators of effects observed in the lab.
- Concerns about internal validity of survey studies — limitations of statistical and design approaches to address in cross-sectional surveys; cost of longitudinal surveys.

Researchers’ ability to conduct experiments in surveys has been heavily affected by technology and is continuing to expand with technological advances.
Two final requests...

Webinar evaluation coming your way

Send ideas for future webinars
(Allyson@uic.edu)

Questions?

THANKS FOR ATTENDING!
ALLYSON@UIC.EDU