10-9-2014

Using Survey Methodology to Improve Surveys Related to Science and Engineering

Jolene Smyth
University of Nebraska-Lincoln, jsmyth2@unl.edu

Follow this and additional works at: http://digitalcommons.unl.edu/dberspeakers

Part of the Quantitative, Qualitative, Comparative, and Historical Methodologies Commons, and the Social Statistics Commons

http://digitalcommons.unl.edu/dberspeakers/64

This Presentation is brought to you for free and open access by the Discipline-Based Education Research Group at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in DBER Speaker Series by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Abstract for DBER Group Discussion on 2014-10-09

Presenter(s), Department(s):
Dr. Jolene D. Smyth
Associate Professor, Department of Sociology
Director, Bureau of Sociological Research
University of Nebraska-Lincoln

Title:
Using Survey Methodology to Improve Surveys Related to Science and Engineering

Abstract:
Dr. Smyth will describe her work with the National Science Foundation’s National Center for Science and Engineering Statistics (NCSES) on the Survey of Earned Doctorates, Survey of Microbusiness Innovation Science and Technology, National Survey of College Graduates, National Survey of Recent College Graduates, and Survey of Doctorate Recipients. She will also present results of research examining the impact of survey mode preference on nonresponse and measurement, a topic of great interest to NCSES as they look for ways to maximize response and data quality for their surveys.
Using Survey Methodology to Improve Surveys Related to Science and Engineering

Jolene Smyth
Associate Professor, Department of Sociology
Director, Bureau of Sociological Research
j Smyth2@unl.edu
The funding mechanism

• 2010-2015 Cooperative Agreement between UNL and USDA-National Agricultural Statistics Service (NASS)
  – Financial support from NSF-National Center for Science and Engineering Statistics through an interagency agreement with USDA-NASS

• Different from a grant
  – 10% indirect rate
  – Continued agency involvement in the work
  – Work needs to be of mutual interest/benefit to UNL and NASS/NCSES
    • Graduate student training is a priority (i.e., RA funding)
Objective

• Conduct survey methodology research to assist with design improvements and/or the redesign of surveys related to science, engineering, and agriculture.
  – Questionnaire design and measurement in web, mail, and mixed-mode surveys
  – Survey recruitment and implementation
  – Post-survey adjustments and imputation
  – Provide consultation to the agencies
  – Provide seminars to the agencies
Questionnaire design and measurement in web, mail, and mixed-mode surveys

- Review questionnaire design with an eye toward
  - Unimode design within surveys
  - Unnecessary design differences across surveys

- Surveys:
  - National Survey of Recent College Graduates (NSRCG) – Cross-sectional survey of science, engineering, or health bachelor’s degree recipients in prior 2-3 years. Tracks trends in education, employment, and salaries of recent grads. Sample institutions and then graduates.
  - National Survey of College Graduates – Longitudinal survey of college graduates used to examine characteristics of the college-educated, occupation, work activities, salary, etc. Sample from ACS and previous NSCG.
  - Survey of Earned Doctorates (SED) – An annual census of doctorate recipients used to assess characteristics of the doctoral population and trends in doctoral education.
  - Survey of Doctorate Recipients (SDR) – Longitudinal survey of doctoral recipients in science, engineering, and health. Followed through their careers from degree to age 76 to understand educational and occupational achievements and career movements of the nationals doctoral scientists and engineers. Sample from SED.
These surveys have had to change with the times

• The SED began in 1957 and the NSRCG, NSCG and SDR all began in early 1970s
  – Survey methods have changed a lot since 1957 and even since the 1970s!

• Most were originally conducted using paper and pencil questionnaires.

• Over time, telephone and web modes were added, raising several challenges.
  – Unimode design
  – Protection of trendlines with changes in data collection
  – Measurement across the surveys contributing to the Scientists and Engineers Statistical Data System (SESTAT) – SDR, NSCG, NSRCG
    • Different vendors (SDR = NORC; NSCG = Census; Mathematica Policy Research)
Example Finding #1: The paper and web versions sometimes unnecessarily used different question formats

**Paper**
- Open-ended boxes
- Horizontally arranged boxes
- “Month” and “Year” labels above boxes
- “Last worked” label to left of boxes
- No symbol use
- Segmented answer boxes
- Separate check box for “never worked”

**Web**
- Drop-down and open-ended boxes
- Vertically arranged boxes
- “Month” and “Year” labels to left
- No “last worked” label
- Use of YYYY symbols
- No segmentation
- No way to report “never worked”
Example Finding #2: The information provided for some questions differed across modes

A27. To what extent was your work on your principal job related to your first U.S. doctoral degree? Was it...

Mark one answer.

☐ Closely related

☐ Somewhat related

☐ Not related

Go to question A30

D2. (If Yes) What type of degree did you earn?

If you completed more than one degree, mark the level for the highest degree awarded.

Mark one answer.

☐ Bachelor's degree (e.g., BS, BA, AB)

☐ Master's degree (e.g., MS, MA, MBA)

☐ Doctorate (e.g., PhD, DSc, EdD)

☐ Other professional degree (e.g., JD, LLB, MD, DDS, DVM) – Specify □

☐ Other – Specify □
Example Finding #2: The information provided for some questions differed across modes

A38. Counting all jobs held in 2009, what was your total earned income for 2009, before deductions?

Include all wages, salaries, bonuses, overtime, commissions, consulting fees, net income from businesses, summertime teaching or research, or other work associated with scholarships.

TOTAL 2009 EARNED INCOME: $___,___,___,___00

© Jolene Smyth, 2014

2010 NSCG
Example Finding #3: “Other” options were dealt with very differently across modes.
etc.

- We identified many more changes across the modes.
  - Things were most messy when there were three modes (paper, web, and telephone).

- Many of the differences were unnecessary and can easily be eliminated
  - Do more quality control of the web instruments contractors put together.

- Others were harder to deal with (e.g., field of study and occupation questions)
  - Paper – find the code from a two-page “field of study” list printed at the end of the questionnaire.
  - Web – type in a response. If it matches an option in the data base, it will be auto-coded. If not, go through a series of items that start broad and narrow down to the specific answer.
  - Phone – Go through a series of items that start VERY broad and narrow down.
Some findings triggered ideas for future experiments (e.g., the “if none” problem)

A separate check box. Where should it go?

An instruction to enter “0”. But where should it be placed?

A filter question?

Did you directly supervise any people?

(If Yes) How many people did you typically directly supervise?
Survey recruitment and implementation

• Currently reviewing the recruitment and implementation protocols for the SDR, NSRCG, and NSCG and synthesizing the last 10 years of internal research on implementation for these surveys.

• We have also examined the concept of “mode preference”
  – Conventional Wisdom: People will be more likely to respond and will give better answers if we survey them in a mode they prefer.
  – NCSES collects mode preference in the SDR and uses that information to tailor later data collections.
  – But they haven’t tested this idea well.
  – We were able to test it on Nebraskans.
Research Questions

• Is “mode preference” a useful predictive or explanatory concept?
  – Does catering to mode preference increase response rates?
  – Does catering to mode preference affect measurement?
Why wouldn’t “mode preference” be a useful concept?

- Previous research suggests self-reports of mode preference are highly influenced by the mode in which the self-report is given (Groves & Kahn 1979).

- Yet many surveyors reason that people have a mode preference and that if we cater to it, we can increase response rates.

- They design survey protocols as if mode preferences are real.

We needed an empirical test of catering to mode preference

• We carried out that test with two questions driving us?
  
  – Are mode preferences meaningful (i.e., do they predict anything)?
  
  – Is catering to mode preferences a good practice?
  
  – If people are more likely to answer and give better responses when answering in their preferred mode then there is something real and meaningful to self-reports of mode preference.
Basic Design

Step 1. Measure Mode Preference

Step 2. Evaluate Participation
Mode preference was measured in the 2008 Nebraska Annual Social Indicators Survey (NASIS)

- Statewide omnibus survey conducted February-August 2008
- Conducted by telephone with a listed landline sample of Nebraska residents age 19+
- AAPOR RR3: 38%
- 1,370 of 1,811 (75.6%) respondents indicated willingness to participate in future social research.
Self-reported mode preferences were as follows:

- If you received a request to do another survey like this one, would you prefer to participate in…
  - An in-person interview at your home
  - An interview on your home phone
  - An interview on your cell phone
  - A paper survey sent by mail, or
  - A survey on the internet

- Home Phone: 49.2%
- Mail: 24.6%
- Internet: 19.7%
- Cell Phone: 0.4%
- Missing: 4.5%
- In person: 1.7%
In 2009, two follow-up surveys were conducted in which some 2008 NASIS respondents were surveyed in their preferred mode and others were surveyed in a non-preferred mode.
Step 1. Measure Mode Preference

2008 NASIS
Feb. – Aug. 2008
1370 Respondents were willing to participate in additional research

Step 2. Evaluate Participation

2009 Quality of Life in a Changing Nebraska Survey

- Phone
  - N=1000
  - RR2=55.5%

- Web
  - N=314
  - RR2=50.0%

Mail
- Mail
  - N=297
  - RR2=53.9%

- Mail
  - N=327
  - RR2=53.2%

Web
- Web
  - N=291
  - RR2=25.4%

2009 NASIS
March – July 2009

1370 Respondents were willing to participate in additional research
Step 1. Measure Mode Preference

2008 NASIS
Feb. – Aug. 2008
1370 Respondents were willing to participate in additional research

Step 2. Evaluate Participation

Phone
N=1000
RR2=55.5%

56.2% preferred phone

Mail
N=297
RR2=53.9%

23.9% preferred mail

Mail
N=314
RR2=50.0%

Web
N=327
RR2=53.2%

Web
N=291
RR2=25.4%

20.6% preferred web

2008 NASIS
Feb. – Aug. 2008
1370 Respondents were willing to participate in additional research
Catering to mode preference seems to make a positive difference, BUT...

- It does not overcome low overall web response rates. Those who prefer the web respond to mail and phone surveys at higher rates than they respond to web surveys.

Response rates are higher for those being surveyed in their preferred mode.

This finding holds when respondent characteristics are taken into account.

For the phone survey, it is driven by higher cooperation rates, not higher contact rates.
Step 1. Measure Mode Preference

2008 NASIS
Feb. – Aug. 2008
1370 Respondents were willing to participate in additional research

Step 2. Evaluate Participation

Phone
N=1000
RR2=55.5%

Mail
N=297
RR2=53.9%

Web
N=314
RR2=50.0%

Mail
N=327
RR2=53.2%

Web
N=291
RR2=25.4%

16.2% preferred mail
23.6% preferred web

Step 1: Measure Mode Preference

2008 NASIS
Feb. – Aug. 2008
1370 Respondents were willing to participate in additional research

Step 2: Evaluate Participation

Phone
N=1000
RR2=55.5%

Mail
N=297
RR2=53.9%

Web
N=314
RR2=50.0%

Mail
N=327
RR2=53.2%

Web
N=291
RR2=25.4%

16.2% preferred mail
23.6% preferred web
There is no difference in response rates between those initially offered their preferred mode and those initially offered a non-preferred mode.

Response Rates in Mixed-Mode Treatments by those Answering in Preferred and Non-Preferred Modes

- **Mail then Web**
  - Preferred Mode: 52.9%
  - Non-Preferred Mode: 49.4%

- **Web then Mail**
  - Preferred Mode: 46.8%
  - Non-Preferred Mode: 55.2%
Does catering to mode preference get responses quicker?

People who prefer mail do not participate in a mail survey faster than those who prefer other modes.

People who prefer web do participate faster in the web survey than those who prefer other modes.
Does catering to mode preference get responses quicker?

But those who prefer a non-web mode participate at much higher rates when it is made available after a web mode.

Offering a second mode after web overcomes any advantage of catering to mode preference.
Mode preference does predict participation, but it is not a remedy for falling response rates

- Catering to mode preference increases response rates in web-only and phone-only surveys, but not in mixed-mode (web and mail) surveys.

- It cannot overcome the overall low response rates to web surveys.

- People who prefer the web answer web surveys faster than people who prefer other modes.

- Self-report mode preference does seem to have some predictive value; it is not all measurement artifact.

Do those answering in their preferred mode show better response behavior?

<table>
<thead>
<tr>
<th>Experimental Treatments</th>
<th>Format encourages satisficing or makes answering more difficult</th>
<th>Format discourages satisficing or makes answering easier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answering in Preferred Mode</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Answering in a non-preferred mode</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

• Those answering in a preferred mode SHOULD NOT take advantage of opportunities to satisfice or have difficulty with tougher items.

• Those answering in a non-preferred mode SHOULD take advantage of opportunities to satisfice or have difficulty with tougher questions.
### Encourages Satisficing/High Difficulty

29. Thinking about your personal finances, which of the following have you done in the past 12 months? Check all that apply.

- Delayed/canceled purchasing a home
- Delayed/canceled plans to buy a car
- Delayed/canceled plans to make a major household purchase
- Cut back on vacation spending
- Cut back on eating out
- Cut back on home internet access
- Cut back on home landline telephone services

**Outcome:** Number of items selected

### Discourages Satisficing/Low Difficulty

29. Thinking about your personal finances, please indicate whether or not you have done each of the following in the past 12 months.

- Delayed/canceled purchasing a home
- Delayed/canceled plans to buy a car
- Delayed/canceled plans to make a major household purchase
- Cut back on vacation spending
- Cut back on eating out
- Cut back on home internet access
- Cut back on home landline telephone services

**Outcome:** Item nonresponse rate, answer length, % with multiple themes, % with elaboration, # of elaborations

### Encourages Satisficing/High Difficulty

32. All things considered, during the next year, what do you think will be your biggest challenges?

**Outcome:** Item nonresponse rate & straightlining

### Encourages Satisficing/High Difficulty

14. Please indicate how satisfied or dissatisfied you are with the availability of each of the following in your community.

<table>
<thead>
<tr>
<th>Parks</th>
<th>Very Dissatisfied</th>
<th>Dissatisfied</th>
<th>Neither Satisfied nor Dissatisfied</th>
<th>Satisfied</th>
<th>Very Satisfied</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bicycle paths</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outdoor areas to hunt, fish, or hike</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sporting events</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restaurants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fine arts (museums and theatres)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell phone service</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internet access</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Libraries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Outcome:** Item nonresponse rate & straightlining
Does catering to mode preference affect measurement?

<table>
<thead>
<tr>
<th>Check-all vs. Forced-choice</th>
<th>Answering in a preferred mode</th>
<th>Answering in a non-referred mode</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
• Our data suggests that those answering in a non-preferred mode take advantage of opportunities to satisfice more so than those answering in a preferred mode.
Summary

• Through our work with NCSES we have been able to identify possible threats to data quality and some solutions to existing problems.
  – Unimode design is needed!
  – Catering to mode preference has some utility.

• We are continuing this partnership to examine
  – Mode effects on measurement (i.e., does answering in different modes impact responses).
  – Ways to improve response rates.
Thank you!
Study 2: Does data quality differ across those answering in preferred and non-preferred modes?

• Data quality is affected by respondent motivation, ability, and the burden of the task.

• We assert that respondents prefer modes they find the least burdensome and the most motivating.
  – In these modes they can focus more energy on answering questions and have to focus less on the survey technology.

• We expect higher data quality in preferred versus non-preferred modes.

• We can examine these issues using the same data as Study 1.
We examine the link between mode preference and data quality in multiple-answer, open-ended, and grid items.

- We chose these types of questions because they are particularly difficult and respondents are more likely to engage in satisficing behaviors on difficult questions.

- For each question type, QLCN sample members were assigned to one of two formats that make it easier or more difficult to provide an optimal answer.
Step 1. Measure Mode Preference

2008 NASIS
Feb. – Aug. 2008
1370 Respondents were willing to participate in additional research

Step 2. Evaluate Participation

<table>
<thead>
<tr>
<th>Phone</th>
<th>2009 NASIS</th>
<th>March – July 2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=1000</td>
<td>RR2=55.5%</td>
<td></td>
</tr>
</tbody>
</table>

Mail
- N=297
  - RR2=53.9%
- N=327
  - RR2=53.2%
- N=327
  - RR2=53.2%

Web
- N=314
  - RR2=50.0%
- N=291
  - RR2=25.4%
- N=291
  - RR2=25.4%

Form A
- n=79
- n=79
- n=83
- n=37

Form B
- n=81
- n=78
- n=91
- n=37
Check-All-That-Apply vs. Forced-Choice

Previous research suggests the check-all format allows satisficing response behaviors, resulting in fewer items selected.

We expect those answering in a preferred mode to be willing and able to process all items, resulting in few differences across the format.

We expect those answering in a non-preferred mode to be less motivated and therefore to take advantage of the opportunity to satisfice in the check-all format leading to larger differences in the mean number of items selected between these two formats.
Those in a non-preferred mode select significantly fewer items in the check-all format than in the forced-choice format.

The significant interaction between format and responding in a preferred vs. non-preferred mode holds in multivariate models controlling for respondent demographics.
Small vs. Large Open-Ended Answer Boxes

Previous research shows that larger answer boxes communicate that more information is needed and thus obtain longer, more detailed responses.

Expectations:
• No differences in response quality across box sizes for those answering in a preferred mode.
• Higher item nonresponse rate in the large box than the small box treatment for those answering in a non-preferred mode.
• Shorter answers with fewer themes and elaborations in the small box versus the large box treatment for those answering in a non-preferred mode.
Those in a non-preferred mode are significantly more likely to leave the item blank when it has a large rather than small answer box.

The significant interaction between box size and responding in a preferred vs. non-preferred mode holds in multivariate models controlling for respondent demographics.
• There was no significant interaction between box size and responding in a preferred vs. non-preferred mode for
  – mean number of words
  – percent giving multiple themes,
  – percent who elaborated
  – mean number of elaborations (among those who elaborated)
Top vs. Fully Labeled Grid

Organizing items into a grid format saves space, but requires horizontal and vertical processing, which is difficult. Grids often produce high item nonresponse and straightlining.

A fully labeled grid reduces difficulty by eliminating the need for vertical processing.

Expectations:
- No difference in item nonresponse and straightlining across formats for those responding in their preferred mode.
- Significantly higher item nonresponse and straightlining rates in the top labeled version than the fully labeled version for those responding in a non-preferred mode.

*Mail mode only; this comparison was not done on the web.
The fully labeled grid had significantly fewer items left blank than the top labeled grid regardless of response mode.

**Mean Number of Grid Items Left Blank by Labeling in Preferred and Non-Preferred Modes**

- **Overall**
  - Top Labeled: 0.4
  - Fully Labeled: 0.2

- **Preferred Mode**
  - Top Labeled: 0.6
  - Fully Labeled: 0.4

- **Non-Preferred Mode**
  - Top Labeled: 0.3
  - Fully Labeled: 0.2

* p<.01
• Straightlining did not significantly differ by the type of grid or whether or not they responded in their preferred mode.
Study 2 Summary

• Those answering in a non-preferred mode did take advantage of some opportunities to satisfice.

• Those answering in a preferred mode did not seem to take advantage of opportunities to satisfice.

• Our findings were not as strong as we would have liked. More research with larger sample sizes is needed before we draw hard conclusions about mode preference and data quality.