

Designing Better UX Surveys

Jean E. Fox

US Bureau of Labor Statistics

UX Research & Strategy

January 20, 2022



How I came to be talking about surveys in UX

- BS from the Human Factors program at Tufts University
 - ▶ Worked at American Institutes for Research
 - ▶ Then continued on at Virginia Tech and George Mason University
- Ended up at the Bureau of Labor Statistics
 - ▶ Consumer Price Index, Unemployment Rate, Jobs Report
 - ▶ Use surveys to collect data
 - ▶ Survey Methodologists help design our surveys



UX in the US Government

(Just a quick side note supporting UX career opportunities in government)

- Human Factors type work
 - ▶ Military, Transportation
- UX type work
 - ▶ Websites, systems, services
- All levels of government
- Some US agency Innovation Labs follow human-centered design approaches
 - ▶ OPM, State, USAID
- Rewarding, fun, interesting challenges, great colleagues



What are surveys?

- Protocols seeking information from groups of people systematically
 - ▶ Goal is to generalize or to provide estimates of the larger population
- Can be called Questionnaires or Surveys
 - ▶ Recommendations apply to Polls or Interviews as well
 - ▶ Online or on paper
- Surveys have many uses in UX:
 - ▶ Usability tests (post-task and post-test)
 - ▶ Ethnographic work
 - ▶ Stakeholder feedback
 - ▶ Market research



Goals for Today

- Why survey design is important for UX work
- Learn from field of Survey Methodology
- Share best practices for designing (and testing!) better UX surveys
 - ▶ Selecting topics to be covered in your survey
 - ▶ Writing survey questions
 - ▶ Testing survey questions



What we won't be covering today

■ Sampling

- ▶ Are you asking all the right people to complete your survey?
- ▶ Are they responding?

■ Motivation

- ▶ What encourages people to complete your survey?
- ▶ Do incentives help?

■ Data cleaning

- ▶ How do you decide whether a response is valid?
- ▶ What do you do with outliers?

■ See Caroline Jarrett's presentation on planning surveys

- ▶ <https://www.effortmark.co.uk/survey-octopus-getting-valid-data-surveys/>



Who's doing survey work in UX?

- MeasuringU.com

- ▶ Jeff Sauro

- ▶ Jim Lewis

- Caroline Jarrett at Effortmark

- ▶ Relevant books: “Forms that Work” and “Surveys that Work”

Why talk about surveys?

- They're not the latest UX method
 - ▶ Or the hottest technology
- But they are everywhere in UX
 - ▶ And we don't talk about them much
- And they are so easy to do poorly
 - ▶ There are lots of examples at [@BadSurveyQ](#) on Twitter

Problems in surveys do exist

Cancel App feedback

How was your [redacted] experience today?
Required – select one

Neither good nor poor

Very poor

Very satisfied

Satisfied

Poor

- Scale is inconsistent
- Options are randomly ordered
- Not all options answer the question directly

Posted by @Shminkleberry 9/5/2019

Problems in surveys do exist (2)

- Similar situation, where the responses don't answer the question.

During your most recent visit, how satisfied were you with the following COVID-19 measures in-store....?

Sanitation in-store

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree	Did not notice
-------------------	----------	----------------------------	-------	----------------	----------------

1	2	3	4	5	6
---	---	---	---	---	---

N/A

More Examples

“ I don't care about sports, I don't understand how some people are so obsessed! ”

DEFINITELY DISAGREE (red box with white X) TEND TO DISAGREE (orange box with white X) TEND TO AGREE (light green box with white checkmark) DEFINITELY AGREE (green box with white checkmark)

Neither agree nor disagree ►

Posted by @BadSurveyQ 9/5/2019

- The 2 statements are independent (it's double-barreled)
- Colors and icons could bias responses
- Option for “neither” is hidden

More Examples (2)

“ I don't care about sports, I don't understand how some people are so obsessed! ”

DEFINITELY DISAGREE TEND TO DISAGREE TEND TO AGREE DEFINITELY AGREE

Neither agree nor disagree ▶

Posted by @BadSurveyQ 9/5/2019

- Another example of a similar scale.

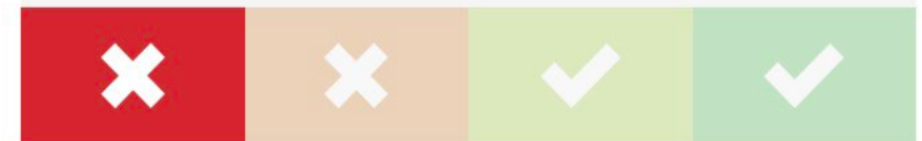
Do you agree or disagree with the following statement?



DEFINITELY DISAGREE



I'm not bothered by flashing features in cars



Neither agree or disagree

Posted by @BadSurveyQ 9/24/2020

More Examples (3)

6. On a scale from 1 to 10 how much do you miss going to the theatre?



Posted by @markrt 11/24/2020

- How do you tell a non-response from a “5”?
- What is the actual midpoint of a scale from 1-10?

More Examples (5)

Which of the following best describes how you feel about the aircraft that you just flew?

- I would actively avoid flying this aircraft again
- I don't want to fly this aircraft again
- I would like to fly this aircraft again
- I would go out of my way to fly this aircraft again

- Options are missing (“No opinion”)
- What happens if this is a required question?
 - ▶ People will give you bad data or drop out.

More Examples (6)

- Be sure you include all the options
 - ▶ Missing 6-12 months
 - ▶ 2-3 years
 - ▶ 4-5 years, etc.

What's the age of the youngest child?

- 0 - 3 months old
- 4 - 6 months old
- 1 - 2 years old
- 3 - 4 years old
- 5 - 6 years old
- 7 - 8 years old
- 9 - 10 years old
- 11 - 15 years old
- Prefer not to say

Survey Design is Important!

- You want to make it as easy as possible for your respondents to answer your questions
- You want to be sure you are getting the information you need
- Important decisions are often based on survey results
 - ▶ BUT Designing good surveys is harder than it seems
- Fortunately, there are resources
 - ▶ Academic research
 - ▶ Survey Methodologists



Selecting Content



Selecting Survey Topics

- What do you really need to know?
 - ▶ Have clear research questions
 - ▶ Don't ask just what might be interesting
 - ▶ Can you get the data somewhere else?
- Are you asking questions that your target audience can answer?
 - ▶ They may not know the answer
 - ▶ They may not remember the answer
- How will you analyze all the data?
 - ▶ Especially for open-ended questions



Writing Questions



Survey Response Model

■ Comprehension

- ▶ Understanding the question

■ Retrieval

- ▶ Identify the information you need to respond

■ Judgement

- ▶ Deciding on your answer

■ Response

- ▶ Mapping your answer to the response available

Tourangeau, Rips, and Rasinski, *The Psychology of Survey Response*, 2000



Best Practices for...

- Rating Scales
- Rankings
- Double-barreled questions
- Agree/Disagree questions



What are rating scales?

- A continuum of one construct
- Covering the full range of the continuum
 - ▶ Sometimes “Don’t know” or “Not applicable” are needed
- Not a list of categories



General guidance for rating scales

- Response options should directly answer the question posed
 - ▶ Consider the survey as a conversation between you and a respondent
 - ▶ How would they answer the question if it were open-ended
- Choose qualifiers (like “Not at all,” “Very”) to create an evenly distributed scale
 - ▶ As best you can
- Cover the full range of options



Types of Rating Scales (Bi-polar)

- Bi-polar (opposites at end of scale)
 - ▶ Be sure ends are opposite and scale is symmetric
 - ▶ Frequently “Very Not-X” to “Very X”
 - ▶ Often called Likert-type scales

	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
I would use this tool in my work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Types of Rating Scales (Uni-Polar)

- Uni-polar (from none to a lot)

Was the documentation helpful?

- Extremely Helpful
- Moderately Helpful
- Somewhat Helpful
- Not at all Helpful



Example

Neutral:

- Doesn't answer question directly
- Not on the continuum

23. How useful would you find the following content items on the website:

	Not useful at all	Somewhat useful	Neutral	Useful	Very useful
A brief overview before I "dive deeper" into the content	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>

Posted by @BotchedEmerald July 29

Size of Rating Scale

(Number of Response Options)

- How many response options should there be?
 - ▶ Scales with 3 points don't give you any intensity and have lower reliability
 - ▶ Scales with 5-7 options are the most reliable
 - ▶ But it depends on the survey
 - Experts may be able to detect more differences
 - May want more options on bi-polar than uni-polar scales.

Size of Rating Scale (Odd or Even Number)

- Odd or even number of options?
 - ▶ Odd number gives a midpoint for bi-polar scales
 - There is no neutral midpoint on uni-polar scales
 - ▶ For UX work, a midpoint on bi-polar scales is usually good
 - Unless you want to find which way respondents lean

Response Labels (1)

■ Labeling all response options

- ▶ Especially valuable for objective scales like frequency
- ▶ Less important for subjective scales (but use at least endpoint anchors)

■ Using numbers

- ▶ Avoid numbers unless they are meaningful
- ▶ Use negative numbers judiciously, because respondents may avoid negative responses.

Response Labels (2)

- Be sure the scale is balanced

My recent stay at The [REDACTED]

	Completely Agree	Agree Very Much	Agree	Somewhat Disagree	Disagree
Made me feel like a valued guest	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Was a good value for the price paid	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

NEXT

- This scale has 3 “agree” options and only 2 “disagree” options, so the results will skew towards “agree”



Ratings vs Rankings

- **Rating:** Select a value for individual items from a scale
- **Ranking:** Select an order for the items, comparing each against all the others.



Rating or Ranking?

You are going on a diet and need to stop eating certain foods.

Rate the following foods from 1 to 10 where 1 is the first food you would give up and 10 is the last food you would give up.

	1	2	3	4	5	6	7	8	9	10
Bread	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cookies	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cake	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Candy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pasta	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Potatoes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Soda	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fruit Juice	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ice Cream	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cereal	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- Would you give ratings or rankings?
- For this question, people answered as follows:
 - ▶ Rank = 57%
 - ▶ Rate = 37%
 - ▶ Can't tell = 2%
 - ▶ No response = 4%
- Had to analyze rankings and ratings separately



Ranking

- Ranking is difficult
 - ▶ Requires many comparisons
 - ▶ Gets harder with longer lists
- Resulting data are less useful
- Recommendations
 - ▶ Use ratings when you can
 - Determine ranks from average ratings
 - ▶ Use rankings only if you need respondents to make decisions about the order of options



Double-Barreled Questions

- Double-barreled questions force respondents to provide a single response to what may be multiple questions.

	Not Applicable	Limited	Fair	Neutral	Good	Very Good
How would you rate your ability to plan and evaluate (to organize work, set priorities, and determine the resources you need to accomplish your job)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- They assume that the topics will logically group together, which may or may not be true.

Double-Barreled Questions (2)

■ Recommendations

- ▶ Watch for “and” in questions, which can indicate a double-barreled question
- ▶ To address double-barreled questions
 - Break them into multiple questions
 - Focus on one construct and delete everything else
 - Leave them as is (they may be ok)

Agree / Disagree Items

■ Why use them

- ▶ They are easy to write.
- ▶ You can cover lots of items with one scale.
- ▶ You can use them in grids.
- ▶ The scale is fairly standard and familiar to respondents.

■ Why NOT use them

- ▶ People like to be agreeable, so responses may be biased towards “agree” (the “acquiescence bias”).
- ▶ They require an additional level of cognitive processing to translate the response to the agree/disagree scale.

Agree / Disagree Items (2)

■ Recommendations

- ▶ Avoid qualifiers like “very,” as in “I thought the task was very easy.”
- ▶ Avoid Agree/Disagree items if there is a good alternative
- ▶ Use “construct specific” responses that reflect the construct you are measuring.

The product was easy to use.

- Strongly Agree
- Agree
- Neither Agree nor Disagree
- Disagree
- Strongly Disagree

Was the product easy or difficult to use?

- Very Easy
- Easy
- Neither Easy nor Difficult
- Difficult
- Very Difficult

Summary of Best Practices (page 1)

- Consider the survey as a conversation with the respondents
 - ▶ Be sure the responses match the question
 - ▶ Avoid jargon unless the respondents are familiar with it
- Use required questions carefully
- Rating scales options should be evenly distributed, balanced, and follow a continuum
- For rating scales, use text labels that are appropriate for the question (either anchors only or all options)

Summary of Best Practices (page 2)

- Include all possible options
 - ▶ Be sure scales are complete
 - ▶ You may need “Don’t know,” “Not applicable,” or “Other”
- Use ratings rather than rankings when possible
- Avoid double-barreled questions
- Consider construct-specific formats rather than agree/disagree questions when you can

Testing the Survey



Why test surveys?

- Be sure questions are measuring what you want
- Just like any other effort, you design following best practices, but still conduct usability testing
- There can be all kinds of problems
 - ▶ Complicated instructions and question wording
 - ▶ Unfamiliar jargon
 - ▶ Incorrect assumptions about what your respondents know
 - ▶ Unexpectedly sensitive topics
 - ▶ Questions that are difficult to answer
 - ▶ Answers that are difficult to recall

Testing Surveys

- First step, review the final survey in whatever tool you are using
 - ▶ It may look different online than on paper
- Look for typos or other errors in setting up the questions



Improper question formatting

4.2 Does your institution provide *mandatory or optional* information security training?

	Yes	No
No	<input type="radio"/>	<input type="radio"/>
Yes, mandatory training	<input type="radio"/>	<input type="radio"/>
Yes, optional training	<input type="radio"/>	<input type="radio"/>
Don't know	<input type="radio"/>	<input type="radio"/>

Posted by @ewinsberg 5/21/2020

How to Test Surveys

- Cognitive Interviewing
 - ▶ Similar to usability testing, but with some differences
- Have participants complete the survey
- Afterwards, ask them questions (probes), such as
 - ▶ In your own words, what was the question asking?
 - ▶ What did you consider in determining your response?
 - ▶ Was there anything difficult about this question?
- Focus your analysis on their answers to the probes
 - ▶ Unlike usability testing, where focus is on actual behaviors



Summary



Decide what you really need to know



Write the questions following best practices



Test the survey

Contact Information

Jean Fox
US Bureau of Labor Statistics

Fox.Jean@bls.gov
[@JeanHarrisFox](#)

