



World Leaders in Research-Based User Experience

Iterative Design of a Survey Question: A Case Study

Summary: Through 4 rounds of revisions, we made a survey more specific and usable. Running pilot studies before conducting a full-scale survey ensures you'll reach your research goal.

By [Feifei Liu](#) on October 6, 2019

Topics: [Research Methods](#)

Survey questions are delicate things. Even small details in wording can affect how your respondents interpret and answer them. A carelessly written question can ruin a study, so it's worth a little extra time to perfect your survey.

Case Study: Survey on How People Use the Web

Recently, we decided to replicate [a study](#) conducted 21 years ago by researchers at Xerox PARC. The original study investigated how the information found online affects people's decision making. The study consisted of a large-scale survey in which 3,292 respondents described in detail a situation where online content impacted their decisions or actions.

Today, people rely even more heavily on information found on the web than 20 years ago. From purchasing a house to deciding where to eat for dinner, the web helps users make a huge variety of decisions. Thus, [we replicated the study to see if the important online information-seeking behaviors have changed over two](#)

[decades](#).

In the Xerox survey, the researchers asked the following single question:

Please try to recall a recent instance in which you found important information on the World Wide Web, information that led to a significant action or decision. Please describe that incident in enough detail so that we can visualize the situation.

While we wanted the responses to be comparable to the 1998 study, we realized that we would likely need to tweak the question's wording to ensure we'd collect information that reflects today's use of online services. Through 4 rounds of [pilot testing](#) the survey, we were able to refine the question.

1st Round of Testing

We wanted to keep the question as close to the previous version as possible to make a valid comparison.

For the first version of the question, we only changed "World Wide Web" to "online" to reflect current terminology. [Google Ngram](#) showed that in 1998, the word "online" only appeared about 1.5 times more than the phrase "World Wide Web" in the Google Books corpus, but in 2008, the frequency of the word "online" was already more than 100 times higher than that of "World Wide Web." Also, [Google Trends](#) showed that the related queries of the word "online" included "online films" and "online games", while the ones of the phrase "World Wide Web" included "World Wide Web Wikipedia" and "who created World Wide Web," suggesting that today people use the word "online" to refer to the things they can do on the World Wide Web.

Thus, we rephrased the survey question as follows.

*Please try to recall a recent instance in which you found important information **online**, information that led to a significant action or decision. Please describe that incident in enough detail so that we can visualize the situation.*

For this phase, we recruited 11 participants who filled out a written survey and we collected their verbal feedback at the end of the survey.

Four of these pilot participants reported that this question was too general, and they were not sure what we wanted. This was probably not a problem 21 years ago, but now it is because of the pervasiveness of the internet. According to [a report by USC](#), the time spent online increased from 9.4 hours per week in 2000 to 23.6 hours per week in 2016. An article by [Clickz](#) showed that, on average, in 2019, people spent 6 hours and 42 minutes online per day. Gathering information online has become such a frequent and mundane task for many people that they struggled to pick out a specific instance to report.

To address this problem, we added an explanatory sentence in the second design.

2nd Round of Testing

Please try to recall a recent instance in which you found important information online, information that led to a significant action or decision. Please describe that incident in enough detail so that we can visualize the situation.

A significant action or decision can be any change in your plans, thoughts, or actions that you consider to be meaningful.

We thought that giving a bit more explanation on “significant” could ease people’s concerns that their actions might not meet our standards. This version of the question was tested with 5 users; the survey was remote and unmoderated.

In this second pilot, people were constrained by the explanatory text and talked only about the *changes* they made because of the online information. For example, a participant wrote, “I looked at the weather on the app on my phone before I left for work in the morning. It said the temperature was colder than I expected it to be. So I put on a warmer coat and a hat”. Another job-seeking participant talked about how online information “changed her applying strategy”

participant talked about how online information changed her applying strategy

and made her focus on certain types of companies. Almost all the responses were related to some specific changes, but change should not be a necessary aspect of a significant decision or action. We realized that adding explanations to "significant" could bias the respondents' answers. We decided to remove the clarifying sentence and try another approach.

3rd Round of Testing

For the third round of testing, we tried adding a multiselect question before the main question about the respondents' significant activity.

Which of the following online activities have you done in the past month? (Please select all that apply)

- ☐ *Bought something*
- ☐ *Watched a TV show or movie*
- ☐ *Planned a vacation*
- ☐ *Sent an email*
- ☐ *Posted on social media (for example, Facebook or Instagram)*
- ☐ *Researched a topic*

We hoped that this question could help users reflect on their recent online activities, and that this process may help them answer the following question. We carefully balanced different kinds of activities — from entertaining to serious ones. We invited 4 users to fill out the revised version of the survey, and also conducted a [cognitive walkthrough](#) with 3 participants to gain insights on the language of the survey.

Unfortunately, all of the participants in this group ended up reporting activities that sounded too similar to our multiselect responses. Seven users were all talking about the research they did online, like “researching information about tax base transfer in California” and “looking up information regarding weight-loss surgery.” Not all significant decisions or actions have to be related to research, so we realized that the multiselect responses were **priming** our respondents. Namely,

the [last response option](#) in the first question, “researched a topic”, primed the participants to come up with research-related answers in response to the second question. We decided to remove the priming question from the survey.

4th Round of Testing

At this point, we were quite confident that the biggest problem was that people had too many online activities to choose from. They needed reassurance that they could choose just one to report. That could help explain why pilot participants were confused when presented with the original question and why they were easily influenced by the changes we’d tested: people weren’t sure which decisions counted as “significant” and which ones didn’t, so they tried to find clues from other information researchers provided. This was probably not a problem during the original PARC study because, at that time, the Internet was not pervasive and didn’t impact people’s lives as much.

Based on this insight, we revised the question again, to include a clarification that could help respondents if many instances came to mind.

Please try to recall a recent instance in which you found important information online, information that led to a significant action or decision. Please describe that incident in enough detail so that we can visualize the situation.

If you can recall several such instances, please describe the one that was the most important to you.

With this addition, we reassured users that they could reply to the question with the one example they believe is the most significant to them.

We piloted the survey online and collected 50 responses.

The 50-person pilot survey went well; we got a diverse set of responses. Besides researching to make decisions, a participant mentioned that “Buying my current phone, Google Pixel 2 XL. Kept seeing commercials on Hulu about it”, which showed that an ad influenced her decision. Another response described how she

got a ticket for one of her favorite-band's concerts because of a notification she received on her phone.

Satisfied with the detail and variety of the data we had collected, we decided to run the full study based on this version of the main question. We collected 700 responses that we analyzed both quantitatively and qualitatively to better understand the current profile of online information-seeking behavior. (Our findings from the final study will be reported in a subsequent article.)

Tips for Survey Design

1. **Make sure that your research questions can be investigated with your survey methodology.**

Surveys cannot answer all research questions. They are good at helping us capture attitudinal data, but not behavioral data. The details and the contextual information they can provide are also limited. In our case, we wanted to identify online information-seeking behaviors that could lead to significant decisions and actions. A [survey](#) can address this goal. But if we want to understand why people choose certain types of information-seeking behaviors instead of others or when and where they engage in these behaviors, surveys are not appropriate. Instead, [user interviews](#) or [field studies](#) can work better in these situations.

2. **Avoid priming or asking leading questions.**

Keep the language of the survey questions neutral. People are social animals who can interpret subtle clues and [try to behave](#) as (they assume) researchers want them to, even subconsciously. As we saw in this case study, minor changes in phrasing the same question or adding another question before it can result in dramatically different responses.

3. **Run pilot studies. You can test several versions at the same time.**

Sometimes, you may not be able to tell if your survey language is neutral

enough until you run it with real people. For your very first pilot, your colleagues or people in a coffee shop could act as testers. However, conduct at least one round of pilot testing with respondents from your demographic of interest — don't rely just on your coworkers. Ask your participants to think aloud as they are completing the survey, to help you identify any interpretation issues or potential leading questions. Testing 5–10 users for each version of your pilot should work fine.

4. Pay attention to the timing of collecting the responses.

Sometimes the time when you send out an online survey can impact the number and quality of your answers. In our study, half of the participants were sent the survey on a weekend and half on weekdays. We did that to avoid biased results related to the timing of response collecting. If your users are likely to be busy during the daytime, sending out a survey at 9:00 A.M. may prevent you from collecting high-quality data.

Poor phrasing, ambiguity, or the wrong sequence of questions can easily result in skewed survey results. Iron out any such issues before you spend the money to collect your data. Like user-interface designs, surveys need to be tested. In fact, a survey instrument *is* a design, so treat it as such.

References

Morrison, J.B., Pirolli, P. and Card, S.K., 2001, March. A taxonomic analysis of what World Wide Web activities significantly impact people's decisions and actions. In *CHI'01 extended abstracts on Human factors in computing systems* (pp. 163–164). ACM.

USC Annenberg Center for the Digital Future. [The 2017 Digital Future Report.](#)

ClickZ. [Internet growth + usage stats 2019: time online, devices, users.](#)

Share this article: [Twitter](#) | [LinkedIn](#) | [Email](#)

Copyright © 1998-2022 Nielsen Norman Group, All Rights Reserved.