Interview with an Expert Series – Usability Testing in Survey Research

Nicholas Parker*

Tags: usability, eye-tracking, survey, usability testing, pretesting

Survey Practice

Vol. 11, Issue 2, 2018

This installment of the Interview with an Expert Series focuses on the topic of Usability Testing. I had the pleasure of interviewing two experts in the field, Emily Geisen, Survey Methodologist and Usability/Cognitive Testing Manager with RTI International, and Jennifer Romano Bergstrom, the Director of User Experience Research at Bridgewater Associates. Emily and Jennifer came to this field from different avenues, but both have developed a deep appreciation for the value of testing early and often when developing a survey instrument. In the interview, they discuss their career paths, the usability testing process, and much more.

Interview conducted by Nicholas Parker, Survey Director with the Federal Reserve Bank of Atlanta

This installment of the Interview with an Expert Series focuses on the topic of Usability Testing. I had the pleasure of interviewing two experts in the field, Emily Geisen, Survey Methodologist and Usability/Cognitive Testing Manager with RTI International, and Jennifer Romano Bergstrom, the Director of User Experience Research at Bridgewater Associates. Emily and Jennifer came to this field from different avenues, but both have developed a deep appreciation for the value of testing early and often when developing a survey instrument. In the interview, they discuss their career paths, the usability testing process, and much more.

1. JUST SO THAT WE'RE ALL ON THE SAME PAGE, WHAT IS USABILITY AND USABILITY TESTING AS IT RELATES TO SURVEYS?

Emily: Usability is the extent to which users – typically respondents, but sometimes interviewers – can use a survey with accuracy, efficiency and satisfaction.

Jennifer: Right, in usability testing, we want to make sure people can complete tasks – typically these tasks include answering questions, navigating, and finding information. We assess if they can do these things quickly and accurately. Usability testing is primarily just watching your users try to accomplish tasks on the survey. You observe what they do – the steps they take, and any errors they make; you note what they say during think aloud

* **Institution:** Federal Reserve Bank of Atlanta

or in response to any questions (probes) you ask them; you can also collect implicit data like eye-tracking data that can also help you understand what they're doing. Then you think about what you observed, figure out the major issues, recommend changes to fix them and test again until optimal usability has been achieved.

Emily: Or until you run out of time or money. Haha.

2. YOU BOTH HAVE SIGNIFICANT EXPERIENCE IN THE USABILITY TESTING ARENA. CAN YOU TELL ME MORE ABOUT YOUR BACKGROUNDS IN USABILITY TESTING?

Emily: My educational background is in survey methodology. Although I learned about questionnaire design and survey pretesting, there was very little discussion about usability or user experience in the courses I took. At that time though, nobody was doing surveys on mobile phones so usability was just not as popular as it is now. I did not learn about usability testing until I started working at RTI International in the Survey Research Division. We had a few staff who were experimenting and learning about assessing the usability of self-administered web surveys when I started. I thought usability was fascinating, and it tied in really well with my questionnaire design and pretesting experience, so I tried to learn more about it from my coworkers and by reading articles by Don Dillman and Cleo Redline and Jakob Nielsen. I took a usability workshop from Steve Krug, author of *Don't Make Me Think*. It was a good course, but not specific to surveys so there was still a lot to learn about applying usability best practices to survey research. Mostly, I learned from doing my own usability testing on lots of different types of surveys. As technology advanced and screen sizes became smaller and smaller, the importance of usability become more obvious and I started doing more research and exploration in the area.

Jennifer: I have always loved psychology, and after majoring in psychology in undergrad, I ended up in an Experimental Psychology PhD program, studying implicit learning and cognitive aging. In my last year of grad school, I applied for an internship at the US Census Bureau's Usability Lab. I did not know what usability even was at that time! But the job posting mentioned running studies with adults, analyzing and presenting findings - and I knew how to do those things well. I quickly learned about usability testing and fell in love with it. I loved how we could use psychological principles to understand behavior and improve products. It was there that I challenged myself to grow by deeply understanding eye tracking. We were using it, but not in great depth. So I took it upon myself to learn more about the data we could collect, and I trained others on how to use it. Little did I know that eye tracking would end up being an area of expertise for me! From there, I landed a leadership position at Fors Marsh Group (FMG), an applied research firm in Arlington VA, where I formed, trained, and led a UX Research team. After three years, I ran into an old colleague from the Census Bureau at an American Association for Public

Opinion Research (AAPOR) conference (we used to see each other annually – Note: the power of conferences and networking!). He had since started at Facebook and soon after the conference, he reached out to me regarding a position. I was so excited – really because I loved San Francisco, but also the opportunity to conduct research on a product that connects millions of people was really enticing! I led UX research for a number of teams at Facebook, and it was here that I learned how to conduct international research and in-home interviews. The skills I gained when learning to conduct in-home interviews varied greatly from the lab settings that I was used to! After two years at Facebook, I moved over to Instagram, where I worked largely with a group of engineers. This was very different from my prior studies, which focused more on design; yet it was similar because the end goal was to understand users' interactions and how to improve the experience for them. Now I am at Bridgewater Associates, where I am leading UX Research for a team of designers and engineers. We are working on a product designed to promote "idea meritocracy" by way of an artificially-intelligent coach that helps to ensure the best ideas rise to the top, regardless of one's prior experiences. While the field continues to move forward quickly, I too continue to learn and grow.

3. HOW WOULD YOU DESCRIBE THE GOALS OF USABILITY TESTING?

Emily: When pretesting survey questions, the main goals are to improve data quality by reducing measurement error and reducing cognitive burden. These two goals are not mutually exclusive. Often we can both improve data quality and reduce cognitive burden. There are other times, though, when the steps it takes to improve data quality increase the length of the question or require more steps to complete, which can increase respondent burden. During those times, we have to decide which goal is most important, and it will vary with the context and the project goals. For example, given the option of lengthening the questionnaire by five minutes or accepting a small amount of error due to a usability issue, you might opt for keeping the shorter questionnaire and accept the small amount of error.

4. WHAT IS INVOLVED IN USABILITY TESTING FROM A PROCESS PERSPECTIVE?

Jennifer: First you need to figure out what the study objectives are - what are you aiming to learn with the usability test? As Emily mentioned, the goal is usually to reduce measurement error, but in usability testing, we are specifically concerned with how people use the survey - how they understand terminology, how they select responses, how they proceed through the survey - and how that affects measurement. By observing and listening to users as they complete tasks, we assess things that are going well - things that are clear - and we assess things that are not going well - for example, when people have trouble responding to an item or finding information. When things do not go well and we identify issues, we determine why people had those issues and what we can do to fix

them. For example, was it the placement of the button that was problematic or the button label? We then report what we have learned and either provide recommendations to improve the product, or we brainstorm with the team about what to change. What actually gets changed depends on who you are working with and what your relationship is with the team.

5. HOW DOES USABILITY TESTING DIFFER FROM COGNITIVE INTERVIEWING?

Emily: They both have the same goal – improve data quality and reduce respondent burden so the distinction can sometimes be difficult to see. The main difference is that cognitive testing focuses on whether the respondent can understand and answer the question and usability testing focuses on whether people can use the survey as intended. Someone may understand the survey question perfectly and have no trouble coming up with their answer, but they select the wrong response because the radio buttons are too small and placed too close together. The way people understand questions and the way people use a survey often interact. Therefore, researchers often conduct combined cognitive and usability testing with the goal of exploring both how people understand the question and how they interact with the survey tool.

6. WHEN IS THE BEST TIME TO CONDUCT USABILITY TESTS?

Emily: As early as you can. You don't have to wait for the final version of the survey. This is an area where survey researchers can benefit from the techniques used for website and software development. Usability testing is often conducted continuously throughout product development. That way the user feedback can guide development. For example, if you're considering the use of grids in your survey, you may want to test a few options for optimizing these questions for mobile devices before finalizing the questionnaire content. You may learn that these items would work better as individual questions rather than as grids. We often do some early rounds of usability testing while we are still finalizing the question wording.

Jennifer: Absolutely - the worst thing you can do is to wait for the survey to be complete before testing. If you find big issues at this point, the likelihood of getting them fixed is low. Start early, test often. You can even put paper prototypes of surveys in front of users to identify issues early on!

7. WHO SHOULD BE INVOLVED IN THE USABILITY TESTING PROCESS?

Emily: It's best to include a multidisciplinary team. My experience is in questionnaire design, but I am not a programmer, so I like to involve the programmers in the process and have them observe as many interviews as possible. That way the programmers can help with problem-solving by indicating what is an easy fix and what would be extremely complicated to program. That way, I don't waste time recommending design changes that are

simply not feasible.

It's also great when your clients or other decision makers can observe as many usability tests as possible. Seeing is believing, and it's way more interesting than reading a 60-page report on the topic. It can speed up the turnaround process considerably.

8. WHAT IS THE MOST COMMON ISSUE UNCOVERED THROUGH THIS PROCESS?

Emily: Honestly, I learn something new every time. I hope that's because I am learning from my mistakes, so I am not making the same mistakes over again. But it's also because the survey landscape is changing constantly. Errors we made previously are obsolete because of new technological options. But new technological features lead to new usability issues.

Jennifer: We often find navigation issues and terminology issues. These seem to come up the most. But honestly, there is always something new, something unexpected. It never gets old, and even if I have conducted a review before testing, and I think I found all the issues, inevitably, we learn something new once we put the survey in front of actual users. Their experience is different from mine, so the unique perspectives they bring enable us to discover different usability issues.

9. COST CONTAINMENT IS A CONCERN FOR MANY SURVEY RESEARCHERS. HOW WOULD YOU CHARACTERIZE THE RETURN ON INVESTMENT FOR USABILITY TESTING?

Emily: Usability testing is not a huge investment. Most of the equipment needed (besides eye tracking, which is not required) is inexpensive. At a minimum you will need a high-quality webcam (<\$90), a USB microphone (\$25), and sharing/screen recording software (\$15-50). For researchers who do a lot of testing, higher-end usability software (\$1,000+) can be useful because it offers ways to automate some of the processes and provides more sophisticated recording options. For mobile testing, a mobile "sled" is useful. It's a platform that holds the device and a webcam so that the device screen and participant's fingers can be recorded. Higher-end sleds (\$350) are adjustable – fitting any size mobile device – and portable, allowing participants to pick up the devices and use them as they normally would. Less expensive options include document cameras or stands that you rig up yourself and simply add a webcam to.

Because it's primarily qualitative, usability studies rely on small numbers of participants. Plus, use of screen sharing and screen recording make it so that travel is not necessary. As a result, the return on improved quality is high. Major errors that lead to break-offs or keep people from accessing the survey would probably be identified in a pilot test. But there are other errors that you would never realize without usability testing. That's where the return on investment is highest. For example, during one usability study we found that when users

swiped to the left to navigate to the next question on their mobile device, they were inadvertently changing their answers without realizing it. The team decided to remove swipe navigation and use Next and Previous buttons only. If we had fielded the survey without usability testing, we would have gotten bad data and not even known it.

Jennifer: Those are what I call the "silent killers" because it's not immediately obvious there's a problem with your data until you start digging in and realize that it doesn't make sense. By then, it's too late to fix. Think about how costly it is to have a survey that does not work or does not capture the data you think it is capturing, and you have to scrap it or redo it! It is far less expensive to conduct usability testing, identify issues, and improve the survey before launching than it is to move forward without a single usability test.

10. IF YOU COULD GIVE ONE PIECE OF ADVICE TO THOSE INTERESTED IN INCORPORATING USABILITY TESTING INTO THEIR SURVEY DESIGN WORKFLOW, WHAT WOULD IT BE?

Emily: Any usability testing is better than no testing. Add it wherever you can – start with a few coworkers just before you plan to launch. Once you see the value, it will be easier to start planning for usability testing and incorporating into the survey development lifecycle.

Jennifer: Absolutely. Anything is better than nothing. Start with a few colleagues, and you will see the value immediately. But once you learn the value, be sure to recruit non-employees, as external participants will be unbiased and will not know the history of the survey, they won't be familiar with the culture and jargon, and they will be less likely to tolerate usability issues.

11. THE GROWTH IN MOBILE SURVEYS HAS INFLUENCED MANY AREAS OF THE SURVEY SPHERE. HOW HAS IT INFLUENCED THE FIELD OF USABILITY TESTING?

Jennifer: We now have to think about the mobile experience along with the desktop experience. Different issues will be discovered when you test with mobile. The screen size plays a big role in how much information people can see at once, or where their fingers land when they click, and there are so many different screen sizes! People also complete surveys on the go - they are not all sitting at their desktop computers when they complete surveys. So it is really important to test mobile too. Ideally you want to conduct usability testing with all possible devices that will be used in the real world once your survey launches. This is not always possible, and this could mean the study will take longer or be more expensive. But it is really important to try to catch as many issues as you can before launch. One way we include different devices in testing is to ask participants to use the device they would use in the real world. Another strategy is to alternate the device that participants use in testing. You can also have participants start on one device then move to another device at some point

in the session (e.g., half way through the survey). You can get creative in how you include mobile in testing.

12. HOW WOULD YOU ADVISE SOMEONE WHO WAS INTERESTED IN SPECIALIZING IN USABILITY TESTING?

Emily: Jen and I recently published a book, titled *Usability Testing for Survey Research* (Elsevier, 2017). It's a great place to get more information on what usability testing is and how to conduct it. We also offer occasional short courses and online courses on the topic.

In addition to reading our book, a strong background in questionnaire design, particularly for web surveys is helpful. I recommend reading *Designing Effective Web Surveys* (Couper, 2008) and *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method* (2014).

I also think that general experience with qualitative interviewing and survey pretesting is helpful.

Jennifer: I recommend the User Experience Professionals Association (UXPA) – International for more information. Check out uxpa.org for webinars, short courses, journal articles, and magazine articles. Full disclosure – I am President of UXPA International. We are always working hard to bring more resources to people and to connect folks in the field. Also feel free to reach out to us! We would love to continue the discussion: contact us at Jennifer Romano Bergstrom (Jennifer@romanocog.com), Emily Geisen (egeisen@rti.org) or Nick Parker (nkprkr@gmail.com).

13. WHAT JOB TITLES ARE TYPICALLY ASSOCIATED WITH THIS SPECIALTY?

Emily: I'm not sure. I primarily think of myself as a *Survey Methodologist*, but that's a pretty broad title and does not mean someone knows anything about usability. However, I also have the title of *Cognitive and Usability Lab Manager*. However, I don't love the "lab" part. It sounds like we're hooking people up to all sorts of equipment when we're not. Plus, I keep getting on mailing lists for catalogs selling beakers and test tubes! I might change it to *Cognitive and Usability Testing Manager*.

Jennifer: There are so many - *User Experience (UX) Researcher* is hot right now; *Research Psychologist* is used quite a bit, but unless you are a Psychologist with a PhD, I don't recommend using this title. *Usability Specialist* is pretty clear.