

The State of UX Research

Carol Barnum

Director, User Research

and Partner

UX Firm

Atlanta, GA

USA

carolbarnum@uxfirm.com

Introduction

I have been a UX researcher for 25 years. I did not come up through the usual degree programs available at the time, such as cognitive psychology and human factors. Rather, I came to the field from technical communication, seeing that there was a role for technical communicators to play in advocating for the user and promoting usability testing to understand the user experience, even if it meant conducting stealth testing on the documentation at the end of the product development cycle.

I am self-taught, but I learned from the best. My usability gurus were the authors of the first two usability handbooks by Dumas and Redish (1993, 1999) and Rubin (1994; with Chisnell, 2008). My mentor was a seasoned practitioner who owned a consulting company that did usability testing. We collaborated to teach a usability testing course for students in a graduate program in technical communication at my university. My side of the arrangement was to research the literature in the field. His side was to provide the methodology and the lab.

That collaboration led to a grant to fund my first lab in 1994. From that point to the time I left the university to start my own consulting firm six years ago, I have continued to expand my capabilities and knowledge in the field by reading, attending seminars and conferences, publishing articles, conducting hundreds of usability tests and other forms of UX research, and eventually writing two books on the subject.

What's the point of this personal history? It's to reflect on my path, the path of my students, and the changing landscape of today with the rapid spread of a myriad of educational and training opportunities and the rapid growth of tools and technology in the UX research space.



My reflections in this essay focus on the state of UX research. I am thinking specifically of usability testing, but also more generally of other common forms of UX research, such as contextual inquiry, heuristic evaluation, and in-depth interviews.

My key concerns can be presented in the following questions:

- Are short courses teaching rigorous usability research methods?
- Is the methodology for conducting valid small studies well understood?
- Is the democratization of UX research diluting the practice of UX research?
- Is the proliferation of online tools diminishing the utility of the UX researcher?
- With current job descriptions, should we drop the U in UX?

Are Short Courses Teaching Rigorous Usability Research Methods?

When I was seeking educational opportunities to learn about usability/UX, program offerings were limited. Not so today. Opportunities abound for those who want to get a formal education in UX, which often includes some exposure to UX research. Degree granting programs are available in traditional and flexible formats, including many online programs at both the undergraduate and graduate level. Many if not most of these programs include foundational courses on UX practice and research. A newer format program, the Center Centre (centercentre.com), is structured around course modules taught by industry experts with partnerships from companies like MailChimp. Students complete one course at a time over a 2-year period, receiving a diploma authorized by the Tennessee Higher Education Commission.

However, for many entering the profession today, faster paths to jobs can come from the growing number of training programs offered by companies like General Assembly, which conducts a 10-week program in User Experience Design and which includes usability testing in the description of the curriculum. Certificates and certifications are available through various organizations, including Nielsen Norman Group, Human Factors International, University of California San Diego Extension, Udemy Academy, and others. Short bootcamps are also available from universities like Georgia Tech and consulting companies like MeasuringU. For a partial list of certificate and certification programs, see usertesting.com/blog/ux-certification-programs.

In addition to, or instead of, enrolling in the growing number of programs, certificates, and certifications available, those interested in learning about UX research can avail themselves of daily articles on *Medium* (medium.com) and numerous blogs, podcasts, webinars, and YouTube videos available from companies and independent practitioners.

Knowledge is a good thing, and the widespread availability of this knowledge is both a reflection of the growth of the profession and a gateway to expand the pool of people interested in doing UX research.

Clearly, I can't speak directly to the quality and rigor of the curriculum for all of the available programs and course offerings. What I am left with is to speculate about the coverage of the fundamentals of conducting unbiased, effective research that is based on a well-designed research plan, a skilled moderator or interviewer, and a team or individual with the training to understand how to interpret the findings and act on them. Are we losing rigor in our UX research studies because newcomers to the field aren't learning these critical fundamentals?

Is the Methodology for Conducting Valid Small Studies Well Understood?

Regardless of the program or course offering, if it addresses UX research, it is likely to present a group of methodologies in a UX researcher's toolkit with an emphasis on not only how to effectively use each tool in the toolkit but also how to choose the correct tool or tools to meet project goals.

The program or course offering also likely addresses the value of small studies. Small studies work well in many development scenarios, but they work particularly well in Agile development methodologies. Because they can be done quickly and inexpensively, they are the method of choice for many practitioners. However, when small studies are conducted, is the original

research that formed the basis for the validity of small studies known and understood by those setting up the studies?

Do practitioners understand that the research conducted by Jakob Nielsen (1989), Robert Virzi (1990, 1992), Jakob Nielsen and Tom Landauer (1993), James Lewis (1993, 1994), and others in the late 1980s to mid-1990s proved the validity of small studies when certain specific conditions were in place? These included working with a defined set of tasks/scenarios and working with a specific subset of the user population who engage in thinking out loud.

When these conditions are in place, the researchers documented that a study with four to five participants will uncover 80–85% of the usability findings associated with the specific tasks in the study, and that additional users are less and less likely to experience new problems. This outcome results when the average likelihood of problem discovery is around 0.30, which will often be the case when developing new software or websites, less so when the low-hanging fruit has already been picked, based on findings from usability testing. If the likelihood of problem discovery is lower, more participants will be needed to achieve 80–85% discovery, and if it's higher, the problem discovery results will likely be higher than 85%. In any case, the process should not stop after a single study, but be repeated with a different set of users, tasks, or a new design.

Are the small studies being conducted today based on these guiding techniques?

What Happens if You Don't Apply these Techniques?

If you don't apply the techniques of testing with small numbers of participants, as espoused by Nielsen, Virzi, Lewis, and other researchers, are the results still valid? Will they hold up?

Jared Spool and Will Schroeder drew a lot of attention to the topic in their CHI presentation/paper (2001), stating that this approach does not work when testing websites. However, their study did not have a prescribed, duplicatable set of tasks. Instead, it asked participants to shop for something they were interested in, resulting in participants ranging far and wide within the four websites under study. Because there was so little overlap in where participants engaged with any website, new findings that prevented users from completing their purchase were still uncovered after 18 participants, leading Spool and Schroeder to conclude that it would take many more users to uncover 85% of the findings on any website.

In addition to the work of Nielsen, Virzi, and Lewis in breaking the "intimidation factor" for usability testing, we have Steve Krug to thank for bringing the subject to the masses in his two popular books: *Don't Make Me Think* (2000, 2005, 2014) and *Rocket Surgery Made Easy* (2010). These books have raised awareness of the value of usability testing to managers at all levels. As well, they have made the case for how easy it is to do—not rocket surgery. Krug's premise, like that of Nielsen, is that some testing is better than no testing.

Unlike Nielsen, however, Krug espouses DIY testing with the following approach to make it happen fast: (1) recruit participants based on the motto—"Recruit loosely and grade on a curve"; (2) test with only three participants because it takes little time and effort and still results in plenty of findings; (3) conduct a findings meeting immediately following the last test so that the most serious problems can be prioritized by the team; (4) "tweak" the design to make the least possible change to address the problems.

My concern with the premise of DIY testing is that the changes to the design are based on observing only three people who may not be "real users" and then having the team determine if the "fix" is acceptable without doing additional testing. These changes to promote fast fixes place too much reliance on non-users to determine what works for real users. Making it so easy to do DIY testing may allow practitioners and managers to check the box without validating the research.

What is the Goal of UX Research?

In my view, the goal of UX research is to identify or uncover *valid* findings, fix them, and iterate the process to uncover more. The validity of the findings comes from the validity of the research plan, the engagement of real users, and the skill of the moderator or interviewer in executing the plan.

I don't agree with those who believe that the goal of UX research is consistency in outcome, regardless of who is conducting the research.

Rolf Molich has demonstrated an apparent lack of consistency in his Comparative Usability Evaluation (CUE) studies, which currently number ten. He has raised the question as to whether there is a replicable process that can be employed to get consistent results regardless of who is conducting the UX research and has used the findings from these studies to suggest that such precision does not exist because of "the evaluator effect." (CUE Studies 1–9 are available at dialogdesign.dk/CUE.html.)

Having participated in four of Molich's CUE studies, including the current CUE-10 study of moderation practices, I hold the opinion that there need not be a replicable process. Some researchers may uncover more findings, others less. If the process is grounded in valid methods and executed without bias, the results are highly likely to be valid.

In an analysis of 33 studies that examined the evaluator effect, Jeff Sauro stated a similar view: "Agreement isn't necessarily the goal. . . . The goal of usability testing is ultimately a more usable experience. Diversity of perspective can be a strength" (Sauro, 2018, Summary & Takeaways section).

Is the Democratization of UX Research Diluting the Practice of UX Research?

If support continues to grow for good user experiences and the marketplace supports the growth of the profession, then everyone with an interest can become a UX researcher. Some have called this the democratization of UX research. It's a topic promoted by books such as Tomer Sharon's *It's Our Research* (2012) and Erica Hall's *Just Enough Research* (2013), which advocate for stakeholder engagement throughout the process.

Stakeholder engagement is also fundamental to the popular approach to design thinking as advocated by Google Ventures and presented in *Sprint: How to Solve Big Problems and Test New Ideas in just Five Days* (Knapp, 2016). This approach engages stakeholders across an organization in a five-day workshop, as opposed to the traditional approach modeled on the Stanford Design School and promoted by design company IDEO to engage designers or architects in a comprehensive design methodology process.

Nielsen Norman group promotes its own collaborative approach, calling it Design Thinking 101 (Gibbons, 2016). Their process employs six steps: (1) empathize, (2) define, (3) ideate, (4) prototype, (5) test, and (6) implement. Each phase is meant to be iterative, with a goal of repeating phases until the team believes the product is ready to launch.

The fact that this rapid, collaborative process begins and ends with a focus on the user is indicative of the current understanding of the importance of user experience in the design process. However, as much as this democratization of research seems a good thing in engaging all stakeholders in the process, I question whether this inclusive approach results in a dilution of the practice of UX.

Is the Proliferation of Online Tools Diminishing the Utility of the UX Researcher?

Online software tools, provided by UserTesting, UserZoom, Loop11, and other platforms, have made it easy for companies to get fast user feedback at relatively low cost. As a result, companies that have subscribed to these software platforms may no longer need any experienced UX researchers in house or in a consulting capacity. In the place of the researcher, the online platforms provide templates to screen participants and set up a quick study, with video recordings available the same day or soon after. Where these platforms were once providers of unmoderated remote usability testing, with some focused on qualitative small studies and others on quantitative large studies, they have, for the most part, expanded and blended services so that they all generally offer the same types of services, including live moderated remote sessions.

Whoever is using any of these platforms, the question is one of rigor: Who writes the research plan, who writes the participant screener, who analyzes the results, who determines the findings upon which recommendations are made? What level of experience do they bring to these tasks? Have the changes in training and education, the need for speed in getting user feedback in Agile development cycles, and the ready access to online tools to support fast

findings diminished the rigor of the process? For example, in usability testing, if the tasks are not well designed and the questions are potentially leading or misleading, will the results be valid?

The benefit of these online tools is that they spread the capability for anyone to conduct UX research. The risk is that they trivialize the place of UX researchers in the process. David Siegel and Susan Dray warn that

the proliferation of shortcut methods in UX practices . . . can undermine the professional status of the field. . . . Sadly, this is something UX folks are already starting to encounter in industry. For example, usability evaluation used to have relatively high status in organizations. It now has become routinized and deskilled to the point that there are companies offering automated usability testing. There are also usability jobs that require knowledge of many applications, with usability evaluation, user research, and/or other UX-type tasks taking a backseat. And these jobs sometimes do not require even a bachelor's degree. . . . For whatever reason, perhaps because UX is no longer a novelty, or because some perceive UX's contribution as marginal, or because we ourselves have promoted the idea that quick-and-dirty techniques are enough, our profession is at risk of being relegated to the corporate basement. (2019, Empathy Shortcuts section)

In addition to the threat to the profession described by Siegel and Dray, Gerry Gaffney warns against the

existential threat that may blind us . . . that at least some of the work traditionally done by UX practitioners could not be carried out by artificial intelligence (AI). . . . Combine a similar capability with automated remote unmoderated testing, and one can readily envisage an AI that not only designs user experiences, but can improve them through testing and by applying evolutionary algorithms to approach an optimal solution. (2017, p. 94)

With Current Job Descriptions, Should We Drop the U in UX?

Bill Gribbons (2017) addressed the issue of the diminished role of the UX researcher, advocating for dropping the U (user) in UX so that the focus is on Experience Design. This shift would allow us to expand the profession to encompass design thinking, customer experience, and other developments that may be dwarfing what we have traditionally seen as our space. Does this proposed name change imply that user research should take a back seat to design?

One experienced UX researcher, Mark Richman, wrote about this growing phenomenon of job postings that place a high priority on experience designers, with a lesser interest in capabilities for research (2018). Although Richman did find some job postings for researchers, he stated that these were the exception. He concluded that "Many of the ads I saw reserved their most specific and vivid language for the visual design end of things Are other companies just paying lip service to the current buzz by hiring visual designers and labelling them 'UX'?" (para 9). One UX researcher who responded to the article wrote, "I have been a UX practitioner for 9 years now and it recently took me nearly a year to get a new job—I discovered the visual interface design 'gotcha' was the problem. . . . Companies really want a VISUAL designer and don't understand that UX is NOT UI" (comment posted to Richman's article by Yvonniam).

If the trend is for one person to have a combined role as UX and visual designer, is that person assuming the role of UX researcher by default? And if so, are they doing their own user experience research on the products they are developing? Bill Albert (2015) warned that when this is the case, "the fox is guarding the usability lab." Can a designer construct and conduct an unbiased usability test of his or her own design? And can this same designer review and interpret the results without bias? Erin Friess (2011) reported on a study of novice designers that revealed explicit or latent bias present in their oral reports that did not match what study participants did or said. She noted that the designers used their oral reports to, in some cases, present findings out of context and in other cases omit findings that did not support their own pre-determined issues.

What I See as a UX Research Consultant

In my role as a UX research consultant, I have seen much of my business shift away from requests for thoughtful, well planned research studies to increasingly fast, informal studies. I have also experienced the impact of companies adopting subscriptions to the online tools in place of my services.

I am willing and able to adapt to the need for faster research, faster findings, less formal methods. I have supported the use of online tools for clients by conducting the analysis of video recordings from unmoderated remote sessions and reporting findings and recommendations. In some cases, when clients are using these online tools, I have offered to write screener questions and tasks/scenarios. My expertise has occasionally been used, but more often it has not. My concern is that potential loss of rigor may result from inexperienced or inadequately trained practitioners setting up studies and analyzing the results.

Counter to this trend is the work I do in the field of medical device human factors validation testing for FDA approval of products coming to market. Following the release of the FDA guideline *Applying Human Factors and Usability Engineering to Medical Devices* (U.S. Department of Health and Human Services, Food and Drug Administration, 2016), companies now understand the requirement to conduct testing with a minimum of 15 participants in each user group. This results in big studies supported by detailed reports and the associated big budgets.

While my work as a UX researcher engaged in big studies for medical devices is interesting and rewarding, I still love small, qualitative, iterative studies, so my concerns are in this area of UX research.

I know I am not alone in airing these concerns, as the mention of several sources in this article, all by seasoned researchers, indicates. I welcome feedback from others in our community.

Acknowledgments

Special thanks to Ginny Redish, Steve Krug, and Jim Lewis for providing excellent suggestions for making this essay stronger and more coherent.

References

- Albert, B. (2015). The fox guarding the usability lab. *Journal of Usability Studies*, 10(3), 96–99.
- Dumas, J. S., & Redish, J. C. (1993). *A practical guide to usability testing*. Norwood, NJ: Ablex.
- Dumas, J. S., & Redish, J. C. (1999). *A practical guide to usability testing* (revised ed.). Bristol, United Kingdom: Intellect.
- Friess, E. (2011). Discourse variations between usability tests and usability reports. *Journal of Usability Studies*, 6(3), 102–116.
- Gaffney, G. (2017). The revolution will not be handheld. *Journal of Usability Studies*, 12(3), 92–94.
- Gibbons, S. (2016, July 31). Design thinking 101. Retrieved from <https://www.nngroup.com/articles/design-thinking/>
- Gibbons, B. (2017). Is it time to drop the “U” (from UX)? *Journal of Usability Studies*, 13(1), 1–4.
- Hall, E. (2013). *Just enough research*. New York, NY: A Book Apart.
- Knapp, J. (with Zeratsky, J., & Kowitz, B.; 2016). *Sprint: How to solve big problems and test new ideas in just five days*. New York, NY: Simon & Schuster.
- Krug, S. (2000, 2005, 2014). *Don't make me think: A common sense approach to web usability* (1st, 2nd, 3rd eds.). Berkeley, CA: New Riders.
- Krug, S. (2010). *Rocket surgery made easy: The do-it-yourself guide to finding and fixing usability problems*. Berkeley, CA: New Riders.

- Lewis, J. R. (1993). Problem discover in usability studies: A model based on the binomial probability formula. *Proceedings of the Fifth International Conference on Human-Computer Interaction* (pp. 666–671). Orlando, FL, USA: Elsevier.
- Lewis, J. R. (1994). Sample sizes for usability studies: Additional considerations. *Human Factors*, 36(2), 368–378.
- Nielsen, J. (1989). Usability engineering at a discount. In G. Salvendy & M. J. Smith (Eds.), *Proceedings of the Third International Conference on Human-Computer Interaction on Designing and Using Human-Computer Interfaces and Knowledge Based Systems (2nd ed.)*; 394–401). New York, NY: Elsevier.
- Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. *Proceedings of ACM INTERCHI'93 Conference* (pp. 206–213). Amsterdam, The Netherlands: ACM.
- Richman, M. (2018, March 6). Are we taking the “U” out of UX? Retrieved from <http://boxesandarrows.com/are-we-taking-the-u-out-of-ux/>
- Rubin, J. (1994) *Handbook of usability testing: How to plan, design, and conduct effective tests*. New York, NY: Wiley.
- Rubin, J., & Chisnell, D. (2008) *Handbook of usability testing: How to plan, design, and conduct effective tests (2nd ed.)* Indianapolis, IN, USA: Wiley.
- Sauro, J. (2018, March 27). How large is the evaluator effect in usability testing? Retrieved from <https://measuringu.com/evaluator-effect/>
- Sharon, T. (2012). *It's our research: Getting stakeholder buy-in for user experience research projects*. Waltham, MA: Morgan Kaufmann.
- Siegel, D., & Dray, S. (2019, February). The map is not the territory: Empathy in design. *Interactions*(26), 82–85. Retrieved from <http://interactions.acm.org/archive/view/march-april-2019/the-map-is-not-the-territory>
- Spool, J., & Schroeder, W. (2001). Testing web sites: Five users is nowhere near enough. *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, (pp. 285–286). Seattle, WA, USA: ACM.
- Virzi, R. A. (1990). Streamlining the design process: Running fewer subjects. *Proceedings of the Human Factors Society 34th Annual Meeting: Vol. 1* (pp.291–294). Orlando, FL, USA: HFES.
- Virzi, R. A. (1992). Refining the test phase of usability evaluation: How many subjects is enough? *Human Factors*, 34(4), 457–468.
- U.S. Department of Health and Human Services, Food and Drug Administration, Center for Devices and Radiological Health, Office of Device Evaluation. (2016, February 3). *Applying human factors and usability engineering to Medical Devices: Guidance for industry and food and drug administration staff*. Retrieved from <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/applying-human-factors-and-usability-engineering-medical-devices>

About the Author



Carol Barnum

Dr. Barnum is Director, User Research and Founding Partner of UX Firm based in Atlanta, GA. She is the author of *Usability Testing Essentials: Ready, Set...Test!*