Usability Testing as Part of a Broader System of UX Design:

A Literature Review on Modes of Inquiry that ‘Dwell in Possibility’

Liv Riggins

University of Minnesota-Twin Cities
Abstract

This paper examines the recent history of usability testing and the evolution of user experience design, in hopes of providing background and grounding for further research in user experience design for an online vocal health intervention. In the course of examining this history, I engage with debates on usability testing methodology and summarize findings on the emerging user experience subtopic of design for engagement. By engaging with debates on openness and control in usability testing, I hope to show how usability testing that is adaptable, open, and non-replicable can nonetheless provide valuable information, especially when conceived of as part of a broader ecosystem of interacting research methods. This understanding of usability testing will inform test design and research methods for results that enable practitioners to recommend design changes that enhance user experience and increase motivation and engagement, rather than simply removing errors.
Usability testing for online interfaces emerged in the 1980s from activities taking place across a combination of fields, namely human factors, software development, and technical communications (Redish, 2010). The increasing prevalence of computer use among a general population previously generally unexperienced with computers led to a need for documentation, design, and study of how computers and computer documentation could be developed to allow people with minimal training to quickly grasp and accomplish work tasks (Sullivan, 2018). The work of researchers including Johnson, Dumas, and Sullivan describes early understandings of usability (1960s-early 1980s), particularly in the types of usability testing more closely connected to human factors, as focusing on developing products and systems that maximized productivity and efficiency, controlling the user’s possible actions to reduce error and uncertainty (Johnson, 1994; Dumas, 2007; Sullivan, 2018). The user could be seen as an “unfortunate human factor,” an interference in the smooth functioning of a system (Johnson, 1994).

User Centered Design

However, in the early 1980s, a key turn occurred in the conceptualization and practice of usability research and testing that reoriented study of usability around pursuing technological development molded to the service of human goals and desire, rather than seeking to control and mold human action to the requirements of a technological system—engineering products and
systems for humans, rather than seeking to mold and control human actions to fit a given product or system (Johnson, 1994). Don Norman’s work in exploring, describing, practicing, and advocating for user-centered design was a crucial part of reorienting the design and testing process around intuitive understanding and ease of use as primary goals. In Norman’s original formulation of user-centered design, the term is used to describe a design process in which user needs take primacy, to the extent of dismissing ‘secondary’ concerns such as aesthetics, pleasure, and emotion (Norman, 2013).

While less encompassing of the user experience and motivations than usability would later become, this change in conception was a “reorientation of technological development around people rather than artifacts” (Johnson, Meredith, & Salvo, 2007, n.p.). It was expanded later in the decade through a reconceptualization of the computer “as a communication tool, rather than a computation device,” leading to an understanding of usability research as the study of the increasingly social, networked, and interactional exchanges taking place between people via computer systems (Johnson et al., 2007, n.p.). These shifts together changed the goal of usability research from creating a highly controlled process that maximized productivity and minimized ‘error’ to examining interaction, communication, and human goals in service of optimal human experience of a technological tool.

**User Experience in Usability Testing**

In addition to enabling ease in completing actions—and designing systems for actions users actually want to complete—conceptions of usability have been expanded to include analyzing and designing the experience of interacting with a particular product or system. This shift was, in many cases, informed by the gaps in action and error-focused usability testing
USABILITY TESTING AS PART OF A BROADER SYSTEM OF UX DESIGN

(termed “post-modern” testing in Lund (2006)). User-centered usability testing was largely assimilated to software design processes, but the resulting software did not necessarily become popular or widely embraced, and could even be a commercial failure (Lund, 2006). Despite an intuitive interface and high ease-of-use scores, processes of usability testing that treated it as a “debugging” of the user interface, analogous to debugging the software, resulted in technologies that could be entirely usable, but not necessarily useful, compelling, or inviting of interaction.

Concurrently, and perhaps in response to the results of error-focused usability testing as it was being practiced, Norman reconsidered his characterization of aesthetics and emotion as a secondary factor in usability and design, emphasizing the crucial role that aesthetics, pleasure, and emotion play in experience and use of designs in his 2003 *Emotional Design*. Here, he eschews the term user-centered design in favor of human-centered design, which he redefines as “an approach that puts human needs, capabilities, and behavior first, then designs to accommodate those needs, capabilities, and ways of behaving” (Norman, 2013, p.8). Experience of the gaps in this more user-centered, but nonetheless utilitarian and limited definition of usability testing led to the realization that “ease of use is only important in the extent to which it enables users to experience more of the value that matters to them” (Lund, 2006, p.5). The “post-modern” concept of usability left out crucial elements, now included in what Lund terms “post-post-modern” usability, and Norman terms human-centered design: human satisfaction, enrichment, values, and contextuality.

Contrasting with the task- and interface-focused set of heuristics that Norman developed, Cheryl Gesiler articulates a set of heuristics and design principles for user engagement that illustrates the differences between a task-focused and more comprehensive, experience-focused
The emerging field of experience architecture builds on this shift towards the importance of researching, examining, and designing for a comprehensive experience, and further emphasizes the iterative, interactional, and contextual understanding of user experience beginning to be articulated in ‘post-post-modern’ usability and human-centered design. In Potts and Salvo’s *Rhetoric and Experience Architecture* (2018), usability testing is characterized as one research tool among many others, including content strategy, app development, project management, interaction design, findability, and web development. Usability’s placement within the broad “umbrella” of experience architecture reflects an understanding of it as “not a single answer,” but one part of “a diverse range of solutions that together show a rich collection of potential trajectories” (Potts & Salvo, 2018, p.9). Today, usability is often grouped under or in relation to user experience research, which seeks to address many of the same goals as usability, but in a broader sense that takes account of the increasing contextual and interrelated nature of designing and delivering content experiences. In the richly intertextual current online environment, context and interaction with other systems must be considered: Potts and Salvo state that “Gone is the
moment in which we thought we could build for a simple singular task, if we ever could. Here we must understand the context in which our participants are engaging with these experiences” (Potts & Salvo, 2018, p.5). This reconceptualization of usability not as an answer by itself, but as part of a series of tools and conceptual frameworks that may be brought to bear on a particular situation or product, reflects increasing understanding of the complexity of designing for user experiences.

Along with recognition of the interrelations and complexity between the different disciplines that contribute to creating an experience of content or a product, this framing of usability within experience architecture calls for increased flexibility, adaptability, and openness—what Johnson and colleagues call “dwelling in possibility” (Johnson et. al, 2007, n.p.). This concept of usability emphasizes its liminal qualities and ability to make meaning by straddling conflicting boundaries, namely “the space between the known and unknown,” and “both problems and potential.” Here, usability research involves to analyzing the problems that exist, then not only envisioning technological solutions but also consider the context in which they occur, the rhetorical framing, and the possible ways to rethink how the entire system is framed.

This understanding of usability’s strength in not only seeing what is happening, but also examining why it is happening, how it is, and proposing what else could be happening, reinforces Sullivan’s 1989 call for usability research to become “a broad and culturally-aware field of inquiry, built around the challenges of understanding and improving the relationship between humans and our technologies” (Sullivan, 1989). By adopting a mode of inquiry and embracing the power to sit across boundaries of known and unknown, existing and
Motivation and Engagement in Experience Architecture

In the broad, experience-focused, boundary-crossing understanding of a usability testing that “dwells in possibility,” analyzing not just what is happening but collecting new insights on why it is and what else may be, concepts and practices from communication theory become more central, shifting the focus from mechanical insights to communication and psychological insights.

The emphasis in experience architecture and user experience on “gathering stories” on user motivations and habits involves analysis of symbolism and production of meaning, areas in common with communication theory and semiotics. Andrew and Miriam Mara pursue “a triple focus on individual motivation, connections, and interactions” in their UX research on an app-based art scavenger hunt (Mara & Mara, 2018, p.189). To understand what drives individual motivation and connection and consequently how to generate these experiences through the scavenger hunt and app design, Mara and Mara (2018) foreground writing and observation in their methodology, drawing on ethnography’s concepts and methods. They then analyze their observations through inquiry into how meaning is produced for app users, which leads them to consider what symbols and functions are meaningful for app users, why they are, and how they produce meaning that then drives engagement. In their discussion, they parse user motivations into a few major themes, similar to Christiansen and Howard’s (2018) more structured practice
of developing *habitus* rather than personas to summarize user motivations and inform design decisions through understanding motivations and modes, rather than demographics or characteristics alone.

Based on their groupings of user motivations, Mara and Mara seek to drive engagement through building an app with symbols and narratives that mirror user motivations. They seek to engage the motivations they’ve parsed through the scavenger hunt and map design, including “creating meaningful pauses in the event, connecting participants with the narrative thread of the event (a progression through twenty-six points), and with connecting narratives that participants want to build into their lives” (Mara & Mara, 2018, p. 194). Especially in this description of their recommendations for the app, Mara and Mara foreground their story-focused, communications-informed approach to experience architecture, emphasizing that experiences are built through storytelling and socializing, and so designers must take interest in what stories are being told, how they are being told, and their effects.

In situations where the focus of the experience is not task completion or work so much as learning, enjoyment, or exploration, such as Mara and Mara’s scavenger hunt and potentially the health education module discussed here, this focus on engagement, storytelling, and meaning production becomes central to defining what practitioners seek to study and engage through research.

**Designing for Engagement in Online Health Intervention Applications**

This focus not only on experience, but also on the engagement and motivation aspects of user experience, is especially relevant in online health intervention and education applications. While there certainly may be specific tasks users need to complete to accomplish their goals, the
USABILITY TESTING AS PART OF A BROADER SYSTEM OF UX DESIGN

The purpose of these types of applications is education or behavioral change in the service of better health outcomes, not efficient completion of assigned tasks. As such, analysis on these types of applications focuses on factors of user experience that have to do with satisfaction, engagement, and motivation, all of which are more crucial factors in helping users achieve their goal of better health through self-tracking, behavioral changes, reflection, and education. These are especially pertinent to online health interventions for patients with chronic disorders, where motivation and engagement may need to be sustained over time to have the desired impact (Kostkova, 2015).

Because of these unique goals and characteristics of designing for online health interventions, the concept of designing for user engagement becomes particularly relevant. User engagement has variously been defined as “the state of mind that we must maintain in order to enjoy a representation of an action” (Laurel, 1993, pp. 112-113), a dimension of usability (Quesenbery, 2003), and a category of user experience (O’Brien, 2008). Here, I will view it “within the context of an overall interactive user experience with an emphasis on what compels people to become engaged and sustain their use of a technology,” drawing on the definitions put forth by Jones (1998), Sutcliffe (2010), and O’Brien (2016).

Based on previous studies on designing for user engagement in self-directed health interventions for chronic disease, Turchioe and colleagues (2019) grouped the potential features that promote this type of engagement into six categories, which I will also use to organize discussion of potential features and modalities for enhancing engagement: external incentives (Spring et. al, 2018; Niendam et. al, 2018; Hales et. al, 2017), health expert feedback (Clarke, 2014), social connectedness (Bisafar, Ponnada, Shamekhi, & Parker, 2017), goal setting, self discovery (Korinek et. al, 2018; Mamyinka, Smaldone, & Bakken, 2015) and decision support.
USABILITY TESTING AS PART OF A BROADER SYSTEM OF UX DESIGN

(Hollis et. al, 2017; Sas et. al, 2014). Underlying all of these features is the element of tailoring and personalization, which Lustria et. al (2009, 2013) identified through meta-analysis of studies on web-based health interventions as significant factors in achieving desired health outcomes through online education and intervention.

Balancing Control and Openness in Test Design

Unlike many fields of study, usability does not have a centuries-long history that can be used to define it, resulting in varying definitions, conceptions, and methods that provide a lively and inquisitive field, but also leave it open to charges of lack of rigor and credibility. The mutability and unpredictability in usability testing has drawn criticism as being a methodologically immature, and therefore less valuable, form of research. In the qualitative approach to knowledge production typical of the scientific method, variability and lack of replicability mean that findings are not true, and cannot be considered to contribute to the bulk of accepted knowledge in a field. In the comparative usability evaluation (CUE) studies, Molich and Dumas (1998) attempted to test the legitimacy of usability methods and, consequently, the validity of findings from usability testing by testing the replicability of test findings. Molich asked usability research team to conduct their standard test procedure with the same research questions and data, then compared recommendations across the different teams’ reports. They found that the results differed radically across teams—no two teams reported the same problem (Molich & Dumas, 1998). From this, Molich and other usability practitioners in favor of increased control in test design have concluded that the teams’ failure to replicate was “convincing evidence of the field’s immature and erratic approach” (Christiansen & Howard,
This critique of a lack of replicability, rigor, and discipline conventions has had a particular impact because of the newness of the field of usability.

However, other usability researchers and practitioners propose that systemicity, rigor, and replicability are not relevant standards for usability testing. Sullivan’s (2018) *Beckon, Counter, Experience* focuses on addressing the nature of usability research and proposing the best balance in control and openness in usability research. Though usability testing often collects quantitative data, Sullivan positions usability testing as a fundamentally qualitative type of research based in observation, most similar to ethnography. Instead, Sullivan maintains that the insight and unique capabilities of usability and user experience research come in their ability to solicit unexpected encounters and therefore produce new knowledge. This type of usability testing has methods that include “gathering stories that we otherwise would miss or to discovering themes, variables, and research questions in areas that are mysterious, new, or under-investigated,” an endeavor that is necessarily unpredictable and non-replicable (Sulllivan, 2018, p. 30). Other researchers and practitioners have similarly highlighted the surprising insights, perspectives, and technological workarounds that users bring (Latour & Woolgar, 1979; Vilbergsdottir, Hvannberg, & Law, 2014; Spinuzzi, 2003). The purpose of bringing in users to do the testing is that they have insights a usability practitioner cannot foresee; making the testing process replicable and predictable therefore defeats the purpose of bringing in users.

Giving up control, having openness to contingency, and creating conditions for a wide range of user encounters goes against the grain of what researchers are taught to do and how quantitative research is conceptualized, yet it is crucial for gaining new insights to a system and its potential issues and improvements.
It is for this reason that Sullivan calls for “supple, plastic methods that do not become blind in their rigidity” (Potts & Salvo, 2018, p.9). Adopting this position, of flexibility, adaptability, and openness is an act of respect for the knowledge of the user, who is ostensibly at the center of user experience design, and is also crucial to creating valuable, new knowledge through usability research.

The work of Sullivan, Walter, and Johnson further indicates that, while there should certainly be a balance between structure/control and openness in usability research, there is already sufficient focus on control. This is a result of Molinch and others’ critiques of the lack of replicability and rigor in usability testing methods and that openness provides more potential for unique findings. Most importantly, the user-centered design model that usability operates on is already more controlled than other design models, notably participatory design, essentially baking control into the research and testing process. Drawing on Ritter, Baxter and Churchill (2014), Rose et. al (2017) characterizes the tradition of user-centered design that usability research operates within as a practice with North American origins that operates on a representational model of participation, where users engage with research but are not incorporated in design decisions or included on the design team. In contrast, approaches rooted in participatory design draw on the Scandinavian design tradition for a deeply user-involved process, where “participation is conceived as full engagement by members of the community throughout the entire design process” (Rose et al., 2017, n,p). Echoing Sullivan’s argument on the valuableness of unexpected and unpredictable insights through fuller user engagement, Walton, Zraly, & Mugengana, (2015) note that a participatory design framework can have significant benefits in exposing unexpected and innovative issues, solutions, or improvements
that researchers would not have thought of without direct user engagement: “well-designed, well-conducted community-based research encounters unexpected challenges and serendipitous surprises because power is not centralized with researchers” (Walton et al., 2015, p. 62).

Rose’s work in usability and user experience research similarly emphasizes the importance of flexibility and adaptability in methods and use of multiple testing methods at a time, especially when working with highly-specific user populations. In an article on usability testing conducted in two Ghanian villages, Gorman, T., Rose, E., Yaaqoubi, J., Bayor, A., & Kolko, B. (2011) establish the importance of partnering with local staff or organizations to conduct testing and recruitment, the challenges of doing so, and the necessity of flexibility and adaptation in methods and procedure based on user population and setting, as do Putnam, Johnson, Rose, and Kolko (2009) in a study on adapting user-centered design methods for mobile technology in Kyrgyzstan.

Rose et al. (2017) discussed the challenges and requirements for connecting with a highly specific user community—in the case of that project, immigrant patients without health insurance. The article emphasizes the need for close collaboration and shared decision making with a community partner organization in test design, research questions, methods, results, and recruiting, as well as flexibility and adaptability in response to the needs and goals of the partner organization and participants alike. Flexibility, adaptability, and collaboration are crucial in these cases because usability testing methods may not have been designed with these users, community partners, and their unique situations and needs in mind, and so must be changed or adapted as the situation and collaborators require. While these studies all took place in a different
contexts and with different user populations, the call for partnership with organizations closer to the users and flexibility and adaptation of testing methods remains the same across these studies.

To resolve some of the tension between the need for mutability and need for standards and control, researchers and practitioners across the field of usability testing encourage use of multiple accepted methods in one test. For example, Racadio, Rose, and Boyd (2012) recommend combining field tests with lab studies to “leverage recruiting strengths” and reap benefits of observation in different contexts. This reinforces the understanding of usability testing as one part of a larger toolkit established by experience architecture, and also responds to the criticism of usability testing as insufficiently rigorous to consistently catch errors and issues. Molich calls practitioners to “realize that there is no foolproof way to identify usability flaws,” “place less focus on finding ‘all’ problems,” “realize that single tests are not comprehensive,” and focus on quality of the test procedure (Molich & Dumas, 1998, p. 74).

These recommendations constituted a decentering of usability testing as a comprehensive method in itself, complimenting recommendations from Sullivan (1989) and Redish (2010) to embrace a broader understanding of usability testing that understands it as one potential tool, that can be combined with other test methods and research methods to gain a more accurate picture of the user’s experience. Reason’s Swiss cheese model of risk management aligns with these recommendations to use multiple ways of usability testing; the holes in the procedure must be made smaller and more layers of verification must be added to reduce the chance that the ‘holes’ in a procedure align and allow a significant issues to slip through. By adopting this idea of layering, and broadening the conception of usability research beyond testing alone, researchers and practitioners will be able to catch more issues and gain a more qualitative understanding of
how users interact with and experience an application, leading to better information to inform design decisions.

**Conclusion**

In discussing the evolution of usability testing, its expansion to include focuses on experience and contextuality, and the role of motivation and engagement in creating meaningful patient experiences, I seek to provide a grounding background for research in user engagement design, particularly in patient communities. The developments in usability testing that I have described led to an articulation of usability as an iterative, interactional, user-centered, adaptable, action-oriented, and contextual field of inquiry. By engaging with debates on openness and control in usability testing, I hope to show how usability testing that is adaptable, open, and non-replicable can still have valuable information, especially when conceived of as part of a broader ecosystem of interacting research methods, as no one test or method will capture all the information that is relevant to the study of user experience.
References


[https://doi.org/10.3389/fpubh.2015.00134](https://doi.org/10.3389/fpubh.2015.00134).


USABILITY TESTING AS PART OF A BROADER SYSTEM OF UX DESIGN

https://doi.org/10.1016/J.JPSYCHIRES.2017.10.017.


USABILITY TESTING AS PART OF A BROADER SYSTEM OF UX DESIGN


USABILITY TESTING AS PART OF A BROADER SYSTEM OF UX DESIGN


