User Experience Methods in Research and Practice

Abstract

User experience (UX) researchers in technical communication and beyond still need a clear picture of the methods used to measure and evaluate UX. This article charts current UX methods through a systematic literature review of recent publications (2016-2018) and a survey of 52 UX practitioners in academia and industry. Our results indicate that contemporary UX research favors mixed methods, and that usability testing is especially popular in both published research and our survey results. Other methods, such as surveys and ethnography, get widely used in published research but prompt varied reactions from survey respondents. This article presents these findings as a snapshot of contemporary research methods for user experience.

Keywords: User Experience, Methods, Literature Review, Survey Research

Introduction

This special issue asks scholars in technical communication to take a holistic look at what constitutes user experience (UX), both within technical communication and more broadly in the loose coalition of fields that comprise all of user experience. Among the many questions posed by the special issue is one on methods: "What methods do we use and approaches do we take to research usability and user experience design within the field?" The question has long plagued UX, as researchers have struggled to outline the parameters and consistent practices of UX (Scapin, Senach, Trousse, & Pallot, 2012; Forlizzi & Battarbee, 2004; Kaasinen et. al., 2015) and articulate the “multi-disciplinary needs of the user experience industry,” (Instone, 2005, p. 1). In 2010, Vermeeren et. al. argued that “A multitude of methods for UX design and evaluation exist, but a clear overview of the current state of the available UX evaluation methods is missing” (p. 521). Progress on this project is steady but incomplete. In 2006, Hazzenzahl and Tractinsky outlined a productive, forward-looking research agenda for UX, but in 2018, Hazzenzahl argued that “we are far from having a coherent understanding of what user experience actually is” (p. 301).

In theory, technical communication scholars should be well-poised to answer this question about user experience methods, given the prominent overlap between tech comm and UX goals. As Redish and Barnum (2011) write, “Technical communicators are by training and necessity user-centered. Their focus is always the audience, the people who will use whatever they are creating. Their goal is to make even complex interactions understandable and usable” (p. 91). In 2016, Lauer and Brumberger argued that “Developing information products—e.g., instructional
documents—for specific users has long been the core of technical communication work” (p. 249). Technical communicators have long used their methods to improve user experience, whether they considered their work through that terminology or not. The field has also offered its own definitions of UX, such as Williams’ technical communication-influenced definition that addresses “people's emotional, psychological, physical, and social needs, as well as their

cognitive ones and thereby redefining a 'good' text as one that embraces the complexity of humanity" (p. 12).

We know that technical communication scholars employ a variety of methods to measure usability and user experience, including usability testing (Kastman Breuch, Zachary, & Spinuzzi, 2001), think aloud protocols (Cooke, 2010), and surveys, interviews, and observations (Sullivan, 1989). Yet we have been slow to systematically map user experience methods within and beyond technical communication. In her discipline-defining mapping of research questions, Rude (2009) mentions user experience only once. The introduction to a special issue on contemporary research methodologies in technical communication likewise only mentions user experience once, in passing (McNely, Spinuzzi, & Teston, 2015). One of the most exhaustive reviews of empirical technical communication scholarship (Meloncon & St. Amant, 2018) does not specifically mention user experience but does find that usability testing in technical communication research is a troublesome category. Meloncon and St. Amant write that "the usability studies in this data set clearly show that TPC researchers approach usability testing in varied and diverse ways" (p. 16). Authors claiming to conduct usability tests do not always provide a full description of their methods or a rationale for why those methods were chosen and how they contribute to knowledge-building. According to Meloncon and St. Amant, the field needs empirical research which provides "a clear connection between the research question, the methods chosen, and the data collected. This is the key to sustainable research for TPC" (p. 17). Other technical communication texts provide little guidance on usability methods; as recently as 2016, Chong found that technical communication textbooks, foundational texts, and course syllabi devoted little attention to issues of usability, let alone user experience.

This article addresses the need for a more sustainable, robust understanding of UX methods by charting the methods, artifacts studied, and research questions (or lack thereof) in 218 empirical studies about UX published between 2016-2018. Our goal is to catalog the very recent methods employed by a variety of UX researchers near the time of this publication to provide a contemporary look at the field. Our work updates literature reviews about user experience methods from Bargas-Avila and Hornbaek (2012) and Robinson, Lanius, and Weber (2017). However, we also present data from a survey of 52 user experience practitioners as a way to compare our literature review results with the working experiences of UX professionals, both in and beyond academia. Taken together, our findings suggest that UX research prioritizes mixed methods. While stalwart methods like usability testing, surveys, interviews, and ethnography dominate published research, our survey respondents favored a variety of methods, including expert reviews, data capture, and biometrics, to measure the user experience. This article builds on past research that attempts to catalog UX methods through literature reviews and surveys, presents the findings of our own literature review and survey, and then discusses the relevance of these findings to both the UX discipline broadly conceived and technical communication specifically.

Literature Review

Scholars have performed several literature reviews to chart the trends in UX research throughout the 21st century. Scapin, Senach, Trousse, and Pallot (2012) catalog methods in several UX studies and find that the field is incredibly multidisciplinary and variable, leading them to conclude that with the need for "some converging, non-polysemous, agreed upon definitions of UX that cover the various domains and territories that are involved" (p. 5). Maia and Furtado (2016) reviewed 25 UX studies and determined that these papers favored qualitative methods to analyze ready-made products. Gross and Bongartz (2012) chart the development of UX research in three-year intervals between 2000 and 2010. González-Pérez, Ramírez-Montoya, & García-Peñalvo (2018) performed a literature review of studies implementing user-centered design (UCD) in repositories "with the purpose of emphasizing the need to develop a framework of best practices of UCD" (p. 84). Yusop, Grundy, and Vasa (2017) reviewed 57 usability studies and identified a number of methodological issues, including inconsistent terms and mixed data. Urgas et. al. (2016) reviewed 109 studies of website usability and found that most studies employed questionnaires and usability testing. Nie and Sun (2017) performed text analysis and data analysis on a huge sample of literature published between 2004-2015 and found tremendous growth in studies covering interaction design, which encompasses user experience. Bargas-Avila and Hornbaek (2012) analyzed 51 studies (containing 66 experiments) published between 2005-2009 and concluded that questionnaires were the most common research method and that 50% of studies were qualitative, 33% were quantitative, and 17% involved mixed methods. Robinson, Lanius, and Weber (2017) analyzed 431 UX studies from 2000-2016 and found a steady increase in qualitative and mixed methods studies in the 2010s. Robinson and Lanius (2018) analyze specifically how geography and discipline affect UX publications. Other literature reviews have examined UX-related trends from research in specific disciplines, including HCI (Hooper and Dix, 2012), research involving particular technologies, such as mobile devices and applications (Ismail, Ahmad, Kamaruddin, & Ibrahim, 2016; Hao, Chong, Man, Liu, & Shi, 2016; de Sa & Carriço, 2011; Kjeldskov & Graham, 2003), VR and AR (Kim, 2012), ATMs (Mkpojiogu & Asuquo, 2018), and smart homes (Kim, Oh, Cho, Lee, & Kim, 2013), and research trends in particular cultures (Kurosu et. al., 2004).

Other scholars have analyzed UX through surveys and interviews with practitioners (MacDonald, 2015; Law, Schiak, & Roto, 2014; Dove, Halskov, Forlizzi, & Zimmerman, 2017; Chilana et. al., 2011; Mao, Vredenburg, Smith, and Carey, 2005). A survey by Vredenburg, Mao, Smith, and Carey (2002) found "a major discrepancy between the commonly cited measures and the actually applied ones" (p. 471). Ibargoyen, Szotak, and Bojic (2013) asked 77 experience professionals to define their discipline, and found that "even seasoned professionals often lacked the vocabulary to describe a discipline in which they partake on a daily basis" (p. 2087). Lallemand, Gronier, and Koenig (2015) surveyed 758 UX practitioners and researchers about the perception of statements and definitions about UX, in order to expand many of the findings from a survey of UX practitioners by Law et. al. (2009). Gerea and Herskovic (2015) also replicated the Law et. al. study with UX professionals in Latin America, while Rajanen et. al. (2017) found in a cross-cultural analysis that "UX professionals diverge when defining UX" (p. 218). Jääskeläinen and Heikkinen (2010) conducted a survey which found huge differences in

the definitions of UX between practitioners and end users. Salgado, Amaral, Freire, and Fortes (2016) surveyed 26 Brazilian business about their UX practices and found that usability tests and heuristic analysis were the most common methods for understanding user needs. In a survey of 59 gaming companies, Rajanen and Tapani (2018) elicited definitions of game usability and the methods and artifacts used to study it. Gray (2016) interviewed 13 UX practitioners about their methods and found that they primarily mentioned personas, interviews, and focus groups as consistent methodologies. UX researchers needs both "core set of design methods" (p. 4052) alongside more soft skills to negotiate complex, collaborative environments.

To continue the scholarly exploration of UX methods, this article expands the literature review approach used in studies like Bargas-Avila and Hornbaek (2012) and Robinson, Lanius, and Weber (2017), but we also include survey results from UX academics and practitioners. Studies that combine literature reviews and empirical data about UX professionals are rare, though Lachner et. al. (2016) combined a review of 30 publications and interviews with 24 UX experts to develop a tool that supports "a common organizational understanding of a product's UX and the selection of further in-depth UX evaluations" (p. 1). Our study also owes a debt to Vermeeren et. al. (2010), which collected UX methods through a literature search, a focus group, the authors' professional experience, and a survey. Like these authors, we believe that combining a literature review and survey data allows us to compare the perspectives and methods of academics and other practitioners, and provides a more holistic look at the current state of the discipline.

This article answers the following research questions:

RQ1. What UX methods do user experience professionals prioritize in their work?
RQ2. How do researcher's attributes influence their choice of UX methods?
RQ3. How do methods differ between user evaluation and user research?

Methods

This study was conducted using a two pronged approach to update an extant systematic literature review previously published and to consult user experience experts using a survey instrument.

Systematic Literature Review

The systematic review was conducted on data from 2016-2018 collected for articles that referenced UX or user experience from a variety of publications and fields. The process of collecting the data was accomplished in the phases described below. First, following practices established by a previous systematic review study (Robinson, Lanius, and Weber, 2017), we identified peer-reviewed publications that met our criteria by using Google Scholar (GS). According to Gehanno, Rollin, and Darmoni (2013), "the coverage of GS is much higher than previously thought for high quality studies. GS is highly sensitive, easy to search and could be the first choice for systematic reviews or meta-analysis. It could even be used alone" (p. 4).

We performed the search through a function in the reference management application Papers 3. In order to find studies, we used search terms "user experience" and “UX” found in the abstract or title. We hoped that this configuration would 1) limit false positives and 2) ensure we did not miss relevant publications. This search schema retrieved 273 publications and resulted in 218 publications to analyze after they were screened for language and empiricism. We coded the articles in Google Sheets using a set of codes developed in Robinson, Lanius, and Weber (2017) to identify the method, research questions, artifacts of study, and other features of the articles not reported below, such as the sample sizes of users tested. Coding required some interpretation for the method because of inconsistencies about the vocabulary for methodologies (for instance, any observation of users in laboratory or natural settings was coded as ethnography, whether or not the authors specifically used that term). Many articles also used more than one method or studied more than artifact, so the number of method codes exceeds the number of articles we reviewed.

Survey Instrument

The survey used a series of closed-ended questions to understand how user experience researchers conceptualize methods as scientific research or evaluation tools. The survey collected demographic information on respondents, asked them to define user experience research, and asked them to sort various methods into either research or evaluation categories and then rank the methods within that category.

We distributed the survey through email to researchers who had published a user experience article in the last two decades and used professional platforms, such as LinkedIn and email listservs (e.g. ATTW and CPTSC) in an open call for respondents. We also shared the survey with industry contacts who research user experience and asked them to share the survey widely. The survey was open for three months (December 2018 to February 2019). Of the 89 responses, only a portion were fully completed, and those 52 responses were retained for analysis. The survey instrument is found in Appendix A.

Survey Respondents: The majority of the participants were Caucasian (65%) with the remainder of the survey comprised of African American (10%), Hispanic (10%), and Asian (5%) respondents. The average age was 40 years old with most participants coming from the United States (74%). There was some international perspective with representation from Europe (17%) and at least one respondent from South America, Asia, the Middle East, and Africa.

Industry	Academia	Industry / Government	Students	Other	Total
Count	29	15	4	3	52

Gender	Men	Women	Prefer Not To Say	Total
Count	23	26	3	52

Education	BA	MA	PhD or PhD Candidate	Certificate	Total
Count	6	15	29	2	52

Table 1: Survey Demographics

Survey Analysis Process

Once the survey data was collected, it was downloaded as a .csv file for processing. All identifying information was removed in addition to the response records that were primarily incomplete (less than 98% complete). Descriptive statistics including the mean, standard deviation, and counts were prepared to understand the dimensions of the data. Due to the small number of industry and student participants, hypothesis testing was restricted to the full data set rather than looking for differences in the sub-groups (n = 52 compared to n=15 and n = 4). All textual data was dummy coded for analysis in IBM's SPSS 25 Statistical Analysis software; the accompanying charts were created using Microsoft Excel.

Results

This section includes both systematic review results and the survey results, and our research questions will be reviewed in the discussion section.

Systematic Review Results

From the systematic review of empirical user experience articles we have three sets of results: 1) The number of articles that included either a hypothesis or research question, 2) the breakdown of methods employed, and 3) the percentage of articles containing qualitative, quantitative, and mixed methods.

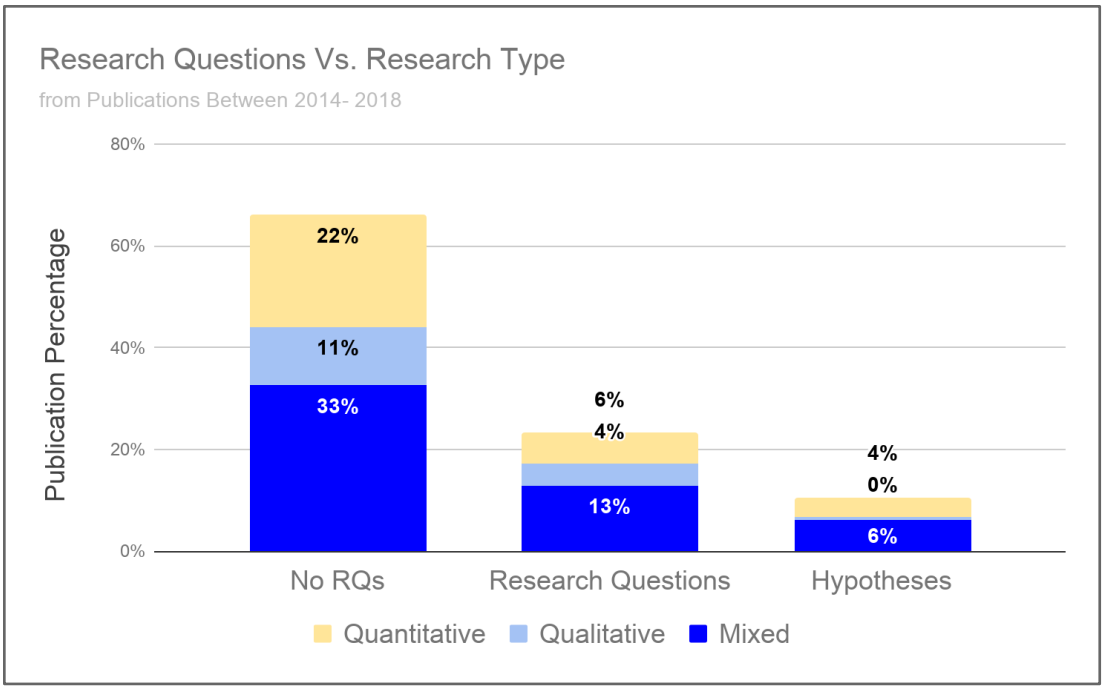

Figure 1: The percentage of articles including research questions and hypotheses.

The results showed that the majority of articles identified no formal research questions or hypothesis; we only coded research questions or hypotheses when the authors of the paper specifically identified them. Just over 20% of articles identified one or more research question, and only a small percentage of articles included at least one hypothesis.

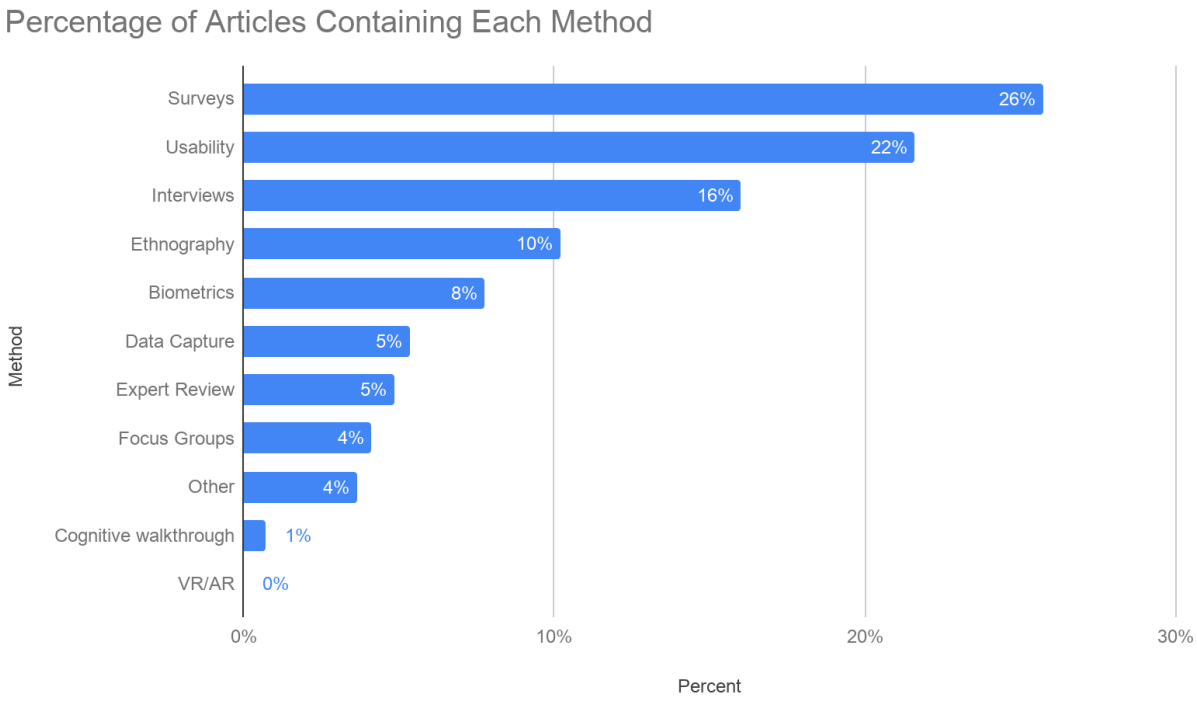

Figure 2: The percentage of articles including specific methods

Because some articles used multiple methods, the total methods deployed is larger than the number of articles. Our analysis showed that surveys were the most popular research method, appearing in 105 articles. Close behind in popularity were usability tests (n = 64), defined by our coding schema as any instance where researchers specifically include the term “usability” in their methods or any instance where researchers asked users to complete tasks with artifacts within a defined timespan. 16% of articles employed interviews (n = 54). About 10% of articles included ethnography (n = 33) or the observation of users in real-world settings. About 8% of articles (n = 26) employed biometric methods, which included a wide range of measures such as eye tracking, EEG, galvanic skin response, heart rate, and gestures.While virtual reality (VR) and artificial reality (AR) have the potential to be used as a method for either research or evaluation, no studies in our sample user virtual reality or artificial reality as a method for UX evaluation, even when those studies examined VR or AR artifacts.

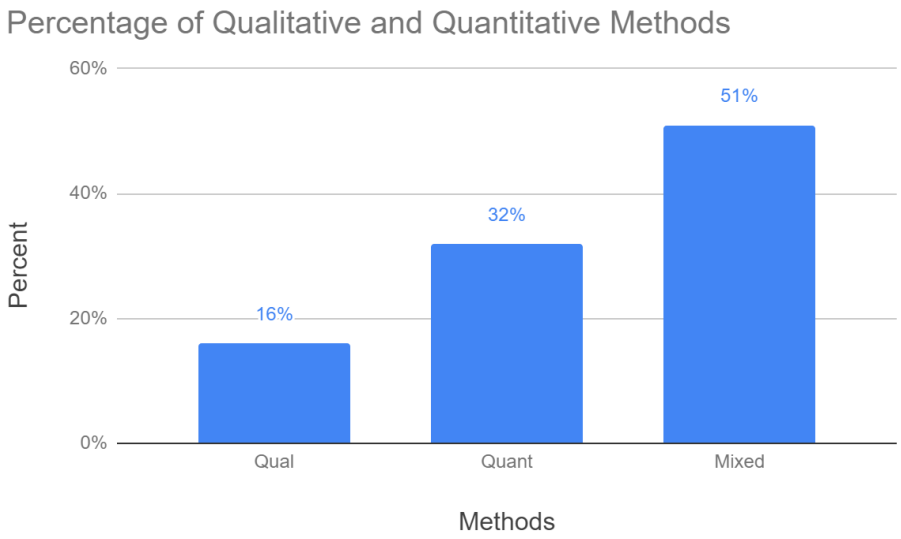

Figure 3: The percentage of articles using qualitative, quantitative, and mixed methods.

The 2016 - 2018 articles show a clear differentiation in preference between exclusively qualitative methods (n = 35) and quantitative (n = 71) methods; the largest share of published research (n = 112) used mixed methods and represents 51% of the field’s latest work.

Survey Results

The first block of questions in our survey related to how individuals would rank twelve methods after they had sorted them as either an evaluation or as a research method. Table 2 presents the results for both evaluation and research methods with their average rank, standard deviation, and distribution for the amount of times they were ranked in the top six slots. A rank of “1” is best, so the methods at the top of the list (those with the lowest averages) are preferred to those at the bottom of the list.

Research Type	Average	Standard Deviation	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
Expert Review	2.27	1.56	5	2	2	0	2	0
Usability	2.41	1.62	7	4	1	3	1	1
Data Capture	2.64	1.26	5	5	7	3	2	0
Biometric	2.96	1.57	5	7	4	3	5	1
Ethnography	2.97	1.80	11	6	7	5	2	6
Structured Interviews	3.00	1.53	7	7	9	6	3	3
Close Ended Questionnaires	3.52	1.70	3	6	6	3	4	5
Open Ended Questionnaires	3.52	1.26	1	6	4	7	7	0
Focus Group	3.57	1.75	5	1	1	8	3	3
Big Data	3.67	1.95	8	1	4	4	6	7
Narrative Interviews	3.88	1.56	1	5	6	3	6	5
Artificial Reality / Virtual Reality	4.22	1.20	0	1	1	3	3	1
Card Sorting	4.45	1.63	0	2	1	3	0	5
Diaries	4.94	1.29	0	1	2	1	5	7

Table 2: The distribution of rankings for the top six research methods.

Both usability and expert review were ranked highly as research and evaluation. Data capture and biometric methods were rated higher when described as research than when placed in the evaluation category. More qualitative methods such as ethnography and narrative interviews rose to the top when participants classified them as evaluation methods. Overall, the standard deviations reveal that evaluation methods are more consistently ranked than are research methods. Research methods, with a wider range of respondent ranks, are more controversial. Interestingly, card sorting, diaries, and artificial reality / virtual reality were in the bottom three methods regardless of whether they were described as research or evaluation by the respondent.

Evaluation Type	Average	Standard Deviation	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
Usability	1.88	1.30	18	9	4	0	2	1
Ethnography	2.50	1.69	3	2	1	0	2	0
Expert Review	2.59	1.41	10	11	10	2	5	1
Narrative Interviews	2.80	1.37	1	8	2	2	1	1
Data Capture	3.08	1.38	4	4	9	6	1	2
Big Data	3.23	1.79	3	2	2	3	1	2
Biometric	3.30	1.81	5	2	4	3	3	3
Focus Group	3.32	1.44	5	0	8	7	4	1
Open Ended Questionnaires	3.44	1.82	3	3	2	3	2	3
Structured Interviews	3.75	1.42	1	1	3	3	3	1
Close Ended Questionnaires	3.94	1.81	2	3	0	4	3	4
Artificial Reality / Virtual Reality	4.00	1.76	2	4	3	4	3	7
Diaries	4.14	1.77	0	2	1	0	2	2
Card Sorting	4.19	1.21	0	2	4	6	6	3

Table 3: The distribution of rankings for the top six evaluation methods.

Correlations

Inferential statistics were performed to determine if differences in the choice of research methods could be expected to appear across the larger user experience community based on certain attributes or if the observed differences are due to random chance. The statistical tests used are non-parametric, so they do not require the data have an underlying normal distribution. Each of the positive results have a p-value of less than .05, meaning that we are sure with 95% confidence that if the study was repeated, these differences would still be present (statistical significance). However, many of the correlation coefficients are small or moderate, making the substantive significance small. While we are confident the following patterns exist, there are other factors that impact the decisions behind which research methods are preferred that are not perfectly captured in our survey data.

By running a Spearman Correlation on how respondents rated different research methods, four distinct patterns emerged. First, open rank questionnaires are moderately, positive correlated with closed rank questionnaires (spearman correlation p = .001, rho (52) = .419). Those who use questionnaires are receptive to both open question formats (qualitative) and closed question formats (quantitative) data types. Outside of this flexibility, there was a noticeable affinity for either qualitative and quantitative methods.

On the qualitative side, ethnography had a small, positive correlation with diaries (spearman correlation p = .047, rho (52) = .257). Diaries and card sorting were moderately, positively correlated (spearman correlation p = .003, rho (52) = .380). These qualitative methods are rated more highly by individuals that value qualitative methods. The same was true for a quantitative affinity: as a person's rating for usability goes up, they rated structured interviews more poorly (spearman correlation p = .012, rho (52) = -.321). Structured interview ratings were also moderately, negatively correlated with diary ratings (spearman correlation p = .01, rho (52) = -.329). Those who value structured interviews tended to not value diaries.

The third pattern is potentially related to access to the various tools used in these methods. Artificial reality and virtual reality were moderately, negatively correlated with open questionnaires (spearman correlation p = .008, rho (52) = -.340); open questionnaires were also negatively correlated with biometric tools (spearman correlations p = .006, rho (52) = -.350). Biometric tools had a strong, positive correlation with AR/VR (spearman correlation p < .001, rho (52) = .597).

Finally, there is a pattern related to trust in the researcher’s expert opinion compared to the use of external data capture tools. As a person's expert review rating goes up, they rate biometric tools more poorly (spearman correlation p = .005, rho (52) = -.357), and the respondents that value expert review rate data capture more poorly (spearman correlations p = .017, rho (52) = -.306). Both of these correlations are moderate, negative correlations.

Tests of Difference

Academic, Industry, Government, Student Roles: By running a Kruskall-Wallis test, we found that students rank biometric methods more highly than those working in industry ($p = 0.04$, $H(3) = 8.287$). For biometric methods, the student rank mean is 18 while the academic rank mean is 26.94, government is 24.10, and industry is 38.70. All other methods did not have a statistically significant difference in how they were evaluated based on the respondent's position.

Gender: A Mann-Whitney test revealed no differences in methods ranking based on respondent's gender.

Effects of Age on Use of Methods: After performing a logistic regression, it became clear that age has a slight effect on whether respondents used a specific method in the past five years or not (a "yes" or "no" state). The r^2 is the coefficient of determination and ranges from 0 to 1. Across the four significant relationships, the coefficient of determination reveals only a slight explanatory power of age on the probability of using a particular method.

1) Usability - Younger respondents are more likely to have used usability methods within the past five years ($\square^2(1) = 4.007$, $p = .045$, Nagelkerke $r^2 = .103$).

2) Ethnography - Older respondents are more likely to have used ethnographic methods within the past five years ($\square^2(2) = 11.095$, $p = .007$, Nagelkerke $r^2 = .230$).

3) Data Capture - Younger respondents are more likely to have used data capture methods within the past five years ($\square^2(2) = 11.936$, $p = .006$, Nagelkerke $r^2 = .248$).

4) Biometrics - Younger respondents are more likely to have used biometric methods within the past five years ($\square^2(2) = 8.714$, $p = .013$, Nagelkerke $r^2 = .187$).

Effect of Percentage of UX Job Duties on Use of Methods: After performing a logistic regression, the percentage of your job that involves user experience does impact whether the respondent used a method in the past five years or not (a "yes" or "no" state).

1) Expert Review - the higher percentage of your job involves user experience, the greater likelihood you performed an expert review in the last 5 years ($\square^2(2) = 8.418$, $p = .008$, Nagelkerke $r^2 = .177$.)

2) Structured Interviews - the higher percentage of your job involves user experience, the greater likelihood you performed structured interviews in the last five years ($\square^2(2) = 6.496$, $p = .042$, Nagelkerke $r^2 = .151$).

3) Data Capture - the higher percentage of your job involves user experience, the greater likelihood you performed data capture within the last five years ($\square^2(2) = 11.936$, $p = .045$, Nagelkerke $r^2 = .248$).

4) Big Data - the higher percentage of your job that involves user experience, the greater likelihood you performed a big data project within the last five years ($\square^2$ (2) = 7.056, p = .009, Nagelkerke r^2 = .164).

5) Card Sorting - the higher percentage of your job that involves user experience, the greater likelihood you performed a card sort within the last five years ($\square^2$ (2) = 9.826, p = .002, Nagelkerke r^2 = .209).

Experience With Methods Effect on Ranking: A Mann-Whitney test showed that experience with closed ended questionnaires in the last five years results in a much higher ranking of that method (p = .003, U (n = 52) = 140). Mean rank for those who use closed ended questionnaires is 26.98, while those who do not rank them at 43.23.

Experience with structured interviews in the last five years also translated into a much higher ranking according to a mann-whitney test (p = .002, U (n = 52) = 164). Mean rank for those who use structured interviews is 26.23, while the mean rank for those who do not use them is 42.25.

No other method ranks have a relationship with use in the last five years which implies that the way UX researchers rate methods is not based on recent personal experience with the exception of the two methods: close ended questionnaires and structured interviews.

Evaluation or Method

In asking respondents to sort the available list of methods into “evaluation” or “research”, we were trying to access the context for when different methods are valued. Evaluation was defined in our survey as the “quick assessment and the improvement of a single item.” Research, on the other hand, was defined as “methods that produce results which are generalizable.”

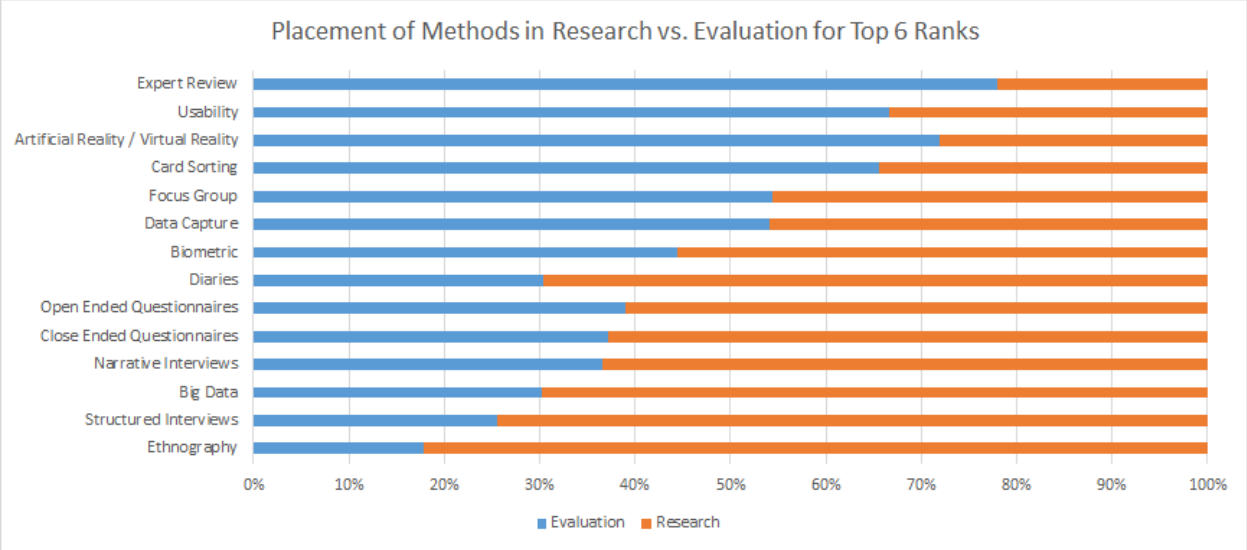

Figure 4: Where UX Respondents Sort Methods as Evaluation or Research

All respondents answered a series of questions about User Experience research methods as evaluation or research. In Figure 4, several items were more firmly established in the research category with at least 60% of respondents sorting them as research: diaries, big data, close and open ended questionnaires, narrative and structured interviews, and ethnographic study. Expert review, usability, AR/VR, and card sorting were firmly on the evaluation side.

Discussion

The discussion section analyzes our results in the order of our research question:

RQ1. What UX methods do user experience professionals prioritize in their work?
RQ2. How do researcher's attributes influence their choice of UX methods?
RQ3. How do methods differ between user evaluation and user research?

The discussion combines the results of the systematic review and survey and compares our findings to previous findings

RQ1. What UX methods do user experience professionals prioritize in their work?

Table 4 compares the top five methods identified by both survey participants for research and evaluation as well as the top five methods used by publications. Survey participants specify using different methods for evaluation and for research. In both cases, usability, ethnography, and expert review appear among the top five. When it comes publications over the last few years, the top five include surveys, usability, interviews, ethnography, and biometrics.

Rank	Research Methods (Survey Results)	Evaluation Methods (Survey Results)	Published Methods (Systematic Review)
1	Expert review	Usability	Surveys
2	Usability	Ethnography	Usability
3	Data Capture	Expert review	Interviews
4	Biometrics	Narrative Interviews	Ethnography
5	Ethnography	Data capture	Biometrics

Table 4: Top Five Methods in Survey Response and Systematic Review

Our results include both expected and unexpected findings. Unsurprisingly, usability testing appears as a highly prioritized method for survey respondents and academic researchers. This indicates that a number of researchers find usability a reliable and adaptable method (or perhaps set of methods) to answer a number of research questions. Clearly, academic researchers and our survey participants believe usability testing can yield valuable results to both generate new knowledge and evaluate existing products. However, the popularity of usability testing may also echo analysis from an analysis by St. Amant and Melocon (2018),

which found that researchers in technical communication use the term “usability study” in such varied and vague ways that it proved challenging to code consistently. The authors note that technical communication “researchers seem to make assumptions that when they use the phrase ‘usability testing’’ that all readers and other researchers will know exactly what they are talking about” (St. Amant and Meloncon, 2018, p. 16). Various disciplines and journals may be experiencing the same problem. Our survey asked users to define UX but not usability, so it is unclear whether participants define usability consistently.

One surprise is the prevalence of expert reviews as our survey participants’ top choice for a research method and third choice as an evaluation method. This represents a significant increase since a study by Vredenberg, Mao, Smith, and Carey (2002), where rankings from over 100 usability professionals resulted in “informal expert review” being ranked 11th in most popular methods and “formal heuristic evaluation” being ranked 7th. Further, only 5% of published articles in our sample used some form of expert review, a result similar to the 2.8% of publications deploying expert review in a sample from 2000-2016 (Robinson, Lanius, and Weber, 2017). Publications in our sample often combined expert review with other methods as part of a multi-phase design process, so results from expert reviews were later tested on actual users. Practitioners may use the same strategy. Still, it is surprising to see survey participants prioritize a method that does not usually involve interactions with users or the collection of user data.

Some findings from the systematic review reflect previous results from UX literature reviews. The top three published methods from 2016-2018--surveys, usability, and interviews--match the top three methods in published studies from 2000-2016--usability, surveys, and interviews--but with usability and surveys trading positions in the top slot (Robinson, Lanius, and Weber, 2017). These results also closely match the literature review of Bargas-Avila and Hornbaek (2011), which found that questionnaires, user observations, and semi-structured interviews were the three most popular UX methods in research published from 2005-2009. Additionally, Urgas et. al. (2016) reviewed 109 studies of website usability and found that most studies employ questionnaires and usability testing. Published UX research has strongly favored surveys for almost 20 years. Scapin, Senach, Trousse, and Pallot (2012) found extensive literature on UX surveys (p. 5). Law et. al. (2014) found that 100% of 58 UX studies they reviewed used questionnaires, while Vermeeren et. al. (2010) found that almost half of the UX methods they catalogued involved surveys. Our results also suggest that many published articles use some form of ethnography, a finding which reflects Robinson, Lanius, and Weber (2017) and Bargas-Avila and Hornbaek (2011). Taken together, these literature reviews suggest that published UX research continues to prioritize long-standing methods, often imported from other fields. These methods have clearly proven fruitful and adaptable for UX research, though publications less often use methods developed specifically for UX purposes.

Despite the popularity of surveys in published research, our respondents did not rank surveys highly for either research or evaluation purposes. Unfortunately, our data is not fine-grained enough to explain why surveys scored so low for this group, though this finding is consistent with data from Vredenberg, Mao, Smith, and Carey (2002). It is possible that the respondents

favored methods designed to research and evaluate specific products, while academics publishing papers want methods designed to produce more generalizable data. Additionally, published literature may rely so heavily on surveys because they provide comparatively quick data to satisfy the publishing expectations of many academic institutions and fields.

One consistent result between our survey and systematic review finding is an apparent increase in biometric data (including eye tracking, heart rate, EEG, and galvanic skin response). While only 8% of published studies included biometric measures in our sample, that represents a notable rise from 5.2% of studies using physiological data (including eye tracking) between 2000-2016 (Robinson, Lanius, and Weber) and 5% of studies between 2005-2009 (Bargas-Avila and Hornbaek, 2009). Given that respondents also prioritized biometric data, our findings suggest that UX is slowly incorporating more physiological measures into research. We speculate further on possible reasons for the rise of biometric research in our discussion of RQ2 below.

RQ2. How do researcher's attributes influence their choice of UX methods?

The survey revealed several potential trends in the community surrounding the choice of user experience methods. The first trend is the softening of the long standing tension between qualitative and quantitative methods. While it is an easy way to discuss different ideological investments in research methods, both published research and the survey results show that the qualitative-quantitative divide is fading. For example, the correlations of different method ranks shows that respondents are committed to genres of methods rather than all qualitative or all quantitative formats. The open ended and closed ended questionnaires had a moderate positive correlation that shows this commitment. Of course, the general affinity for more qualitative methods (ethnography, diaries, and card sorting) persists, but two of these correlations were slight. More than half of the published research in 2016 – 2018 contained a mixed method approach that supports this observed trend in the survey. The percentage of qualitative (16%), quantitative (32%) and mixed methods (51%) of studies in our sample tracks closely with findings from Robinson, Lanius, and Weber (2017). However, these percentages reflect an almost exact inversion from Bargas-Avila and Hornbaek's 2011 analysis, which found that 50% of studies were qualitative, 33% quantitative, and 17% mixed. Other literature reviews have also suggested that UX favors qualitative methods (Maia and Furtado, 2016). Our findings suggest that UX relies on a diversity of methods that capture different facets of the user's experience. The reliance on mixed methods may also indicate that UX is moving towards more robust research, where multiple methods help researchers confirm their findings.

An additional correlation based finding is related to emerging tools and technologies. The single strong correlational relationship found in the ranking of methods was between biometric and AR/VR, which reveals that there is both a stark commitment or a shared apathy towards newer tools. If you like biometrics, you like AR/VR; if you dislike biometrics, you will also dislike AR/VR. The basis for this relationship could be related to limited access to eye tracking, heart rate, artificial reality, or virtual reality. It might also be that these technologies are still expensive to purchase and resource heavy to use, putting them out of reach for many user experience

practitioners. It is possible that as emerging tools become more readily available (accompanied by more user friendly analytical software), members of the community will revisit their commitments and be receptive to these technology dependent methods.

Age also affects some method preference. In one of the tests of difference, students ranked biometrics more highly than did academics, industry, or government respondents. In the logistic regression tests, younger researchers were more likely to have used usability, data capture, and biometrics in the last five years, while older researchers were more likely to have used ethnography. While it is not possible to definitively say why, it is possible there has been a shift in what these different age brackets learned prior to become UX researchers. Alternatively, with age tends to come more seniority and autonomy in the workplace which might create the flexibility to choose an ethnographic method that requires more trust in the researcher. The second possible explanation is further supported by the high prioritization of expert review across the board.

The respondents generally showed confidence in their own expertise, sometimes preferring it to an external data capture / analytical tool. As the individual's work involved a higher percentage of UX tasks, they were more likely to have performed an expert review. There was also a negative correlation where those who ranked expert review highly also ranked biometric and data capture tools poorly. Additionally, expert review was highly ranked regardless of whether it was placed in the evaluation or research categories. Trust in the perception of the researcher may point to why there is a noticeable lack of research questions or hypothesis published in empirical user experience articles: fewer than 40% of articles published since 2016 contain either. Research questions and hypothesis testing are bound to the scientific method which is designed to externalize the process whereby beliefs and theories are tested; expert review is not always bound to that commitment. Experts are allowed to use intuition and experience where the scientific method requires testing and validation.

Finally, personal experience with a method in the last five years did not have an effect on how those methods were ranked with the exception of closed ended questionnaires and structured interviews. These two methodological preferences could be explained by researchers who are already comfortable with questionnaires and interviews being persuaded of the merits of more structure after seeing the results. But fascinatingly, our commitments to the value of different methods is not related to our personal practices.

RQ3. How do methods differ between user evaluation and user research?

Evaluation methods were less controversial as respondents ranked these methods similarly (smaller standard deviations). In the research category, however, respondents were much less consistent, as reflected by the higher average rank for the research category and wider standard deviations. Our findings are consistent with Gray (2016), who in an interview of user experience researchers found that "remarkably few explicit user research methods were shared by participants, with only persona/scenario and interview/focus group being mentioned" (p.

4049). There were a few trends in placement of methods as either evaluation, but they do not follow typical method divisions.

Expert review was placed in the evaluation category 78% of the time. This makes sense as expert review is the most convenient way to assess and improve the design of an artifact. Usability, likewise, was placed in the evaluation category 67% of the time, which corresponds to the origin of usability testing as the final gateway to a product or services launch. AR/VR was surprisingly sorted as evaluation 72% of the time, but it was poorly ranked. This configuration is likely due to the unknown future role of AR/VR as a method in general rather than indicating its current use in evaluation projects.

On the other side, several methods were consistently sorted into the research or generalizable knowledge side. Ethnography (82%) and diaries (79%) were almost always placed in research, with interviews, questionnaires, and big data also appearing as research more than 60% of the time. We suspect that these proportions reflect the intellectual lineage of these methods from various academic disciplines, although the nearly complete sorting of ethnography, an approach that is known for its rich, complicated, and messy nature, was placed under the umbrella of "generalizable knowledge."

Limitations and Directions for Future Research

This study attempts to provide a snapshot of contemporary UX methods used in academic and industry research, but it also contains possible limitations in both the systematic literature and the survey. Our systematic literature review process only captures articles that specifically identify as "user experience" or "UX" and are available through Google Scholar. Capturing publications with "UX" in the metadata is difficult due to the issues associated with two-letter searches in indexes that are not case sensitive (Bargas Avila, 2011; Vermeeren, 2010). It is highly likely that our review missed relevant publications that fell beyond this scope. Additionally, our review identified a few publications that were not available through our interlibrary loan service or not available in English, so those publications are not included in our review.

We distributed our survey through as many channels and contacts as possible, but because we used both convenience and snowballing samples, we missed a wide range of possible participants, especially practitioners outside of the United States. We could also not offer incentives to every participant, which likely decreased participation. Many of the survey responses were incomplete and had to be discarded. Ultimately, our survey responses provide a useful but incomplete picture of contemporary UX practice.

Our results raise many questions that would benefit from more research. Research could examine further some of the findings that emerge here, such as the preference for expert review among our survey respondents or the definitions of usability testing held by both practitioners and academic researchers. Similar research could attempt to explore or resolve some of the controversies that appeared to emerge about some methods. Future research could specifically examine how the broad rankings of UX methods align with preferences of technical

communicators, both in academic publications and industry practice. Research could also look at how academic and industry research conduct research at various phases of the product development cycle and which methods they prefer for early and late-stage evaluation of products.

Conclusion

This project provides a glimpse of UX methods deployed in technical communication and beyond. In collecting this data, we hope to shed light on the strengths and weaknesses of current UX research. This project responds to the call from Scapin et. al. (2012) that “It is very important for well grounded UX future that the nature of data categories to gather is well defined, structured, and coordinated, together with improved methods” (Scapin et. al., 2012, p. 6). We also take seriously the charge from St. Amant and Meloncon (2018) to establish more sustainable research methods.

Our results suggests reasons for enthusiasm, concern, and future questions. The prioritization of mixed methods approaches could suggest that UX is becoming more rigorous and working to determine which methods best capture different facets of the user experience. The prominence of usability testing could suggest that this method proves effective for both research and evaluation; however, our finding could also suggest an over-reliance on usability testing due to vague definitions of usability, limited definitions of user experience, or a tendency to fall back on established methods. A modest rise in biometrics suggests that the field is exploring the effectiveness of physiological measures for understanding user experience, though this rise will require discussion about the usefulness, appropriateness, and ethics that accompany these measures. But perhaps the quickest and most blatant improvement to UX research would be a field-wide expectation, regardless of discipline or journal, that generalizable UX research be guided by distinct questions or hypotheses. These and other steps will help the multi-discplinary field of UX develop consistent, productive research methods.

References

Bargas-Avila, J. A., & Hornbæk, K. (2011). Old wine in new bottles or novel challenges: A critical analysis of empirical studies of user experience, In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 93–102). New York: ACM. http://doi. org/10.1145/1978942.1979336

Chilana, P.K., Ko, A.J., Wobbrock, J.O., Grossman, T., & Fitzmaurice, G. (2011). Post-deployment usability: A survey of current practices. In *CHI '11 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 2243-2246). DOI: 10.1145/1978942.1979270

Cooke, L. (2010). Assessing concurrent think-aloud protocol as a usability test metho’d: A technical communication approach. *IEEE Transactions on Professional Communication*, 53(3), 202-215.

de Sá, M., & Carriço, L. (2011). Designing and evaluating mobile interaction: Challenges and trends. *Foundations and Trends in Human–Computer Interaction*, 4(3), 175-243. http://dx.doi.org/10.1561/1100000025

Dove, G., Halskov, K., Forlizzi, J., & Zimmerman, J. (2017). UX design innovation: Challenges for working with machine learning as a design material. In *CHI '17 Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, (pp. 278-288). DOI: 10.1145/3025453.3025739

Forlizzi, J., & Battarbee, K. (2004). Understanding experience in interactive systems. *DIS '04 Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*, 261-268.

Gehanno, J.-F., Rollin, L., & Darmoni, S. (2013). Is the coverage of Google Scholar enough to be used alone for systematic reviews. *BMC Medical Informatics and Decision Making*, 13(7), 1–5. http://doi.org/10.1186/1472-6947-13-7

Gerea, C., & Herskovic, V. (2015). Measuring user experience in Latin America: An exploratory survey. In *Proceedings of the Latin American Conference on Human Computer Interaction* (pp. 19–4). New York: ACM. http://doi.org/10.1145/2824893.2824914

González-Pérez, L.I., Ramirez-Montoya, M., & García-Peñalvo, F.J. (2018). User experience in institutional repositories: A systematic literature review. *International Journal of Human Capital and Information Technology Professionals*, 9(1), 70-86. DOI: 10.4018/IJHCITP.2018010105

Gray, C.M. (2016). "It's more of a mindset than a method": UX practitioners' conception of design methods. In *CHI '16 Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, DOI: 10.1145/2858036.2858410

Gross, A., & Bongartz, S. (2012). Why do I like it? Investigating the product-specificity of user experience. *NordiCHI* '12, 322-330.

Hao Y., Chong, W., Man, K.L., Liu O., & Shi, X. (2016). Key factors affecting user experience of mobile crowdsourcing applications. In *Proceedings of the International MultiConference of Engineers and Computer Scientists 2016 Vol II*, IMECS 2016, March 16 - 18, 2016, Hong Kong. (pp. 1-6).

Hassenzahl M. (2018). The Thing and I: Understanding the Relationship Between User and Product. In: Blythe M., Monk A. (eds) Funology 2. Human–Computer Interaction Series. Springer, Cham. (301-313). pp. DOI: https://doi.org/10.1007/978-3-319-68213-6_19

Hassenzahl, M., & Tractinsky, N. (2006). User experience - a research agenda. *Behaviour & Information Technology*, 25(2), 91–97. http://doi.org/10.1080/01449290500330331

Hooper, C.J., & Dix, A. (2012). Web science and human-computer interaction: When disciplines collide. In *WebSci '12 Proceedings of the 4th Annual ACM Web Science Conference*, Evanston, IL. (pp. 128-136). DOI: 10.1145/2380718.2380736.

Ibargoyen, A., Szostak, D., Bojic, M. (2013). The elephant in the conference room: Let's talk about experience technology. *CHI 2013 Extended Abstracts*, pp. 2079-2088.

Instone, K. (2005). User experience: An umbrella topic. *CHI 2005*, 1-3.

Ismail, N.A., Ahmad,F., Kamaruddin, N.A., Ibrahim, R. (2016). A review on usability issues in mobile applications. *IOSR Journal of Mobile Computing & Application*, 3(3), 47-52. DOI: 10.9790/0050-03034752.

Jääskeläinen, A., & Heikkinen, K. (2010). Divergence of user experience: Professional vs. end users. In *Proceedings of I-UxSED 2010*, 30.

Kaasinen, E., Roto, V., Hakulinen, J., Heimonen, T., Jokinen, J.P.P., Karvonen, H., Keskinen, T., Koskinen, H., Lu, Y., Saariluoma, P., Tokkonen, H., & Turunen, M. (2015). Defining user experience goals to guide the design of industrial systems. *Behaviour & Information Technology*, 34:10, 976-991, DOI: 10.1080/0144929X.2015.1035335

Kastman, L.M., Zachry, M., & Spinuzzi, C. (2001). Usability instruction in technical communication programs: New directions in curriculum development. *Journal of Business and Technical Communication*, 15(2), 223-240.

Khabsa, M., & Giles, C. L. (2014). The number of scholarly documents on the public web. *PLoS ONE*, 9(5), e93949–6. http://doi.org/10.1371/journal.pone.0093949

Kim, M.J., Oh, M.W., Cho, M.E., Lee, H., & Kim, J.T. (2013). A critical review of user studies on healthy smart homes. *Indoor and Built Environment*, 22(1), 1-11. DOI: 10.1177/1420326X12469733

Kim, S.J.J. (2012). User study trends in augmented reality. In *Proceedings of 2012 International Symposium on Ubiquitous Virtual Reality*, pp. 1-5. DOI 10.1109/ISUVR.2012.17

Kjeldskov, J., & Graham, C. (2003). A review of mobile HCI research methods. In *Mobile HCI 2003: Human-Computer Interaction with Mobile Devices and Services*, (pp. 317-335).

Kurosu, M., Kobayashi, T., Yoshitake, R., Takahashi, H., Urokohara, H., & Sato, D. (2004). Trends in usability research and activities in Japan. *International Journal of Human-Computer Interaction*, 17(1), 103-124. https://doi.org/10.1207/s15327590ijhc1701_8

Lachner, F., Naegelein, P., Kowalski, R., Spann, M., & Butz, A. (2016). Quantified UX: Towards a common organizational understanding of user experience. In *NordiCHI '16*, DOI: 10.1145/2971485.2971501.

Lallemand, C., Gronier, G., & Koenig, V. (2015). User experience: A concept without consensus? Exploring practitioners' perspectives through an international survey. *Computers in Human Behavior*, 43, 35–48. http://doi.org/10.1016/j. chb.2014.10.048

Lauer, B., & Brumberger, E. (2015). Technical communication as user experience in a broadening industry landscape. *Technical Communication*, 63(3), 248-264.

Law, E., van Schaik, P., & Roto, V. (2014). Attitudes towards user experience (UX) measurement. *International Journal of Human-Computer Studies*, 72(6), 526–541. http://doi. org/10.1016/j.ijhcs.2013.09.006

MacDonald, C.M. (2015). User experience librarians: User advocates, user researchers, usability evaluators, or all of the above? In *ASIST '15 Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community*.

Maia C.L.B., Furtado E.S. (2016) A Systematic Review About User Experience Evaluation. In: Marcus A. (eds) *Design, User Experience, and Usability: Design Thinking and Methods. DUXU 2016. Lecture Notes in Computer Science*, vol 9746. Springer, Cham. pp. 445-455.

McNely, B., Spinuzzi, C., & Teston, C. (2015). Contemporary research methodologies in technical communication. *Technical Communication Quarterly*, 24(1), 1-13.

Melocon, L., & St. Amant, K. (2018). Empirical research in technical and professional communication: A 5-year examination of research methods and a call for research sustainability. *Journal of Technical Writing and Communication*, 48(1), 1-28.

Mkpojiogu, E.O.C., & Asuquo, A.E. (2018). User experience of ATM users in Nigeria: A systematic review of empirical papers. *JORIND*, 16(1), 115-125.

Mao, J., Vredenburg, K., Smith, P.W. & Carey, T. (2005). The state of user-centered design practice. Communications of the ACM, 48(3), 105-109.

Nie, B., & Sun, S. (2017). Using text mining techniques to identify research trends: A case Study of design research. *Applied Sciences*, 7(4), 401-423. DOI:10.3390/app7040401

Rajanen, M., & Tapani, J. (2018). A survey of game usability practices in North American game companies. In B. Andersson, B. Johansson, S. Carlsson, C. Barry, M. Lang, H. Linger, & C. Schneider (Eds.), D*esigning Digitalization (ISD2018 Proceedings)*. Lund, Sweden:

Lund University. ISBN: 978-91-7753-876-9. http://aisel.aisnet.org/isd2014/proceedings2018/HCI/1.

Rajanen, D., Clemmensen, T., Iivari, N., Inal, Y., Rızvanoğlu, K., Sivaji, A., Roche, A. (2017). UX professionals' definitions of usability and UX – A comparison between Turkey, Finland, Denmark, France and Malaysia. In: Bernhaupt R., Dalvi G., Joshi A., K. Balkrishan D., O'Neill J., Winckler M. (eds) *Human-Computer Interaction – INTERACT 2017*. INTERACT 2017. Lecture Notes in Computer Science, vol 10516. (pp. 218-239). Springer, Cham. DOI: https://doi.org/10.1007/978-3-319-68059-0_14

Redish, G., & Barnum, C. (2011). Overlap, influence, intertwining: The interplay of UX and technical communication. *Journal of Usability Studies*, 6(3), 90-101.

Robinson, J., & Lanius, C. (2018). A geographic and disciplinary examination of UX empirical research since 2000. In *SIGDOC '18 Proceedings of the 36th ACM International Conference on the Design of Communication*, Milwaukie, WI. DOI: 10.1145/3233756.3233930

Robinson, J., Lanius, C., & Weber, R. (2017). The past, present, and future of UX empirical research. *Communication Design Quarterly*, 5(3), 10-22. DOI: 10.1145/3188173.3188175.

Roto, V., Vermeeren, A., Väänänen-Vainio-Mattila, K., & Law, E. (2011b). User experience evaluation – which method to choose? *Lecture Notes in Computer Science* (pp. 714–715). http://doi.org/10.1007/978-3-642-23768-3_129

Rude, C. D. (2009). Mapping the research questions in technical communication. *Journal of Business and Technical Communication*, 23(2), 174–215. http://doi.org/10.1177/1050651908329562

Salgado, A.D.M., Amaral, L.A., Freire, A.P., Fortes, R.P.M. (2016). Usability and UX practices in small enterprises: Lessons from a survey of the Brazilian context. In *SIGDOC '16 Proceedings of the 34th ACM International Conference on the Design of Communication*, DOI: 10.1145/2987592.2987616.

Scapin, D. L., Senach, B., Trousse, B., & Pallot, M. (2012). User experience: Buzzword or new paradigm? *In Proceedings of the ACHI Fifth International Conference on Advances in Computer-Human Interactions*. Retrieved, from https://wwwsop.inria.fr/axis/pages/bestpaper/dlsPaper20153ACHI2012. pdf

Sullivan, P. (1989). Beyond a narrow conception of usability testing. *IEEE Transactions on Professional Communication*, 32(4), 256-264.

Ugras T., Gülseçen S., Çubukçu C., İli Erdoğmuş İ., Gashi V., Bedir M. (2016) Research Trends in Web Site Usability: A Systematic Review. In: Marcus A. (eds) *Design, User Experience, and Usability: Design Thinking and Methods. DUXU 2016. Lecture Notes in Computer Science*, vol 9746. Springer, Cham. pp. 517-528.

Vermeeren, A., Law, E., Roto, V., Obrist, M., Hoonhout, J., & Väänänen-Vainio-Mattila, K. (2010). User experience evaluation methods: Current state and development needs. In *Proceedings of the 6th Nordic Conference on Human Computer Interaction: Extending Boundaries* (pp. 521–530). New York: ACM. http://doi.org/10.1145/1868914.1868973

Vredenburg, K., Mao, J., Smith, P.W., & Carey, T. (2002). A survey of user-centered design practice. In *CHI '02 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 471-478). New York: ACM. DOI: 10.1145/503376.503460

Williams, S.D. (2007). User experience design for technical communication: Expanding our notions of quality information design. In *2007 IEEE International Professional Communication Conference*, (pp. 1-13). Seattle, WA: IEEE.

Yusop, N.S.M., Grundy, J., & Vasa, R. (2017). Reporting usability defects: A systematic literature review. *IEEE Transactions on Software Engineering*, 43(9), 848-867. DOI: 10.1109/TSE.2016.2638427.

Appendix

Appendix A: Survey - The Politics and Status of UX Research Methods

1. Please select the definition you think best describes User Experience:
2. In the last 5 years, have you used any of the following methods? (Please select all that apply.)
    - Usability
    - Expert Review (heuristic evaluation)
    - Ethnographic Study (contextual inquiry/field studies)
    - Diaries
    - Focus Groups
    - Open-ended Questionnaires
    - Closed end Questionnaires
    - Narrative Interviews
    - Structured Interviews
    - Screen Capture (mouse movement, click through)
    - Physiological (eye tracking, heart rate, galvanic skin response)
    - Big Data
    - Artificial Reality / Virtual Reality (simulations)
    - Card Sorting
    - Other _____
3. What are your top three preferred methods in general?
4. For the following question, please assume that a research method produces results which are generalizable, while evaluation methods are for quick assessment and the improvement of a

single item.
Please place and rank order the following methods where they belong in the field of User Experience according to the two categories: evaluation and research. The ranking should reflect how useful the methods are, with the higher rank having the greatest utility and the lowest rank being least useful.

- Usability
- Expert Review (heuristic evaluation)
- Ethnographic Study (e.g., contextual inquiry/field studies)
- Focus Group
- Open-ended Questionnaires
- Closed End Questionnaires
- Structured Interviews
- Narrative Interviews
- Screen Capture (mouse movement, click through)
- Physiological (eye tracking, heart rate, galvanic skin response)
- Big Data
- Card Sorting
- Diaries
- Artificial Reality / Virtual Reality (simulations)
- Other

5. Please review the following description and rank the methods in the order for which you would deploy them.
   You are a user experience researcher who is charged with evaluation of a new e-commerce website. Your customer "Blurp Products" is very early in the design process, and they need to get a broad sense of their users.
   - Usability
   - Expert Review (heuristic evaluation)
   - Ethnographic Study (e.g., contextual inquiry/field studies)
   - Focus Group
   - Open-ended Questionnaires
   - Closed End Questionnaires
   - Structured Interviews
   - Narrative Interviews
   - Screen Capture (mouse movement, click through)
   - Physiological (eye tracking, heart rate, galvanic skin response)
   - Big Data
   - Card Sorting
   - Diaries
   - Artificial Reality / Virtual Reality (simulations)
   - Other
6. Please review the following description and rank the methods in the order for which you would deploy them.
   You are a user experience researcher who is charged with the evaluation of a new e-commerce website. Your customer "Bloop Products" has a complete product ready to go to market, but one member of their team has convinced the product manager to do a user experience review prior to launch.
   - Usability
   - Expert Review (heuristic evaluation)
   - Ethnographic Study (e.g., contextual inquiry/field studies)
   - Focus Group
   - Open-ended Questionnaires
   - Closed End Questionnaires
   - Structured Interviews
   - Narrative Interviews
   - Screen Capture (mouse movement, click through)

    - Physiological (eye tracking, heart rate, galvanic skin response)
    - Big Data
    - Card Sorting
    - Diaries
    - Artificial Reality / Virtual Reality (simulations)
    - Other
7. What percentage of your job duties are related to usability or user experience?
8. What is your primary and current employment status?
9. What is your gender?
10. What is your racial identity?
11. What year were you born?
12. In which country do you currently reside?
13. In which state do you currently reside?
14. What is your highest level of education you have completed or the highest degree you have received?
15. If you have completed an associate, bachelors, masters, professional school, or doctorate academic degree, please indicate the subject area(s) of ALL your degree(s).
16. In the last 5 years, have you shared any of following methods informally or formally with your colleagues? (Please select all that apply.)
    - Usability
    - Expert Review (heuristic evaluation)
    - Ethnographic Study (e.g., contextual inquiry/field studies)
    - Focus Group
    - Open-ended Questionnaires
    - Closed End Questionnaires
    - Structured Interviews
    - Narrative Interviews
    - Screen Capture (mouse movement, click through)
    - Physiological (eye tracking, heart rate, galvanic skin response)
    - Big Data
    - Card Sorting
    - Diaries
    - Artificial Reality / Virtual Reality (simulations)
    - Other
17. Where are you employed?
18. Which of the following industries most closely matches the one in which you are employed?
19. How many years have you been working in your field or industry?
20. Which program or department do you teach in?
21. How would you describe the academic institution where you teach?
22. In the last 5 years, have you taught any of the following methods in the classroom? (Please select all that apply.)
    - Usability
    - Expert Review (heuristic evaluation)
    - Ethnographic Study (e.g., contextual inquiry/field studies)
    - Focus Group
    - Open-ended Questionnaires
    - Closed End Questionnaires
    - Structured Interviews
    - Narrative Interviews
    - Screen Capture (mouse movement, click through)
    - Physiological (eye tracking, heart rate, galvanic skin response)
    - Big Data
    - Card Sorting
    - Diaries
    - Artificial Reality / Virtual Reality (simulations)
    - Other

23. Please indicate the number of years you have been a teacher for each category:
24. Number of years I have been teaching
25. Please list the class number and name for courses where you teach user experience content (e.g., EH301 Technical Writing, EH445 Usability Methods).
26. Please select the course level(s) you think is appropriate to teach usability and/or user experience content:
27. Please, explain your choice of course level made above.
28. What type of academic institution are you currently attending?
29. What program are you currently enrolled in?
30. What degree are you seeking?
31. Please list the courses where you have encountered usability or user experience content (e.g., Technical Writing, Usability Methods).

Appendix B: Codebook - Deductive Coding

Grouping	Category	Code	Description
1. Research type: as reported by the researcher.			
	Qual	T-Qual	
	Quant	T-Quant	
	Mixed	T-Mixed	All usability should be tagged as mixed
		other	I need another opinion [Selectable (checkbox)]
	Unclear	T-Unclear	
2. Any research questions provided?			
	RQ1	R-R1	Code the text
	RQ2	R-R2	
	RQ3	R-R3	
	RQ +	R-R+	
	None	R-None	
	Hypotheses	hypo	Only if hypothesis
3. Methods (all that apply)			
	Expert review	Expert	

Individual interaction	Ethno	Ethno	
	interviews	Int	
	eyetracking	Eye	
	Screen capture/data	Screen	
	physiological	Physio	
	usability	Usability	
	VR/AR	VR	
	Big data	bigdata	
Group	Focus group	Focus	
	Surveys	Survey	
	Other	Other	
	unknown	Unknown	
4. Artifacts of study (check all that apply)			
Hardware	Mobile		
	Computer		
	Game Console		
	Tablet		
	Interface device		
	VR/AR		

	HARDOther		
Software	Web app		
	Software (local)		
	Website		
	Video/MM		
	Game		
	Objects		
	Imagined object/ prototype		
	OTHER		
UX			code was used to indicate the interest in UX issues (e.g., SUX instruments, UX education)
Attitude			the designers concern about a concept. For example, researchers might want to know how the users felt about an upcoming implementation of Agile or virtual realit
Process			Process" referred to the observation of a method (e.g., search rankings, implementations).
Service			
Other			
5. Details (check all that apply)			
Population	Age (Older adults or children)		
	Race/Color		
	Religion		

	Sex/gender/ Orientation		
	National Origin		
	Adults (general)		
	occupation		
	Disability		
	Unspecified		
	Other		

6. Find the following items...			
	Sample size		
	Keywords		
	Study duration		
	FULL Journal name		
	Check year		